Bernoulli **22**(4), 2016, 2237–2300
DOI: 10.3150/15-BEJ728

A stochastic algorithm finding p-means on the circle

MARC ARNAUDON[1] and LAURENT MICLO[2]

[1]*Institut de Mathématique de Bordeaux, UMR 5251, Université de Bordeaux and CNRS, 351, Cours de la Libération, F-33405 TALENCE Cedex, France. E-mail: marc.arnaudon@u-bordeaux1.fr*
[2]*Institut de Mathématiques de Toulouse, UMR 5219, Université Toulouse 3 and CNRS, 118, route de Narbonne, 31062 Toulouse Cedex 9, France. E-mail: miclo@math.univ-toulouse.fr*

A stochastic algorithm is proposed, finding some elements from the set of intrinsic p-mean(s) associated to a probability measure ν on a compact Riemannian manifold and to $p \in [1, \infty)$. It is fed sequentially with independent random variables $(Y_n)_{n\in\mathbb{N}}$ distributed according to ν, which is often the only available knowledge of ν. Furthermore, the algorithm is easy to implement, because it evolves like a Brownian motion between the random times when it jumps in direction of one of the Y_n, $n \in \mathbb{N}$. Its principle is based on simulated annealing and homogenization, so that temperature and approximations schemes must be tuned up (plus a regularizing scheme if ν does not admit a Hölderian density). The analysis of the convergence is restricted to the case where the state space is a circle. In its principle, the proof relies on the investigation of the evolution of a time-inhomogeneous $\mathbb{L}^2$ functional and on the corresponding spectral gap estimates due to Holley, Kusuoka and Stroock. But it requires new estimates on the discrepancies between the unknown instantaneous invariant measures and some convenient Gibbs measures.

Keywords: Gibbs measures; homogenization; instantaneous invariant measures; intrinsic p-means; probability measures on compact Riemannian manifolds; simulated annealing; spectral gap at small temperature; stochastic algorithms

1. Introduction

The purpose of this paper is to present a stochastic algorithm finding some of the geometric p-means of probability measures defined on compact Riemannian manifolds, for $p \in [1, \infty)$. Its convergence is analyzed in the restricted case of the circle, as a first step toward a more general result which is conjectured to be true.

1.1. The general notion of p-means

The concepts of mean and median are well understood for real valued random variables. They can be extended to random variables taking values in metric spaces in the following way. Let be given ν a probability measure on a metric space M, whose distance is denoted d. For $p \geq 1$, consider the continuous mapping

$$U_p : M \ni x \mapsto \int d^p(x, y)\nu(dy). \tag{1.1}$$

1350-7265

A global minimum of U_p is called a p-mean of ν, at least if this function is not identically equal to $+\infty$ (equivalently, if all its values are finite, as it can be easily deduced from the triangle inequality). The set of p-means will be designated by $\mathcal{M}_p$, it is non-empty as soon as U_p goes to infinity at infinity (in the Alexandroff sense), but in general it is not reduced to a singleton. The notion of intrinsic mean and median correspond, respectively, to $p = 2$ and $p = 1$. If M is $\mathbb{R}$ endowed with its absolute value, one recovers the usual mean and distance.

These extensions are justified by the increasing number of available graph or manifold valued data samples in various scientific fields. Examples of manifold valued data samples are given by sets of parameters for families of laws endowed with Fisher information metric, by Lie groups (rotations, displacements) in control theory, by symmetric spaces in imaging or signal processing.

For some applications (see, e.g., [26]), it may be important to find $\mathcal{M}_p$ or at least some of its elements. In practice, the knowledge of ν is often given by a finite sequence $Y := (Y_n)_{n\in\{1,2,\dots,N\}}$ of independent random variables, identically distributed according to ν. Since $N \in \mathbb{N}$ is in general large enough, we will consider the limit situation where we have at our disposal an infinite sequence $Y := (Y_n)_{n\in\mathbb{N}}$. One is then looking for algorithms using this data and enabling to find some elements of $\mathcal{M}_p$. In this paper, we will be mainly interested in the case where M is the circle, even if the proposed stochastic algorithm can be considered more generally for compact Riemannian manifolds.

Algorithms for finding p-means or minimax centers have been investigated in [1,2,5–8,12, 13,20,27,28]. When possible a gradient descent algorithm is used. When the gradient of the functional to minimize is difficult or impossible to compute, a Robbins Monro-type algorithm is preferred. Either the functional to minimize has only one local minimum which is also global, or (Bonnabel [7]) a local minimum is seeked. The case of Karcher means in the circle is treated in [10] and [15]. In this special situation, the global minimum of the functional can be found by explicit formula.

For generalized means on compact manifolds, the situation is different since the functional (1.1) to minimize may have many local minima, and no explicit formula for a global minimum can be expected.

1.2. The case of the circle

In this subsection, we consider the case where M is the circle $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$ endowed with its natural angular distance d. As above, let $Y := (Y_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables distributed according to a fixed probability measure ν on $\mathbb{T}$. Let $p \in [1, +\infty)$ be fixed, we present now a stochastic algorithm finding some elements of $\mathcal{M}_p$ by using this data. It is based on simulated annealing and homogenization procedures. Thus, we will need, respectively, an inverse temperature evolution $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ and an inverse speed up evolution $\alpha : \mathbb{R}_+ \to \mathbb{R}_+^*$, where $\mathbb{R}_+^*$ stands for the set of positive real numbers. Typically, they are, respectively, non-decreasing and non-increasing and we have $\lim_{t\to+\infty} \beta_t = +\infty$ and $\lim_{t\to+\infty} \alpha_t = 0$, but we are looking for more precise conditions so that the stochastic algorithm we describe below finds $\mathcal{M}_p$ (namely, some elements from this set).

Let $N := (N_t)_{t\geq 0}$ be a standard Poisson process: it starts at 0 at time 0 and has jumps of length 1 whose inter-arrival times are independent and distributed according to exponential random variables of parameter 1. The process N is assumed to be independent from the sequence Y.

We define the speeded-up process $N^{(\alpha)} := (N_t^{(\alpha)})_{t\geq 0}$ via

$$\forall t \geq 0, \qquad N_t^{(\alpha)} := N_{\int_0^t 1/\alpha_s\, ds}. \tag{1.2}$$

Consider the time-inhomogeneous Markov process $X := (X_t)_{t\geq 0}$ which evolves in M in the following heuristic way: if $T > 0$ is a jump time of $N^{(\alpha)}$, then X jumps at the same time, from X_{T-} to X_T which is obtained by following the shortest geodesic leading from X_{T-} to $Y_{N_T^{(\alpha)}}$ at speed 1 during the time $(p/2)\beta_T\alpha_T d^{p-1}(X_{T\_}, Y_{N_T^{(\alpha)}})$. Almost surely, the above shortest geodesic is unique and there is no problem with its choice. Indeed, by the end of the description below, $X_{T\_}$ will be independent of $Y_{N_T^{(\alpha)}}$ and the law of $X_{T\_}$ will be absolutely continuous with respect to the Lebesgue measure λ on $\mathbb{T}$ renormalized into a probability measure. It ensures that almost surely, $Y_{N_T^{(\alpha)}}$ is not the opposite point of $X_{T\_}$ on $\mathbb{T}$. The schemes α and β will satisfy $\lim_{t\to+\infty}\alpha_t\beta_t = 0$, so that for sufficiently large jump-times T, X_T will be between X_{T-} and $Y_{N_T^{(\alpha)}}$ on the above geodesic and quite close to X_{T-}.

To proceed with the construction, we require that between consecutive jump times (and between time 0 and the first jump time), X evolves as a Brownian motion on $\mathbb{T}$ and independently of Y and N. Very informally, the evolution of the algorithm X can be summarized by the equation

$$\forall t \geq 0, \qquad dX_t = dB_t + (p/2)\alpha_t\beta_t d^{p-1}(X_{T\_}, Y_{N_T^{(\alpha)}})\sigma(X_{t-}, Y_{N_t^{(\alpha)}})\, dN_t^{(\alpha)},$$

where $(B_t)_{t\geq 0}$ is a Brownian motion on $\mathbb{T}$ and where $\sigma(X_{t-}, Y_{N_t^{(\alpha)}})$ is 1 (resp., -1) if the shortest way from X_{t-} to $Y_{N_t^{(\alpha)}}$ goes in the anti-clock wise (resp., the clock-wise) direction, in the usual representation of $\mathbb{R}/(2\pi\mathbb{Z})$ in $\mathbb{C}$. In the above equation, $(Y_{N_t^{(\alpha)}})_{t\geq 0}$ should be interpreted as a fast auxiliary process. The law of X is then entirely determined by the initial distribution $m_0 = \mathcal{L}(X_0)$. More generally at any time $t \geq 0$, denote by m_t the law of X_t.

The first main result of this paper states that at least if ν is sufficiently regular, the above algorithm X finds in probability at large times the set $\mathcal{M}_p$ of p-means:

Theorem 1.1. *Assume that ν admits a density with respect to λ and that this density is Hölder continuous with exponent $a \in (0, 1]$. Then there exist two constants $a_p > 0$, depending on $p \geq 1$ and a, and $b_p \geq 0$, depending on p, such that for any scheme of the form*

$$\forall t \geq 0, \qquad \begin{cases} \alpha_t := (1+t)^{-1/a_p}, \\ \beta_t := b^{-1}\ln(1+t), \end{cases} \tag{1.3}$$

where $b > b_p$, we have for any neighborhood $\mathcal{N}$ of $\mathcal{M}_p$ and for any m_0,

$$\lim_{t\to+\infty} \mathbb{P}[X_t \in \mathcal{N}] = 1. \tag{1.4}$$

Thus, to find an element of $\mathcal{M}_p$ with an important probability, one should pick up the value of X_t for sufficiently large times t.

The constant a_p is the simplest to define, since it is given by

$$a(p) := \begin{cases} a, & \text{if } p = 1 \text{ or } p \geq 2, \\ \min(a, p-1), & \text{if } p \in (1,2). \end{cases} \tag{1.5}$$

The constant $b_p \geq 0$ comes from the theory of simulated annealing (see, e.g., [14]), which will be recalled in next section. For the moment being, we just describe the constant b_p, in the setting of a compact Riemannian manifold M, since there is no extra difficulty and we will need it later on to express a conjecture extending Theorem 1.1. For any $x, y \in M$, let $\mathcal{C}_{x,y}$ be the set of continuous paths $C := (C(t))_{0 \leq t \leq 1}$ going from $C(0) = x$ to $C(1) = y$. The elevation $U_p(C)$ of such a path C relatively to U_p is defined by

$$U_p(C) := \max_{t \in [0,1]} U_p\big(C(t)\big)$$

and the minimal elevation $U_p(x, y)$ between x and y is given by

$$U_p(x, y) := \min_{C \in \mathcal{C}_{x,y}} U_p(C).$$

Then we consider

$$b(U_p) := \max_{x,y \in M} U_p(x, y) - U_p(x) - U_p(y) + \min_M U_p. \tag{1.6}$$

This constant can also be seen as the largest depth of a well not containing a fixed global minimum of U_p. Namely, if $x_0 \in \mathcal{M}_p$, then it is not difficult to see that

$$b(U_p) = \max_{y \in M} U_p(x_0, y) - U_p(y), \tag{1.7}$$

independently of the choice of $x_0 \in \mathcal{M}_p$ (cf. [14]).

Let us now describe a stochastic algorithm, derived from the previous one, which enables one to find some of the p-means of any probability measure ν on $\mathbb{T}$.

For any $x \in \mathbb{T}$ and $\kappa > 0$, consider the probability measure $K_{x,\kappa}$ whose density with respect to the Lebesgue measure $\lambda(dy)$ is proportional to $(1 - \kappa \|y - x\|)_+$. Assume next that we are given an evolution $\kappa : \mathbb{R}_+ \ni t \mapsto \kappa_t \in \mathbb{R}_+^*$ and consider the process $Z := (Z_t)_{t \geq 0}$ evolving similarly to $(X_t)_{t \geq 0}$, except that at the jump times T of $N^{(\alpha)}$, the target $Y_{N_T^{(\alpha)}}$ is replaced by a point W_T sampled from $K_{Y_{N_T^{(\alpha)}}, \kappa_T}$, independently from the other variables.

Theorem 1.2. *Let ν be an arbitrary probability measure on $M = \mathbb{T}$. For $p = 2$, consider the schemes*

$$\forall t \geq 0, \qquad \begin{cases} \alpha_t := (1+t)^{-c}, \\ \beta_t := b^{-1} \ln(1+t), \\ \kappa_t := (1+t)^k, \end{cases}$$

with $b > b(U_2)$, $k > 0$ and $c \geq 2k + 1$. Then, for any neighborhood $\mathcal{N}$ of $\mathcal{M}_2$ and for any initial distribution $\mathcal{L}(Z_0)$, we get

$$\lim_{t \to +\infty} \mathbb{P}[Z_t \in \mathcal{N}] = 1,$$

where $\mathbb{P}$ stands for the underlying probability.

More generally, for any given $p \geq 1$, it is possible to find similar schemes (where c depends furthermore on $p \geq 1$) enabling to find the set of p-means $\mathcal{M}_p$ (see Remark 5.2). Even if ν satisfies the condition of Theorem 1.1, it could be more advantageous to consider the alternative algorithm Z instead of X when the exponent a in (1.3) is too small.

Remark 1.1. The schemes α, β and κ presented above are simple examples of admissible evolutions; they could be replaced, for instance, by

$$\forall t \geq 0, \qquad \begin{cases} \alpha_t := C_1(r_1 + t)^{-c}, \\ \beta_t := b^{-1} \ln(r_2 + t), \\ \kappa_t = C_2(r_3 + t)^k, \end{cases}$$

where $C_1, C_2 > 0$, $r_1, r_3 > 0$, $r_2 \geq 1$ and still under the conditions $b > b(U_p)$, $k > 0$ and $c \geq 2k + 1$. It is possible to deduce more general conditions insuring the validity of the convergence results of Theorems 1.1 and 1.2 (see, e.g., Proposition 4.3 below).

How to choose in practice the exponents c and k satisfying $c \geq 2k+1$ in Theorem 1.2? We note that the larger c, the faster α goes to zero and the faster the algorithm Z is using the data $(Y_n)_{n \in \mathbb{N}}$. In compensation, k can be chosen larger, which means that ν is closer to its approximation by its transport through the kernel $K_{\cdot,\kappa_t}(\cdot)$ (defined before the statement of Theorem 1.2, for more details see Section 5), namely the convergence will be more precise. This is quite natural, since more data have been required at some fixed time. So in practice a trade-off has to be made between the number of i.i.d. variables distributed according to ν one has at his disposal and the quality of the approximation of $\mathcal{M}_p$.

1.3. Numerical illustration

The algorithm X (and similarly for Z) is not so difficult to implement. Let us identify $\mathbb{T}$ with $(-\pi, \pi]$ and construct X_t for some fixed $t > 0$. Assume we are given $(Y_n)_{n \in \mathbb{N}}$, $(\alpha_s)_{s \in [0,t]}$, $(\beta_s)_{s \in [0,t]}$ and X_0 as in the Introduction. We need furthermore two independent sequences $(\tau_n)_{n \in \mathbb{N}}$ and $(V_n)_{n \in \mathbb{N}}$, consisting of i.i.d. random variables, respectively, distributed according to the exponential law of parameter 1 and to the Gaussian law with mean 0 and variance 1. We begin by constructing the finite sequence $(T_n)_{n \in [\![0,N]\!]}$ corresponding to the jump times of $N^{(\alpha)}$: let $T_0 := 0$ and next by iteration, if T_n was defined, we take T_{n+1} such that $\int_{T_n}^{T_{n+1}} 1/\alpha_s \, ds = \tau_{n+1}$. This is done until $T_N > t$, with $N \in \mathbb{N}$, then we change the definition of T_N by imposing $T_N = t$. Next, we consider the sequence $(\check{X}_n, \widehat{X}_n)_{n \in [\![0,N]\!]}$ constructed through the following iteration (where the variables are reduced modulo 2π): starting from $\check{X}_0 := \widehat{X}_0 := X_0$, if $\widehat{X}_n$ was defined, with $n \in [\![0, N-1]\!]$, we consider

$$\check{X}_{n+1} := \widehat{X}_n + \sqrt{T_{n+1} - T_n} V_{n+1}. \tag{1.8}$$

Next, we define

$$\widehat{X}_{n+1} := \check{X}_{n+1} + (p/2)\alpha_{T_{n+1}}\beta_{T_{n+1}} |W_{n+1}|^{p-2} V_{n+1}, \tag{1.9}$$

where W_{n+1} is the representative of $Y_{n+1} - \check{X}_{n+1}$ in $(-\pi, \pi]$ modulo 2π. Then $\check{X}_N$ has the same law as X_t.

Theorems 1.1 and 1.2 provide theoretical results at very large times, but in practice, one has to work with a finite horizon t, for which the best corresponding scheme β may not be of the form of those given in these theorems (see the lectures of [9] for the classical simulated annealing algorithm). Thus, the previous theorems should only be seen as indications of what could be tried in practice. Let us illustrate that by some numerical simulations. On the circle, still identified with $(-\pi, \pi]$, consider the probability distribution $\nu = (\delta_0 + \delta_\pi)/2$. A priori we should resort to Theorem 1.2, but let us just “apply” Theorem 1.1 with $a = 1$, namely with the scheme

$$\forall t \geq 0, \qquad \alpha_t := \frac{1}{1+t}.$$

For $p = 1$ the function U_1 is constant, meaning that the set of medians $\mathcal{M}_1$ is the whole circle. For $p > 1$, the function U_p admits two global minima, $\mathcal{M}_p = \{-\pi/2, \pi/2\}$, and two global maxima, 0 and π. It is easy to see that $b(U_p) = \pi^p(1 - 2^{1-p})$, so that we can take, for instance,

$$\forall t \geq 0, \qquad \beta_t := \frac{2}{\pi^p(1 - 2^{1-p})} \ln(1+t),$$

(for $p = 1$, the factor in front of the logarithm can be chosen freely, one could even choose the scheme β to be constant). With the above notation, let $(Y_n)_{n\in\mathbb{N}}$, $(\tau_n)_{n\in\mathbb{N}}$ and $(V_n)_{n\in\mathbb{N}}$ be independent sequences consisting of i.i.d. random variables, respectively, distributed according to the uniform law on $\{0, \pi\}$, to the exponential law of parameter 1 and to the Gaussian law with mean 0 and variance 1. Let $t > 0$ be fixed. The finite sequence $(T_n)_{n\in[\![0,N]\!]}$ is constructed through the recurrence $T_0 = 100$ and

$$\forall n \in [\![0, N-1]\!], \qquad T_{n+1} := \sqrt{(T_n + 1)^2 + \tau_{n+1}} - 1$$

until $T_N > t$. Starting from $\check{X}_0 := \widehat{X}_0 := 0$, we consider the sequence $(\check{X}_n, \widehat{X}_n)_{n\in[\![0,N]\!]}$ defined via (1.8) and (1.9). The histograms of Figure 1 of the distribution of $\check{X}_N$ correspond to $p = 1.1$ and $p = 2$ and $t = 200$ and $t = 400$ and they are obtained with 100 samples of the procedure described above.

It appears that as time goes on, there is a tendency to concentrate on the set of means $\{-\pi/2, \pi/2\}$, but that this is more difficult to achieve for small $p > 1$, due to the fact that in the limit case $p = 1$, one is trying to sample according to the uniform distribution on $(-\pi, \pi]$.

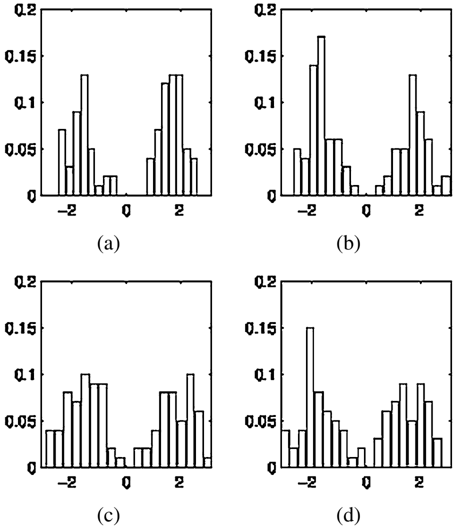

Figure 1. (a) $p = 2$ and $t = 200$, (b) $p = 2$ and $t = 400$, (c) $p = 1.1$ and $t = 200$, (d) $p = 1.1$ and $t = 400$.

Figure 2 is plotting a typical trajectory (observed at the jump times), with $p = 2$, $t = 400$ and for which the simulation gave $N = 150{,}366$ (close to $400^2 - 100^2$). It should be emphasized that if instead of using 100 samples in a Monte-Carlo procedure as above, one rather resorts to the empirical measure generated by one trajectory, one would get similar histograms.

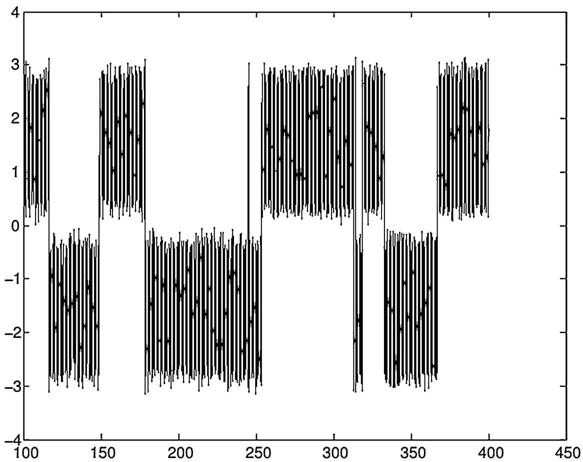

Figure 2. A trajectory for $p = 2$ and $t = 400$.

1.4. The conjecture for Riemannian manifolds

The description of the algorithm given in Section 1.2 can be extended to any compact Riemannian manifold M endowed with its distance d. For general books on Riemannian geometry, we refer to [19].

As above, let $Y := (Y_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables distributed according to a fixed probability measure ν on M. Let $p \in [1, +\infty)$ be fixed. We also need an inverse temperature evolution $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ and an inverse speed up evolution $\alpha : \mathbb{R}_+ \to \mathbb{R}_+^*$, which typically will be non-decreasing and non-increasing and satisfying $\lim_{t\to+\infty} \beta_t = +\infty$ and $\lim_{t\to+\infty} \alpha_t = 0$.

We consider again the speeded-up process $N^{(\alpha)} := (N_t^{(\alpha)})_{t\geq 0}$ via

$$\forall t \geq 0, \qquad N_t^{(\alpha)} := N_{\int_0^t 1/\alpha_s\, ds},$$

where $N := (N_t)_{t\geq 0}$ be a standard Poisson process independent from Y. The time-inhomogeneous Markov process $X := (X_t)_{t\geq 0}$ evolves in M in the following heuristic way: if $T > 0$ is a jump time of $N^{(\alpha)}$, then X jumps at the same time, from X_{T-} to

$$X_T := \exp_{X_{T-}}\big((p/2)\beta_T\alpha_T d^{p-2}(X_{T_-}, Y_{N_T^{(\alpha)}})\overrightarrow{X_{T-}Y_{N_T^{(\alpha)}}}\big).$$

By definition, the latter point is obtained by the following the shortest geodesic leading from X_{T-} to $Y_{N_T^{(\alpha)}}$ at time 1, during a time $s := (p/2)\beta_T\alpha_T d^{p-2}(X_{T_-}, Y_{N_T^{(\alpha)}})$ (and thus may not really correspond to an image of the exponential mapping if s is not small enough). The schemes α and β will satisfy $\lim_{t\to+\infty} \alpha_t\beta_t = 0$, so that for sufficiently large jump-times T, X_T will be between X_{T-} and $Y_{N_T^{(\alpha)}}$ on the above geodesic and quite close to X_{T-}. Almost surely, the above shortest geodesics are unique and there is no problem with their choices in the previous construction. Indeed, by the end of the description below, X_{T_-} will be independent of $Y_{N_T^{(\alpha)}}$ and the law of X_{T_-} will be absolutely continuous with respect to the Riemannian probability λ, namely the volume measure standardized to total volume one. It ensures that almost surely, $Y_{N_T^{(\alpha)}}$ is not in the cut-locus of X_{T_-} (which is negligible with respect to λ) so that there is only one shortest geodesic from X_{T-} to $Y_{N_T^{(\alpha)}}$. To proceed with the construction, we require that between consecutive jump times (and between time 0 and the first jump time), X evolves as a Brownian motion, relatively to the Riemannian structure of M (see, e.g., the book of [18]) and independently of Y and N. Very informally, the evolution of the algorithm X can be summarized by the equation (in the tangent bundle TM)

$$\forall t \geq 0, \qquad dX_t = dB_t + (p/2)\alpha_t\beta_t d^{p-2}(X_{T_-}, Y_{N_T^{(\alpha)}})\overrightarrow{X_{t-}Y_{N_t^{(\alpha)}}}\, dN_t^{(\alpha)},$$

where $(B_t)_{t\geq 0}$ would be a Brownian motion on M and where $(Y_{N_t^{(\alpha)}})_{t\geq 0}$ should be interpreted as a fast auxiliary process. The law of X is then entirely determined by the initial distribution $m_0 = \mathcal{L}(X_0)$. We believe that the above algorithm X finds in probability at large times the set $\mathcal{M}_p$ of p-means, at least if ν is sufficiently regular, as in the case where $M = \mathbb{T}$:

Conjecture 1.1. Assume that ν admits a density with respect to λ and that this density is Hölder continuous with exponent $a \in (0, 1]$. Then there exist two constants $a_p > 0$, depending on $p \geq 1$ and a, and $b_p \geq 0$, depending on p and M, such that for any scheme of the form given in (1.3), where $b > b_p$, we have for any neighborhood $\mathcal{N}$ of $\mathcal{M}_p$ and for any m_0,

$$\lim_{t\to+\infty} \mathbb{P}[X_t \in \mathcal{N}] = 1.$$

So as in Section 1.2, to find an element of $\mathcal{M}_p$ with an important probability, one should pick up the value of X_t for sufficiently large times t.

The constant $b_p \geq 0$ should still coincide with the one defined in (1.7).

Let us now extend the stochastic algorithm Z, which should enable one to find some of the p-means of any probability measure ν on the compact Riemannian manifold M.

For any $x \in M$ and $\kappa > 0$, consider, on the tangent space $T_x M$, the probability measure $\widetilde{K}_{x,\kappa}$ whose density with respect to the Lebesgue measure dv is proportional to $(1 - \kappa\|v\|)_+$ (where the Lebesgue measure and the norm are relative to the Euclidean structure on $T_x M$). Denote $K_{x,\kappa}$ the image by the exponential mapping at x of $\widetilde{K}_{x,\kappa}$. Assume next that we are given an evolution $\kappa : \mathbb{R}_+ \ni t \mapsto \kappa_t \in \mathbb{R}_+^*$ and consider the process $Z := (Z_t)_{t\geq 0}$ evolving similarly to $(X_t)_{t\geq 0}$, except that at the jump times T of $N^{(\alpha)}$, the target $Y_{N_T^{(\alpha)}}$ is replaced by a point W_T sampled from $K_{Y_{N_T^{(\alpha)}},\kappa_T}$, independently from the other variables. We believe that a variant of Theorem 1.2 should hold more generally on compact Riemannian manifolds M. But it seems that the geometry of M should play a role, especially through the behaviour of the volume of small enlargements of the cut-locus of points.

Notice that a major difficulty for implementing an algorithm in a high-dimensional manifold simulating the process X_t is to compute the logarithm map $\overrightarrow{xy} = \exp_x^{-1}(y)$. Moreover, this logarithm can be very instable around the cutlocus of x. In [4], it is proposed to replace it by the gradient of some cost function and then to follow the flow of this gradient.

1.5. Discussion

The purpose of this paper is to propose a stochastic algorithm finding p-means by a sequential use of samples from the underlying probability measure on a Riemannian manifold M, even if the formal proof of its convergence is only shown for the circle, the first non-trivial example.

When ν is an empirical measure $(\sum_{l=1}^N \delta_{x_l})/N$, where the x_l, $l \in [\![1, N]\!]$, are points on the circle, Charlier [10,15] and McKilliam, Quinn and Clarkson [21] proposed algorithms finding the 2-mean with complexities of order $N\ln(N)$ and N for the latter work. Empirical measures can in practice be used to approximate more general probability measures on the circle, but it seems this is not a very efficient method, since for each new point added to the empirical measure, the whole algorithm finding the corresponding mean has to be started again from scratch. To our knowledge, the process of Theorem 1.1 is the only algorithm finding p-means for any $p \geq 1$ and for any probability measure ν admitting Hölderian densities, even in the restricted situation of the circle.

Another strong motivation for this paper is the treatment of the jumps of the algorithms X and Z, situation which is not covered by the techniques of [25] (to the contrary of the jumps of the auxiliary process, which can be more easily dealt with).

In [4], we extend the ideas of the present paper to the situation were $d^p(x, y)$ in (1.1) is replaced by a quantity $\kappa(x, y)$ depending smoothly on the parameters x and y belonging to a compact Riemannian manifold M. Via convolutions with the underlying heat kernel, it leads to an algorithm enabling to deal with mappings κ which are only assumed to be continuous. But due to this regularization procedure, the corresponding algorithm is less straightforward to put in practice than the one presented here. Of course, the direct implementability has a price, since it needs precise information about a crucial object, $L^*_{\alpha,\beta}[\mathbb{1}]$. It will be defined in Section 3 and its investigation has to be divided in several cases depending on the value of p. This is hidden in [4], because we were more interested there in the generalization to general compact manifolds than in practicality considerations.

More technical discussions of the results are partially scattered over the manuscript, when it seems more appropriate to introduce them; see, for instance, Remarks 4.1, 4.2, 5.1 and 5.2.

The paper is constructed on the following plan. In next section, we recall some results about simulated annealing which give the heuristics for the above convergence. Another alternative algorithm is presented, in the same spirit as X and Z, but without jumps. In Section 3, we discuss about the regularity of the function U_p, in terms of that of ν. It enables to see how close is the instantaneous invariant measure associated to the algorithm at large times $t \geq 0$ to the Gibbs measures associated to the potential U_p and to the inverse temperature β_t^{-1}. The proof of Theorem 1.1 is given in Section 4. The fifth section is devoted to the extension presented in Theorem 1.2 and the Appendix deals with technicalities relative to the temporal marginal laws of the algorithms.

2. Principles underlying the proof

Here, some results about the classical simulated annealing are reviewed. The algorithm X described in the Introduction will then appear as a natural modification. This will also give us the opportunity to present another intermediate algorithm.

2.1. Simulated annealing

Consider again M a compact Riemannian manifold and denote $\langle\cdot,\cdot\rangle$, ∇, $\triangle$ and λ the corresponding scalar product, gradient, Laplacian operator and probability measure. Let U be a given smooth function on M to which we associate the constant $b(U) \geq 0$ defined similarly as in (1.6). We denote by $\mathcal{M}$ the set of global minima of U.

A corresponding simulated annealing algorithm $\theta := (\theta_t)_{t\geq 0}$ associated to a measurable inverse temperature scheme $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ is defined through the evolution equation

$$\forall t \geq 0, \qquad d\theta_t = dB_t - \frac{\beta_t}{2}\nabla U(\theta_t)\,dt.$$

It is a shorthand meaning that θ is a time-inhomogeneous Markov process whose generator at any time $t \geq 0$ is L_{β_t}, where

$$\forall \beta \geq 0, \qquad L_\beta \cdot := \tfrac{1}{2}\big(\triangle \cdot - \beta\langle \nabla U, \nabla \cdot \rangle\big). \tag{2.1}$$

Holley, Kusuoka and Stroock [14] have proven the following result.

Theorem 2.1. *For any fixed $T \geq 1$, consider the inverse temperature scheme*

$$\forall t \geq 0, \qquad \beta_t = b^{-1} \ln(T + t),$$

with $b > b(U)$. Then for any neighborhood $\mathcal{N}$ of $\mathcal{M}$ and for any initial distribution $\mathcal{L}(\theta_0)$, we have

$$\lim_{t \to +\infty} \mathbb{P}[\theta_t \in \mathcal{N}] = 1.$$

A crucial ingredient of the proof of this convergence are the Gibbs measures associated to the potential U. They are defined as the probability measures μ_β given for any $\beta \geq 0$ by

$$\mu_\beta(dx) := \frac{\exp(-\beta U(x))}{Z_\beta} \lambda(dx), \tag{2.2}$$

where $Z_\beta := \int \exp(-\beta U(x))\lambda(dx)$ is the normalizing factor.

Indeed, [14] show that $\mathcal{L}(\theta_t)$ and μ_{β_t} become closer and closer as $t \geq 0$ goes to infinity, for instance, in the sense of total variation:

$$\lim_{t \to +\infty} \left\| \mathcal{L}(\theta_t) - \mu_{\beta_t} \right\|_{\mathrm{tv}} = 0. \tag{2.3}$$

Theorem 2.1 is then an immediate consequence of the fact that for any neighborhood $\mathcal{N}$ of $\mathcal{M}$,

$$\lim_{\beta \to +\infty} \mu_\beta[\mathcal{N}] = 1.$$

The constant $b(U)$ is critical for the behaviour (2.3), in the sense that if we take

$$\forall t \geq 0, \qquad \beta_t = b^{-1} \ln(T + t),$$

with $T \geq 1$ and $b < b(U)$, then there exist initial distributions $\mathcal{L}(\theta_0)$ such that (2.3) is not true.

But in general (see, e.g., [24]), the constant $b(U)$ is not critical for Theorem 2.1, the corresponding critical constant being, with the notation of the Introduction,

$$b'(U) := \min_{x_0 \in \mathcal{M}} \max_{y \in M} U(x_0, y) - U(y) \leq b(U)$$

(compare with (1.7), where U replaces U_p and where a global minimum $x_0 \in \mathcal{M}$ is fixed). Note that it may happen that $b'(U) = b(U)$, for instance, if $\mathcal{M}$ has only one connected component.

Another remark about Theorem 2.1 is that the convergence in probability of θ_t for large $t \geq 0$ toward $\mathcal{M}$ cannot be improved into an almost sure convergence. Denote by A the connected component of $\{x \in M : U(x) \leq \min_M U + b\}$ which contains $\mathcal{M}$ (the condition $b > b(U)$ ensures that $\mathcal{M}$ is contained in only one connected component of the above set). Then almost surely, A is the limiting set of the trajectory $(\theta_t)_{t\geq 0}$ (see [23], where the corresponding result is proven for a finite state space but whose proof could be extended to the setting of Theorem 2.1). We believe that all these remarks should also hold in the context of Conjecture 1.1 and Theorem 1.1.

2.2. Heuristic of the proof

Let us now heuristically put forward why a result such as Conjecture 1.1 should be true, in relation with Theorem 2.1. For simplicity of the exposition, assume that ν is absolutely continuous with respect to λ. For almost every $x, y \in M$, there exists a unique minimal geodesic with speed 1 leading from x to y. Denote it by $(\gamma(x, y, t))_{t\in\mathbb{R}}$, so that $\gamma(x, y, 0) = x$ and $\gamma(x, y, d(x, y)) = y$. The process $(X_t)_{t\geq 0}$ underlying Theorem 2.1 is Markovian and its inhomogeneous family of generators is $(L_{\alpha_t,\beta_t})_{t\geq 0}$, where for any $\alpha > 0$ and $\beta \geq 0$, $L_{\alpha,\beta}$ acts on functions f from $\mathcal{C}^2(M)$ via, for all $x \in M$,

$$L_{\alpha,\beta}[f](x) := \frac{1}{2}\triangle f(x) + \frac{1}{\alpha}\int f\big(\gamma\big(x, y, (p/2)\beta\alpha d^{p-1}(x, y)\big)\big) - f(x)\nu(dy) \qquad (2.4)$$

(to simplify notation, we will try to avoid writing down explicitly the dependence on $p \geq 1$). The r is well-defined, due to the fact that $\nu \ll \lambda$ which implies that the cut-locus of x is negligible with respect to ν. Furthermore Fubini's theorem enables to see that the function $L_{\alpha,\beta}[f]$ is at least measurable. Next, we remark that as α goes to 0_+, we have for any $f \in \mathcal{C}^1(M)$, any $x \in M$ and any $y \in M$ which is not in the cut-locus of x,

$$\lim_{\alpha\to 0_+} \frac{f(\gamma(x, y, (p/2)\beta\alpha d^{p-1}(x, y))) - f(x)}{\alpha} = \frac{1}{2}\beta p d^{p-1}(x, y)\big\langle\nabla f(x), \dot{\gamma}(x, y, 0)\big\rangle,$$

for all $\beta \geq 0$, so that for any $f \in \mathcal{C}^2(M)$ and $x \in M$,

$$\forall\beta \geq 0, \qquad \lim_{\alpha\to 0_+} L_{\alpha,\beta}[f](x) = \frac{1}{2}\triangle f(x) + \frac{\beta}{2}p\int d^{p-1}(x, y)\big\langle\nabla f(x), \dot{\gamma}(x, y, 0)\big\rangle\nu(dy).$$

Recall that the potential $U = U_p$ we are now interested in is given by (1.1) and that for almost every $(x, y) \in M^2$,

$$\nabla_x d^p(x, y) = -pd^{d-1}(x, y)\dot{\gamma}(x, y, 0)$$

(problems occur for points x in the cut-locus of y and, if $p = 1$, for $x = y$), thus

$$\nabla U_p(x) = -p\int d^{p-1}(x, y)\dot{\gamma}(x, y, 0)\nu(dy). \qquad (2.5)$$

It follows that or any $f \in \mathcal{C}^2(M)$ and $x \in M$,

$$\forall \beta \geq 0, \qquad \lim_{\alpha \to 0_+} L_{\alpha,\beta}[f](x) = L_\beta[f](x).$$

Since $\lim_{t\to+\infty} \alpha_t = 0$, it appears that at least for large times, $(X_t)_{t\geq 0}$ and $(\theta_t)_{t\geq 0}$ should behave in a similar way. The validity of Theorem 2.1 for any $T \geq 1$ and any initial distribution $\mathcal{L}(\theta_0)$ then suggests that Conjecture 1.1 should hold. But this rough explanation is not sufficient to understand the choice of the scheme $(\alpha_t)_{t\geq 0}$, which will require more rigorous computations relatively to the corresponding homogenization property. The heuristics for Theorem 1.2 are similar, since the underlying algorithm $(Z_t)_{t\geq 0}$ is Markovian and its inhomogeneous family of generators $(L_{\alpha_t,\beta_t,\kappa_t})_{t\geq 0}$ satisfies

$$\forall f \in \mathcal{C}^2(M), \qquad \lim_{t\to+\infty} \left\| L_{\alpha_t,\beta_t,\kappa_t}[f] - L_{\beta_t}[f] \right\|_\infty = 0.$$

For any $\alpha > 0$, $\beta \geq 0$ and $\kappa > 0$, the generator $L_{\alpha,\beta,\kappa}$ acts on functions $f \in \mathcal{C}^2(M)$ via, for all $x \in M$,

$$L_{\alpha,\beta,\kappa}[f](x) := \frac{1}{2}\triangle f(x) + \frac{1}{\alpha}\int f\bigl(\gamma\bigl(x, z, (p/2)\beta\alpha d^{p-1}(x,z)\bigr)\bigr) - f(x) K_{y,\kappa}(dz)\nu(dy).$$

The previous observations suggest another possible algorithm to find the mean of a probability measure ν on M. Consider the $M \times M$-valued inhomogeneous Markov process $(\widetilde{X}_t, Y_{N_t^{(\alpha)}+1})_{t\geq 0}$ where $(N_t^{(\alpha)})_{t\geq 0}$ was defined in (1.2) and where

$$\forall t \geq 0, \qquad d\widetilde{X}_t = dB_t + (p/2)\beta_t d^{p-1}(\widetilde{X}_t, Y_{N_t^{(\alpha)}+1})\dot{\gamma}(\widetilde{X}_t, Y_{N_t^{(\alpha)}+1}, 0)\, dt. \tag{2.6}$$

Again, up to appropriate choices of the schemes $(\alpha_t)_{t\geq 0}$ and $(\beta_t)_{t\geq 0}$, it can be expected that for any neighborhood $\mathcal{N}$ of $\mathcal{M}$ and for any initial distribution $\mathcal{L}(\widetilde{X}_0)$,

$$\lim_{t\to+\infty} \mathbb{P}[\widetilde{X}_t \in \mathcal{N}] = 1.$$

Indeed, this can be obtained by following the line of arguments presented in [25]; see [3].

But the main drawback of the algorithm $(\widetilde{X}_t)_{t\geq 0}$ is that theoretically, it is asking for the computation of the unit vector $\dot{\gamma}(\widetilde{X}_t, Y_{N_t^{(\alpha)}+1}, 0)$ and of the distance $d(\widetilde{X}_t, Y_{N_t^{(\alpha)}+1})$, at any time $t \geq 0$. From a practical point of view, its complexity will be bad in comparison with that of the algorithm $X := (X_t)_{t\geq 0}$, which is not so difficult to implement, as it was seen in Section 1.3.

2.3. Outline of the proof

Since the Gibbs measure μ_β defined in (2.2), with U replaced by U_p, concentrates on $\mathcal{M}_p$ for large β, it will be sufficient to show that the law m_t of X_t becomes closer and closer to μ_{β_t} for large t. To measure this closeness, we use the $\mathbb{L}^2$-discrepancy of m_t with respect to μ_{β_t}

defined by

$$\forall t > 0, \qquad I_t := \int \left(\frac{m_t}{\mu_{\beta_t}} - 1 \right)^2 d\mu_{\beta_t}.$$

(Alternatively, it would be interesting to see if the considerations that follow could be extended to the case where this quantity is replaced by the more natural relative entropy of m_t with respect to μ_{β_t}.) To show that this quantity goes to zero as t becomes large, we study its temporal evolution, by differentiating it. The fact that μ_{β_t} is not the instantaneous invariant measure (namely the probability measure left invariant by the generator at time t), leads to supplementary term with respect to what one usually gets by applying this approach (see, e.g., [22]). This term measures in some sense the distance between μ_{β_t} and the instantaneous invariant measure at time t (which itself is not explicitly known). A large part of the paper is devoted to estimate this supplementary term, the final result being presented in Proposition 3.5. In Proposition 4.1, we deduce a bound on the evolution of the quantity I_t. To conclude in Proposition 4.3 that the obtained ordinary differential inequality is sufficient to conclude that $\lim_{t\to+\infty} I_t = 0$, we need an estimate of the spectral gap of the operator presented in (2.1) for large β. For that, we resort to a result due to [14] recalled in Proposition 4.2.

Let us emphasize that the resort to the object $L^*_{\alpha,\beta}[\mathbb{1}]$ defined and investigated in Section 3 to estimate the discrepancy between a well-known measure and an instantaneous invariant measure, which is more difficult to apprehend, should be of much broader use than the one presented here. Indeed, the function $L^*_{\alpha,\beta}[\mathbb{1}]$ is constructed by using directly only two objects which are supposed to be known: the generator and the convenient measure we choose to replace the instantaneous invariant measure, because $L^*_{\alpha,\beta}$ is just the dual operator of $L_{\alpha,\beta}$ in $\mathbb{L}^2[\mu_\beta]$ and $\mathbb{1}$ is the constant function taking the value 1.

3. Regularity issues

From this section on, we restrict ourselves to the case of the circle. Here, we investigate the regularity of the potential U_p introduced in (1.1) and use the obtained information to evaluate how far are the instantaneous invariant measures of the algorithm X from the corresponding Gibbs measures, as well as some other preliminary bounds.

For any $x \in \mathbb{T}$, we denote x' the unique point in the cut-locus of x, namely the opposite point $x' = x + \pi$. Recall that for $y \in \mathbb{T} \setminus \{x'\}$, $(\gamma(x, y, t))_{t\in\mathbb{R}}$ denotes the unique minimal geodesic with speed 1 going from x to y and that δ_x stands for the Dirac mass at x.

Lemma 3.1. *For any probability measure ν on $\mathbb{T}$, we have for the potential U_p defined in* (1.1), *in the distribution sense, for $x \in \mathbb{T}$,*

$$U_p''(x) = \begin{cases} p(p-1) \int_{\mathbb{T}} d^{p-2}(y, x) - 2p\pi^{p-1} \delta_{y'}(x) \nu(dy), & \text{if } p > 1, \\ 2 \int \left(\delta_y(x) - \delta_{y'}(x) \right) \nu(dy), & \text{if } p = 1. \end{cases}$$

In particular if ν *admits a continuous density with respect to* λ, *still denoted* ν, *then we have that* $U_p \in \mathcal{C}^2(\mathbb{T})$ *and*

$$\forall x \in \mathbb{T}, \qquad U_p''(x) = \begin{cases} p(p-1)\int_{\mathbb{T}} d^{p-2}(y,x)\nu(dy) - p\pi^{p-2}\nu(x'), & \text{if } p > 1, \\ (\nu(x) - \nu(x'))/\pi, & \text{if } p = 1. \end{cases}$$

Proof. We begin by considering the case where $p > 1$. Furthermore, we first investigate the situation where $\nu = \delta_y$ for some fixed $y \in \mathbb{T}$. Then $U_p(x) = d^p(x, y)$ for any $x \in \mathbb{T}$ and we have seen in (2.5) that

$$\forall x \neq y', \qquad U_p'(x) = -pd^{p-1}(x,y)\dot{\gamma}(x,y,0).$$

By continuity of U_p, this equality holds in the sense of distributions on the whole set $\mathbb{T}$. To compute U_p'', consider a test function $\varphi \in \mathcal{C}^\infty(\mathbb{T})$:

$$\begin{aligned} \int_{\mathbb{T}} \varphi'(x)U_p'(x)\,dx &= p\int_y^{y+\pi} \varphi'(x)(x-y)^{p-1}\,dx - p\int_{y-\pi}^{y} \varphi'(x)(y-x)^{p-1}\,dx \\ &= p\big[\varphi(x)(x-y)^{p-1}\big]_y^{y+\pi} - p(p-1)\int_y^{y+\pi} \varphi(x)(x-y)^{p-2}\,dx \\ &\quad - p\big[\varphi(x)(y-x)^{p-1}\big]_{y-\pi}^{y} - p(p-1)\int_{y-\pi}^{y} \varphi(x)(y-x)^{p-2}\,dx \\ &= 2p\pi^{p-1}\varphi(y') - p(p-1)\int_{\mathbb{T}} \varphi(x)d^{p-2}(y,x)\,dx. \end{aligned}$$

So we get that for $x \in \mathbb{T}$,

$$U_p''(x) = p(p-1)d^{p-2}(y,x) - 2p\pi^{p-1}\delta_{y'}(x).$$

If $p = 1$, starting again from

$$\forall x \neq y', \qquad U_1'(x) = -\dot{\gamma}(x,y,0), \tag{3.1}$$

we rather get for any test function $\varphi \in \mathcal{C}^\infty(\mathbb{T})$:

$$\begin{aligned} \int_{\mathbb{T}} \varphi'(x)U_1'(x)\,dx &= \int_y^{y+\pi} \varphi'(x)\,dx - \int_{y-\pi}^{y} \varphi'(x)\,dx \\ &= 2(\varphi(y') - \varphi(y)), \end{aligned}$$

so that

$$U_1'' = 2(\delta_y - \delta_{y'}).$$

The general case of a probability measure ν follows by integration with respect to $\nu(dy)$.

The second announced result follows from the observation that if ν admits a density with respect to λ, we can write for any $x \in \mathbb{T}$,

$$\begin{aligned}\int \delta_{y'}(x)\nu(dy) &= \int \delta_{x'}(y)\nu(y)\frac{dy}{2\pi} \\ &= \frac{\nu(x')}{2\pi}.\end{aligned}$$

□

In particular, it appears that the potential U_p belongs to $\mathcal{C}^\infty(\mathbb{T})$, if the density ν is smooth. Let us come back to the case of a general probability measure ν on $\mathbb{T}$. For any $\alpha > 0$ and $\beta \geq 0$, we are interested into the generator $L_{\alpha,\beta}$ defined in (2.4). Rigorously speaking, this definition is only valid if ν is absolutely continuous. Otherwise, the right-hand side of (2.4) is not well-defined for $x \in \mathbb{T}$ belonging to the union of the cut-locus of the atoms of ν. To get around this little inconvenience, one can consider for $x \in \mathbb{T}$, $(\gamma_+(x, x+\pi, t))_{t\in\mathbb{R}}$ and $(\gamma_-(x, x+\pi, t))_{t\in\mathbb{R}}$, the unique minimal geodesics with speed 1 leading from x to $x+\pi$, respectively, in the anti-clockwise (namely increasing in the cover $\mathbb{R}$ of $\mathbb{T}$) and clockwise direction. If $y \in \mathbb{T} \setminus \{x'\}$, we take as before $(\gamma_+(x, y, t))_{t\in\mathbb{R}} := (\gamma(x, y, t))_{t\in\mathbb{R}} =: (\gamma_-(x, y, t))_{t\in\mathbb{R}}$. Next, let k be a Markov kernel from $\mathbb{T}^2$ to $\{-, +\}$ and modify the definition (2.4) by imposing that for any $f \in \mathcal{C}^2(\mathbb{T})$ and any $x \in \mathbb{T}$,

$$L_{\alpha,\beta}[f](x) := \frac{1}{2}\partial^2 f(x) + \frac{1}{\alpha}\int f\bigl(\gamma_s\bigl(x, y, (p/2)\beta\alpha d^{p-1}(x, y)\bigr)\bigr) - f(x)k\bigl((x, y), ds\bigr)\nu(dy),$$

where ∂ stands for the natural derivative on $\mathbb{T}$. Then the function $L_{\alpha,\beta}[f]$ is at least measurable. But these considerations are not very relevant, since for any given measurable evolutions $\mathbb{R}_+ \ni t \mapsto \alpha_t \in \mathbb{R}_+^*$ and $\mathbb{R}_+ \ni t \mapsto \beta_t \in \mathbb{R}_+$, the solutions to the martingale problems associated to the inhomogeneous family of generators $(L_{\alpha_t,\beta_t})_{t\geq 0}$ (see, e.g., the book of [11]) are all the same and are described in a probabilistic way as the trajectory laws of the processes X presented in the Introduction. Indeed, this is a consequence of the absolute continuity of the heat kernel at any positive time (for arguments in the same spirit, see the Appendix). So to simplify notation, we only consider the case where $k((x, y), -) = 0$ for any $x, y \in \mathbb{T}$, this brought us back to the definition (2.4), where $(\gamma(x, y, t))_{t\in\mathbb{R}}$ stands for $(\gamma_+(x, y, t))_{t\in\mathbb{R}}$, for any $x, y \in \mathbb{T}$.

As it was mentioned for usual simulated annealing algorithms in the previous section, a traditional approach to prove Theorem 1.1 would try to evaluate at any time $t \geq 0$, how far is $\mathcal{L}(X_t)$ from the instantaneous invariant probability μ_{α_t,β_t}, namely that associated to L_{α_t,β_t}. Unfortunately, for any $\alpha > 0$ and $\beta \geq 0$, we have little information about the invariant probability $\mu_{\alpha,\beta}$ of $L_{\alpha,\beta}$, even its existence cannot be deduced directly from the compactness of $\mathbb{T}$, because the functions $L_{\alpha,\beta}[f]$ are not necessarily continuous for $f \in \mathcal{C}^2(\mathbb{T})$. Indeed it will be more convenient to use the Gibbs distribution μ_β defined in (2.2) for $\beta \geq 0$, where U is replaced by U_p. It has the advantage to be explicit and easy to work with, in particular it is clear that for large $\beta \geq 0$, μ_β concentrates around $\mathcal{M}_p$, the set of p-means of ν.

The remaining part of this section is mainly devoted to a quantification of what separates μ_β from being an invariant probability of $L_{\alpha,\beta}$, for $\alpha > 0$ and $\beta \geq 0$. It will become clear in the next section that a practical way to measure this discrepancy is through the evaluation of $\mu_\beta[(L^*_{\alpha,\beta}[\mathbb{1}])^2]$, where $L^*_{\alpha,\beta}$ is the dual operator of $L_{\alpha,\beta}$ in $\mathbb{L}^2(\mu_\beta)$ and where $\mathbb{1}$ is the constant function taking the value 1. Indeed, it can be seen that $L^*_{\alpha,\beta}[\mathbb{1}] = 0$ in $\mathbb{L}^2(\mu_\beta)$ if and only if μ_β is invariant for $L_{\alpha,\beta}$. We will also take advantage of the computations made in this direction to provide some estimates on related quantities which will be helpful later on.

Since the situation of the usual mean $p = 2$ is important and is simpler than the other cases, we first treat it in detail in the following subsection. Next, we will investigate the differences appearing in the situation of the median. The third subsection will deal with the cases $1 < p < 2$, whose computations are technical and not very enlightening. We will only give some indications about the remaining situation $p \in (2, \infty)$, which is less involved.

Some other preliminaries about the regularity of the time marginal laws of the considered algorithms will be treated in the Appendix. They are of a more qualitative nature and will mainly serve to justify some computations of the next sections, in some sense they are less relevant than the estimates and proofs of Propositions 3.1, 3.2, 3.3 and 3.4 below, which are really at the heart of our developments.

3.1. Estimate of $L^*_{\alpha,\beta}[\mathbb{1}]$ in the case $p = 2$

Before being more precise about the definition of $L^*_{\alpha,\beta}$, we need an elementary result, where we will use the following notation: for $y \in \mathbb{T}$ and $\delta \geq 0$, $B(y, \delta)$ stands for the open ball centered at y of radius δ and for any $s \in \mathbb{R}$, $T_{y,s}$ is the operator acting on measurable functions f defined on $\mathbb{T}$ via

$$\forall x \in \mathbb{T}, \qquad T_{y,s} f(x) := f\big(\gamma\big(x, y, sd(x, y)\big)\big). \tag{3.2}$$

Lemma 3.2. *For any $y \in \mathbb{T}$, any $s \in [0, 1)$ and any measurable and bounded functions f, g, we have*

$$\int_{\mathbb{T}} g T_{y,s} f \, d\lambda = \frac{1}{1-s} \int_{B(y,(1-s)\pi)} f T_{y,-s/(1-s)} g \, d\lambda.$$

Proof. By definition, we have

$$2\pi \int_{\mathbb{T}} g T_{y,s} f \, d\lambda = \int_{y-\pi}^{y+\pi} g(x) f\big(x + s(y - x)\big) \, dx.$$

In the right-hand side, consider the change of variables $z := sy + (1 - s)x$ to get that it is equal to

$$\frac{1}{1-s} \int_{y-(1-s)\pi}^{y+(1-s)\pi} g\left(\frac{z - sy}{1-s}\right) f(z) \, dz = \frac{2\pi}{1-s} \int_{B(y,(1-s)\pi)} f T_{y,-s/(1-s)} f \, d\lambda,$$

which corresponds to the announced result. □

This lemma has for consequence the next result, where $\mathcal{D}$ is the subspace of $\mathbb{L}^2(\lambda)$ consisting of functions whose second derivative in the distribution sense belongs to $\mathbb{L}^2(\lambda)$ (or equivalently to $\mathbb{L}^2(\mu_\beta)$ for any $\beta \geq 0$).

Lemma 3.3. *For $\alpha > 0$ and $\beta \geq 0$ such that $\alpha\beta \in [0, 1)$, the domain of the maximal extension of $L_{\alpha,\beta}$ on $\mathbb{L}^2(\mu_\beta)$ is $\mathcal{D}$. Furthermore, the domain $\mathcal{D}^*$ of its dual operator $L^*_{\alpha,\beta}$ in $\mathbb{L}^2(\mu_\beta)$ is the space $\{f \in \mathbb{L}^2(\mu_\beta) : \exp(-\beta U_2) f \in \mathcal{D}\}$ and we have for any $f \in \mathcal{D}^*$,*

$$\begin{aligned} L^*_{\alpha,\beta} f &= \frac{1}{2} \exp(\beta U_2) \partial^2 \bigl[\exp(-\beta U_2) f\bigr] \\ &\quad + \frac{\exp(\beta U_2)}{\alpha(1 - \alpha\beta)} \int \mathbb{1}_{B(y,(1-\alpha\beta)\pi)} T_{y,-\alpha\beta/(1-\alpha\beta)} \bigl[\exp(-\beta U_2) f\bigr] \nu(dy) - \frac{f}{\alpha}. \end{aligned}$$

In particular, if ν admits a continuous density, then $\mathcal{D}^* = \mathcal{D}$ and the above formula holds for any $f \in \mathcal{D}$.

Proof. With the previous definitions, we can write for any $\alpha > 0$ and $\beta \geq 0$,

$$L_{\alpha,\beta} = \frac{1}{2} \partial^2 + \frac{1}{\alpha} \int T_{y,\alpha\beta} \nu(dy) - \frac{I}{\alpha},$$

where I is the identity operator. Note furthermore that the identity operator is bounded from $\mathbb{L}^2(\lambda)$ to $\mathbb{L}^2(\mu_\beta)$ and conversely. Thus, to get the first assertion, it is sufficient to show that $\int T_{y,\alpha\beta} \nu(dy)$ is bounded from $\mathbb{L}^2(\lambda)$ to itself, or even only that $\|T_{y,\alpha\beta}\|_{\mathbb{L}^2(\lambda)\circlearrowleft}$ is uniformly bounded in $y \in \mathbb{T}$. To see that this is true, consider a bounded and measurable function f and assume that $\alpha\beta \in [0, 1)$. Since $(T_{y,\alpha\beta} f)^2 = T_{y,\alpha\beta} f^2$, we can apply Lemma 3.2 with $s = \alpha\beta$, $g = \mathbb{1}$ and f replaced by f^2 to get that

$$\begin{aligned} \int (T_{y,\alpha\beta} f)^2 \, d\lambda &= \frac{1}{1 - \alpha\beta} \int_{B(y,(1-s)\pi)} f^2 T_{y,-\alpha\beta/(1-\alpha\beta)} \mathbb{1} \, d\lambda \\ &= \frac{1}{1 - \alpha\beta} \int_{B(y,(1-s)\pi)} f^2 \, d\lambda \\ &\leq \frac{1}{1 - \alpha\beta} \int f^2 \, d\lambda. \end{aligned}$$

Next to see that for any $f, g \in \mathcal{C}^2(\mathbb{T})$,

$$\int g L_{\alpha,\beta} f \, d\mu_\beta = \int f L^*_{\alpha,\beta} g \, d\mu_\beta, \tag{3.3}$$

where $L^*_{\alpha,\beta}$ is the operator defined in the statement of the lemma, we note that, on one hand,

$$\begin{aligned} \int g \partial^2 f \, d\mu_\beta &= Z_\beta^{-1} \int \exp(-\beta U_2) g \partial^2 f \, d\lambda \\ &= \int f \exp(\beta U_2) \partial^2 \bigl[\exp(-\beta U_2) g\bigr] \, d\mu_\beta \end{aligned}$$

and on the other hand, for any $y \in \mathbb{T}$,

$$\int g T_{y,\alpha\beta} f \, d\mu_\beta = Z_\beta^{-1} \int \exp(-\beta U_2) g T_{y,\alpha\beta} f \, d\lambda,$$

so that we can use again Lemma 3.2. After an additional integration with respect to $\nu(dy)$, (3.3) follows without difficulty. To conclude, it is sufficient to see that for any $f \in \mathbb{L}^2(\mu_\beta)$, $L^*_{\alpha,\beta} f \in \mathbb{L}^2(\mu_\beta)$ (where $L^*_{\alpha,\beta} f$ is first interpreted as a distribution) if and only if $\exp(-\beta U_2) f \in \mathcal{D}$. This is done by adapting the arguments given in the first part of the proof, in particular we get that

$$\begin{aligned} &\left\| \frac{\exp(\beta U_2)}{\alpha(1-\alpha\beta)} \int \mathbb{1}_{B(y,(1-\alpha\beta)\pi)} T_{y,-\alpha\beta/(1-\alpha\beta)} \big[\exp(-\beta U_2) \cdot \big] \nu(dy) \right\|^2_{\mathbb{L}^2(\lambda) \circlearrowleft} \\ &\quad \le \frac{\exp(2\beta \operatorname{osc}(U_2))}{\alpha^2(1-\alpha\beta)}. \end{aligned}$$

□

Remark 3.1. By working in a similar spirit, the previous lemma, except for the expression of $L^*_{\alpha,\beta}$, is valid for any $\alpha > 0$ and $\beta \ge 0$ such that $\alpha\beta \ne 1$. The case $\alpha\beta = 1$ can be different: it follows from

$$L_{\alpha,1/\alpha} = \frac{1}{2}\partial^2 + \frac{1}{\alpha}(\nu - I),$$

that if ν does not admit a density with respect to λ which belongs to $\mathbb{L}^2(\lambda)$, then the domain of definition of $L^*_{\alpha,1/\alpha}$ is $\mathcal{D}^* \cap \{f \in \mathbb{L}^2(\mu_\beta) : \mu_\beta[f] = 0\}$, subspace which is not dense in $\mathbb{L}^2(\lambda)$ and worse for our purposes, which does not contain $\mathbb{1}$. Anyway, this degenerate situation is not very interesting for us, because the evolutions $(\alpha_t)_{t\ge 0}$ and $(\beta_t)_{t\ge 0}$ we consider satisfy $\alpha_t\beta_t \in (0,1)$ for t large enough. Furthermore, we will consider probability measures ν admitting a continuous density, in particular belonging to $\mathbb{L}^2(\lambda)$. In this case, $L_{\alpha,1/\alpha}$ and $L^*_{\alpha,1/\alpha}$ admit $\mathcal{D}$ for natural domain, as in fact $L_{\alpha,\beta}$ and $L^*_{\alpha,\beta}$ for any $\beta \ge 0$.

For any $\alpha > 0$ and $\beta \ge 0$ such that $\alpha\beta \in [0,1)$, denote $\eta = \alpha\beta/(1-\alpha\beta)$. As seen from the previous lemma, a consequence of the assumption that U_2 is $\mathcal{C}^2$ is that for any $x \in \mathbb{T}$,

$$\begin{aligned} L^*_{\alpha,\beta}\mathbb{1}(x) &= \frac{1}{2}\exp\big(\beta U_2(x)\big)\partial^2 \exp\big(-\beta U_2(x)\big) - \frac{1}{\alpha} \\ &\quad + \frac{\exp(\beta U_2(x))}{\alpha(1-\alpha\beta)} \int \mathbb{1}_{B(y,(1-\alpha\beta)\pi)}(x) T_{y,-\eta}\big[\exp(-\beta U_2)\big](x)\nu(dy) \\ &= \frac{\beta^2}{2}\big(U_2'(x)\big)^2 - \frac{\beta}{2}U_2''(x) - \frac{1}{\alpha} \\ &\quad + \frac{1}{\alpha(1-\alpha\beta)} \int_{B(x,(1-\alpha\beta)\pi)} \exp\big(\beta\big[U_2(x) - U_2\big(\gamma\big(x,y,-\eta d(x,y)\big)\big)\big]\big)\nu(dy). \end{aligned} \tag{3.4}$$

It appears that $L^*_{\alpha,\beta}\mathbb{1}$ is defined and continuous if ν has a continuous density (with respect to λ). The next result evaluates the uniform norm of this function under a little stronger regularity assumption. Despite it may seem quite plain, we would like to emphasize that the use of an estimate of $L^*_{\alpha,\beta}\mathbb{1}$ to replace the invariant measure of $L_{\alpha,\beta}$ by the more tractable μ_β is a key to all the results presented in the Introduction.

Proposition 3.1. *Assume that* ν *admits a density with respect to* λ *which is Hölder continuous, that is, there exists* $a \in (0,1]$ *and* $A > 0$ *such that*

$$\forall x, y \in \mathbb{T}, \qquad \left|\nu(y) - \nu(x)\right| \le A d^a(x,y). \tag{3.5}$$

Then there exists a constant $C(A) > 0$, *only depending on* A, *such that for any* $\beta \ge 1$ *and* $\alpha \in (0, 1/(2\beta^2))$, *we have*

$$\left\| L^*_{\alpha,\beta}\mathbb{1}\right\|_\infty \le C(A)\max\left(\alpha\beta^4, \alpha^a\beta^{1+a}\right).$$

Proof. In view of the expression of $L^*_{\alpha,\beta}\mathbb{1}(x)$ given before the statement of the proposition, we want to estimate for any fixed $x \in \mathbb{T}$, the quantity

$$\begin{aligned}&\int_{B(x,(1-\alpha\beta)\pi)} \exp\left(\beta\left[U_2(x) - U_2\left(\gamma\left(x, y, -\eta d(x,y)\right)\right)\right]\right)\nu(dy)\\ &\quad = \int_{x-(1-\alpha\beta)\pi}^{x+(1-\alpha\beta)\pi} \exp\left(\beta\left[U_2(x) - U_2\left(x - \eta(y-x)\right)\right]\right)\nu(dy).\end{aligned}$$

Lemma 3.1 and the continuity of the density ν ensure that $U_2 \in \mathcal{C}^2(\mathbb{T})$. Furthermore, since this density takes the value 1 somewhere on $\mathbb{T}$, we get that

$$\left\|U_2''\right\|_\infty \le 2A\pi^a \le 2\pi A. \tag{3.6}$$

Since U_2' vanishes somewhere on $\mathbb{T}$, we can deduce from this bound that $\|U_2'\|_\infty \le 4\pi^2 A$, but for $A > 1/(2\pi)$, it is better to use (2.5), which gives directly $\|U_2'\|_\infty \le 2\pi$.

Expanding the function U_2 around x, we see that for any $y \in (x - (1-\alpha\beta)\pi, x + (1-\alpha\beta)\pi)$ and $\eta \in (0,1]$ (this is satisfied because the assumptions on α and β ensure that $\alpha\beta \in (0,1/2)$), we can find $z \in (x - (1-\alpha\beta)\pi, x + (1-\alpha\beta)\pi)$ such that

$$\beta\left[U_2(x) - U_2\left(x - \eta(y-x)\right)\right] = \beta\eta U_2'(x)(y-x) - \beta\eta^2 U_2''(z)\frac{(y-x)^2}{2}.$$

The last term can be written under the form $\mathcal{O}_A(\alpha^2\beta^3)$, where for any $\epsilon > 0$, $\mathcal{O}_A(\epsilon)$ designates a quantity which is bounded by $K(A)\epsilon$, where $K(A)$ is a constant depending only on A (as usual $\mathcal{O}$ has a similar meaning, but with a universal constant). Note that we also have $\beta\eta U_2'(x)(y - x) = \mathcal{O}(\alpha\beta^2)$. Observing that for any $r, s \in \mathbb{R}$, we can find $u, v \in (0,1)$ such that $\exp(r+s) = (1 + r + r^2\exp(ur)/2)(1 + s\exp(vs))$ and in conjunction with the assumption $\alpha\beta^2 \le 1/2$, we can write that

$$\exp\left(\beta\left[U_2(x) - U_2\left(x - \eta(y-x)\right)\right]\right) = 1 + \beta\eta U_2'(x)(y-x) + \mathcal{O}_A\left(\alpha^2\beta^4\right). \tag{3.7}$$

Integrating this expression, we get that

$$\int_{B(x,(1-\alpha\beta)\pi)} \exp\bigl(\beta\bigl[U_2(x) - U_2\bigl(\gamma(x, y, -\eta)\bigr)\bigr]\bigr)\nu(dy)$$
$$= \nu\bigl[B\bigl(x, (1-\alpha\beta)\pi\bigr)\bigr] + \beta\eta U_2'(x) \int_{x-(1-\alpha\beta)\pi}^{x+(1-\alpha\beta)\pi} y - x\,\nu(dy) + \mathcal{O}_A\bigl(\alpha^2\beta^4\bigr).$$

Recalling that ν has no atom, the first term is equal to $1 - \nu(B(x', \alpha\beta\pi)$. Taking into account (2.5), we have $U_2'(x) = -2\int_{x-\pi}^{x+\pi} y - x\,\nu(dy)$, so that the second term is equal to

$$\beta\eta U_2'(x) \int_{x-\pi}^{x+\pi} y - x\,\nu(dy) - \beta\eta U_2'(x) \int_{x'-\alpha\beta\pi}^{x'+\alpha\beta\pi} y - x\,\nu(dy) = -\frac{\beta\eta}{2}\bigl(U_2'(x)\bigr)^2 + \mathcal{O}_A\bigl(\alpha^2\beta^3\bigr)$$

(in the last term of the left-hand side, $y - x$ is to be interpreted as its representative in $(-\pi, \pi]$ modulo 2π). We can now return to (3.4) and recalling the expression for U_2'' given in Lemma 3.1, we obtain that for any $x \in \mathbb{T}$,

$$\begin{aligned}
L^*_{\alpha,\beta}\mathbb{1}(x) &= \frac{\beta^2}{2}\bigl(U_2'(x)\bigr)^2 - \beta\bigl(1 - \nu\bigl(x'\bigr)\bigr) - \frac{1}{\alpha} \\
&\quad + \frac{1}{\alpha(1-\alpha\beta)}\biggl(1 - \nu\bigl(B\bigl(x', \alpha\beta\pi\bigr)\bigr) - \frac{\beta\eta}{2}\bigl(U_2'(x)\bigr)^2 + \mathcal{O}_A\bigl(\alpha^2\beta^4\bigr)\biggr) \\
&= \frac{1}{\alpha(1-\alpha\beta)} - \beta - \frac{1}{\alpha} + \frac{\beta^2}{2}\biggl(1 - \frac{1}{(1-\alpha\beta)^2}\biggr)\bigl(U_2'(x)\bigr)^2 \\
&\quad + \beta\biggl(\nu\bigl(x'\bigr) - \frac{\nu(B(x', \alpha\beta\pi))}{\alpha\beta(1-\alpha\beta)}\biggr) + \mathcal{O}_A\bigl(\alpha\beta^4\bigr) \\
&= \beta\biggl(\nu\bigl(x'\bigr) - \frac{\nu(B(x', \alpha\beta\pi))}{\alpha\beta(1-\alpha\beta)}\biggr) + \mathcal{O}_A\bigl(\alpha\beta^4\bigr) \\
&= \frac{\beta}{1-\alpha\beta}\biggl(\nu\bigl(x'\bigr) - \frac{\nu(B(x', \alpha\beta\pi))}{\alpha\beta}\biggr) - \frac{\alpha\beta^2}{1-\alpha\beta}\nu\bigl(x'\bigr) + \mathcal{O}_A\bigl(\alpha\beta^4\bigr) \\
&= \frac{\beta}{1-\alpha\beta}\frac{1}{2\pi\alpha\beta}\int_{x'-\alpha\beta\pi}^{x'+\alpha\beta\pi} \nu\bigl(x'\bigr) - \nu(y)\,dy + \mathcal{O}_A\bigl(\alpha\beta^4\bigr).
\end{aligned}$$

The justification of the Hölder continuity comes above all from the evaluation of the latter integral:

$$\begin{aligned}
\biggl|\int_{x'-\alpha\beta\pi}^{x'+\alpha\beta\pi} \nu\bigl(x'\bigr) - \nu(y)\,dy\biggr| &\le A\int_{x'-\alpha\beta\pi}^{x'+\alpha\beta\pi} \bigl|x' - y\bigr|^a\,dy \\
&= 2A\frac{(\alpha\beta\pi)^{1+a}}{1+a} \\
&\le 2A(\alpha\beta\pi)^{1+a}.
\end{aligned}$$

The bound announced in the lemma follows at once. □

To finish this subsection, let us present a related but more straightforward preliminary bound.

Lemma 3.4. *There exists a constant $k>0$ such that for any $s>0$ and $\beta\geq 1$ with $\beta s\leq 1/2$, we have, for any $y\in\mathbb{T}$ and $f\in\mathcal{C}^1(\mathbb{T})$,*

$$\int_{B(y,(1-s)\pi)}\big(T^*_{y,s}[g_y](x)-g_y(x)\big)^2\mu_\beta(dx)\leq ks^2\beta^2\bigg(\int(\partial f)^2\,d\mu_\beta+\int f^2\,d\mu_\beta\bigg), \tag{3.8}$$

where $T^*_{y,s}$ is the adjoint operator of $T_{y,s}$ in $\mathbb{L}^2(\mu_\beta)$ and where for any fixed $y\in\mathbb{T}$,

$$\forall x\in\mathbb{T}\setminus\{y'\},\qquad g_y(x):=f(x)d(x,y)\dot{\gamma}(x,y,0)$$

(*neglecting the cut-locus point y' of y*).

Proof. Since the problem is clearly invariant by translation of $y\in\mathbb{T}$, we can work with a fixed value of y, the most convenient to simplify the notation being $y=0\in\mathbb{R}/(2\pi\mathbb{Z})$. Then the function $g\equiv g_0$ is given by $g(x)=-xf(x)$ for $x\in(-\pi,\pi)$.

Due to the above assumptions, $s\in(0,1/2)$ and we are in position to use Lemma 3.2 to see that for $s\in(0,1/2)$ and for a.e. $x\in(-(1-s)\pi,(1-s)\pi)$,

$$T^*_s[g](x)=\frac{1}{1-s}\exp\big(\beta U_2(x)\big)T_{-\eta}\big[\exp(-\beta U_2)g\big](x),$$

with $\eta:=s/(1-s)$ and where we simplified notation by replacing $T^*_{0,s}$ and $T_{0,-\eta}$ by T^*_s and $T_{-\eta}$. This observation induces us to introduce on $(-(1-s)\pi,(1-s)\pi)$ the decomposition

$$T^*_s[g]-g=T^*_s[g]-\frac{1}{1-s}T_{-\eta}[g]+\frac{1}{1-s}\big(T_{-\eta}[g]-g\big)+\frac{s}{1-s}g,$$

leading to

$$\int\big(T^*_s[g](x)-g(x)\big)^2\mu_\beta(dx)\leq\frac{3}{(1-s)^2}J_1+\frac{3}{(1-s)^2}J_2+\frac{3s^2}{(1-s)^2}J_3, \tag{3.9}$$

where

$$J_1:=\int_{-(1-s)\pi}^{(1-s)\pi}\Big(\exp\big(\beta\big[U_2(x)-U_2\big((1+\eta)x\big)\big]\big)-1\Big)^2\big(T_{-\eta}[g]\big)^2\mu_\beta(dx),$$

$$J_2:=\int_{-(1-s)\pi}^{(1-s)\pi}\big(T_{-\eta}[g]-g\big)^2\,d\mu_\beta,$$

$$J_3:=\int_{-(1-s)\pi}^{(1-s)\pi}g^2\,d\mu_\beta.$$

The simplest term to treat is J_3: we just bound it above by $\int g^2\, d\mu_\beta$. Recalling that $g \le \pi^2 f^2$, we end up with a bound which goes in the direction of (3.8), due to the factor $3s^2/(1-s)^2$ in (3.9) and the fact that $\beta \ge 1$.

Next, we estimate the term J_1. Via the change of variable $z := (1+\eta)x$, Lemma 3.2 enables to write it down under the form

$$\begin{aligned}
&(1-s)\int_{\mathbb{T}} \bigl(\exp\bigl(\beta\bigl[U_2\bigl((1-s)z\bigr)-U_2(z)\bigr]\bigr)-1\bigr)^2 g^2(z)\exp(\beta\bigl[U_2(z)-U_2\bigl((1-s)z\bigr)\bigr]\mu_\beta(dz)\\
&\quad = 4(1-s)\int_{\mathbb{T}} \sinh^2\bigl(\beta\bigl[U_2\bigl((1-s)z\bigr)-U_2(z)\bigr]/2\bigr)g^2(z)\mu_\beta(dz).
\end{aligned}$$

Since $\beta s \le 1/2$, we are assured of the bounds

$$\begin{aligned}
\bigl|\beta\bigl[U_2\bigl((1-s)z\bigr)-U_2(z)\bigr]\bigr| &\le \beta\|U_2'\|_\infty \pi s\\
&\le 4\pi^2\beta s \qquad (3.10)\\
&\le 2\pi^2
\end{aligned}$$

and we deduce that

$$J_1 \le 16\pi^4\cosh^2\bigl(\pi^2\bigr)\beta^2 s^2 \int g^2\, d\mu_\beta.$$

Again this bound is going in the direction of (3.8).

We are thus left with the task of finding a bound on J_2 and this is where the Dirichlet type quantity $\int (f')^2\, d\mu_\beta$ will be needed. Of course, its origin is to be found in the fundamental theorem of calculus, which enables to write for any $x \in (-(1-s)\pi, (1-s)\pi)$,

$$T_{-\eta}[g](x) - g(x) = -\eta \int_0^1 g'\bigl((1+\eta v)x\bigr)x\, dv.$$

It follows that

$$J_2 \le \pi^2\eta^2 \int_{-(1-s)\pi}^{(1-s)\pi} \mu_\beta(dx)\int_0^1 dv\bigl(g'\bigl((1+\eta v)x\bigr)\bigr)^2. \qquad (3.11)$$

Recalling the definition of g, we have for any $z \in (-\pi, \pi)$,

$$\bigl(g'(z)\bigr)^2 \le 2\bigl(\pi^2\bigl(f'(z)\bigr)^2 + f^2(z)\bigr),$$

where we used again that $\|U_2'\|_\infty \le 2\pi$ and that $\beta \ge 1$. Next, we deduce from a computation similar to (3.10) and from $\eta \le 2s$ that

$$\frac{\mu_\beta(x)}{\mu_\beta((1+\eta v)x)} \le \exp\bigl(4\pi^2\bigr),$$

so it appears that there exists a universal constant $k_1 > 0$ such that

$$\int_{-(1-s)\pi}^{(1-s)\pi} \mu_\beta(dx) \int_0^1 dv \big(g'\big((1+\eta v)x\big)\big)^2 \le k_1 \int_0^1 dv \int_{-(1-s)\pi}^{(1-s)\pi} \lambda(dx) T_{-\eta v}[h](x),$$

where

$$\forall x \in \mathbb{T}, \qquad h(x) := \big[\big(f'(x)\big)^2 + f^2(x)\big]\mu_\beta(x).$$

The proof of Lemma 3.2 shows that for any fixed $v \in [0, 1]$,

$$\begin{aligned}\int_{-(1-s)\pi}^{(1-s)\pi} T_{-\eta v}[h](x)\lambda(dx) &\le \frac{1}{1+v\eta}\int_{\mathbb{T}} h(x)\lambda(dx) \\ &\le \int_{\mathbb{T}} h(x)\lambda(dx) \\ &= \int_{\mathbb{T}} \big(f'\big)^2 d\mu_\beta + \int_{\mathbb{T}} f^2 d\mu_\beta.\end{aligned}$$

Coming back to (3.11) and recalling that $\eta = s/(1-s)$, we obtain that

$$J_2 \le k_2 s^2 \bigg(\int_{\mathbb{T}} \big(f'\big)^2 d\mu_\beta + \int_{\mathbb{T}} f^2 d\mu_\beta\bigg),$$

for another universal constant $k_2 > 0$. This ends the proof of (3.8). □

3.2. Estimate of $L^*_{\alpha,\beta}[\mathbb{1}]$ in the case $p = 1$

When we are interested in finding medians, the definition (3.2) must be modified into

$$\forall x \in \mathbb{T}, \qquad T_{y,s} f(x) := f\big(\gamma(x, y, s)\big). \tag{3.12}$$

Similarly to what we have done in Lemma 3.2, we begin by computing the adjoint $T^\dagger_{y,s}$ of $T_{y,s}$ in $\mathbb{L}^2(\lambda)$, for any fixed $y \in \mathbb{T}$ and $s \in \mathbb{R}_+$ small enough.

Lemma 3.5. *Assume that $s \in [0, \pi/2)$. Then for any bounded and measurable function g, we have, for almost every $x \in \mathbb{T}$ (identified with its representative in $(y-\pi, y+\pi)$),*

$$\begin{aligned}T^\dagger_{y,s}[g](x) = {}& \mathbb{1}_{(y-\pi+s,y-s)}(x)g(x-s) + \mathbb{1}_{(y-s,y+s)}(x)\big(g(x-s)+g(x+s)\big) \\ &+ \mathbb{1}_{(y+s,y+\pi-s)}(x)g(x+s).\end{aligned}$$

Proof. By definition, we have, for any bounded and measurable functions f, g,

$$2\pi \int_{\mathbb{T}} g T_{y,s} f\, d\lambda = \int_{y-\pi}^{y+\pi} g(x) f\big(x + \operatorname{sign}(y-x)s\big)\, dx.$$

Let us first consider the integral

$$\begin{aligned}\int_y^{y+\pi} g(x)f\big(x+\operatorname{sign}(y-x)s\big)\,dx &= \int_y^{y+\pi} g(x)f(x-s)\,dx\\ &= \int_{y-s}^{y+\pi-s} g(x+s)f(x)\,dx\\ &= \int_{y+s}^{y+\pi-s} g(x+s)f(x)\,dx + \int_{y-s}^{y+s} g(x+s)f(x)\,dx.\end{aligned}$$

The symmetrical computation on $(y-\pi, y)$ leads to the announced result. □

It is not difficult to adapt the proof of Lemma 3.3, to get, with the same notation,

Lemma 3.6. *For $\alpha > 0$ and $\beta \geq 0$ such that $\alpha\beta \in [0,\pi)$, the domain of the maximal extension of $L_{\alpha,\beta}$ on $\mathbb{L}^2(\mu_\beta)$ is $\mathcal{D}$. Furthermore, the domain of its dual operator $L^*_{\alpha,\beta}$ in $\mathbb{L}^2(\mu_\beta)$ is $\mathcal{D}^*$ and we have for any $f \in \mathcal{D}^*$,*

$$L^*_{\alpha,\beta} f = \frac{1}{2}\exp(\beta U_1)\partial^2\big[\exp(-\beta U_1)f\big] + \frac{1}{\alpha}\int T^*_{y,(\alpha\beta)/2}[f]\nu(dy) - \frac{f}{\alpha},$$

where

$$T^*_{y,(\alpha\beta)/2}[f] = \exp(\beta U_1)T^\dagger_{y,(\alpha\beta)/2}\big[\exp(-\beta U_1)f\big].$$

In particular, if ν admits a continuous density, then $\mathcal{D}^* = \mathcal{D}$ and the above formula holds for any $f \in \mathcal{D}$.

To be able to consider $L^*_{\alpha,\beta}\mathbb{1}$, we have thus to assume that ν admits a continuous density, so that $\mathbb{1} \in \mathcal{D}^* = \mathcal{D}$. Furthermore, we obtain then that for almost every $x \in \mathbb{T}$,

$$L^*_{\alpha,\beta}\mathbb{1}(x) = \frac{\beta^2}{2}\big(U_1'(x)\big)^2 - \frac{\beta}{2}U_1''(x) + \frac{1}{\alpha}\bigg(\int T^*_{y,(\alpha\beta)/2}[\mathbb{1}](x)\nu(dy) - 1\bigg).$$

By expanding the various terms of the right-hand side, we are to show the equivalent of Proposition 3.1.

Proposition 3.2. *Assume that ν admits a density with respect to λ satisfying (3.5). Then there exists a constant $C(A) > 0$, only depending on A, such that for any $\beta \geq 1$ and $\alpha \in (0, \pi\beta^{-2})$, we have*

$$\big\|L^*_{\alpha,\beta}\mathbb{1}\big\|_\infty \leq C(A)\max\big(\alpha\beta^4, \alpha^a\beta^{1+a}\big).$$

Proof. From (2.5) and Lemma 3.1, we deduce, respectively, that for all $x \in \mathbb{T}$,

$$\begin{aligned} U_1'(x) &= -\int \dot{\gamma}(x,y,0)\nu(dy) \\ &= \nu\big((x-\pi,x)\big) - \nu\big((x,x+\pi)\big), \end{aligned} \tag{3.13}$$

$$U_1''(x) = \big(\nu(x) - \nu(x')\big)/\pi. \tag{3.14}$$

On the other hand, from Lemma 3.5 we get that for all $s \in [0,\pi/2)$ and for almost every $x \in \mathbb{T}$,

$$\begin{aligned} &\int T^*_{y,s}[\mathbb{1}](x)\nu(dy) \\ &\quad = \nu\big((x+s,x+\pi-s)\big)\exp\big(\beta\big(U_1(x)-U_1(x-s)\big)\big) \\ &\qquad + \nu\big((x-s,x+s)\big)\big[\exp\big(\beta\big(U_1(x)-U_1(x-s)\big)\big)+\exp\big(\beta\big(U_1(x)-U_1(x+s)\big)\big)\big] \\ &\qquad + \nu\big((x-\pi+s,x-s)\big)\exp\big(\beta\big(U_1(x)-U_1(x+s)\big)\big) \\ &\quad = \nu\big((x,x+\pi)\big)\exp\big(\beta\big(U_1(x)-U_1(x-s)\big)\big)+\nu\big((x-\pi,x)\big)\exp\big(\beta\big(U_1(x)-U_1(x+s)\big)\big) \\ &\qquad + \nu\big((x-s,x)\big)\exp\big(\beta\big(U_1(x)-U_1(x-s)\big)\big) \\ &\qquad + \nu\big((x,x+s)\big)\exp\big(\beta\big(U_1(x)-U_1(x+s)\big)\big) \\ &\qquad - \nu\big((x'-s,x')\big)\exp\big(\beta\big(U_1(x)-U_1(x-s)\big)\big)-\nu\big((x',x'+s)\big) \\ &\qquad \times \exp\big(\beta\big(U_1(x)-U_1(x+s)\big)\big). \end{aligned}$$

This leads us to define $s = \alpha\beta/2 \in (0,\pi/2)$, so that we can decompose

$$\frac{2}{\beta}L^*_{\alpha,\beta}\mathbb{1}(x) = I_1(x,s)+I_2(x,s)+I_3(x,s),$$

with

$$\begin{aligned} I_1(x,s) &:= \frac{1}{\pi}\left(\pi\frac{\nu((x-s,x+s))}{s}-\nu(x)\right)-\frac{1}{\pi}\left(\pi\frac{\nu((x'-s,x'+s))}{s}-\nu\big(x'\big)\right), \\ I_2(x,s) &:= \frac{\nu((x-s,x))-\nu((x'-s,x'))}{s}\big[\exp\big(\beta\big(U_1(x)-U_1(x-s)\big)\big)-1\big] \\ &\quad + \frac{\nu((x,x+s))-\nu((x',x'+s))}{s}\big[\exp\big(\beta\big(U_1(x)-U_1(x+s)\big)\big)-1\big], \\ I_3(x,s) &:= \nu\big((x,x+\pi)\big)\frac{\exp(\beta(U_1(x)-U_1(x-s)))-1-s\beta U_1'(x)}{s} \\ &\quad + \nu\big((x-\pi,x)\big)\frac{\exp(\beta(U_1(x)-U_1(x+s)))-1+s\beta U_1'(x)}{s}. \end{aligned}$$

Assumption (3.5) enables to evaluate $I_1(x,s)$, because we have for any $x \in \mathbb{T}$ and $s \in (0,\pi/2)$,

$$\begin{aligned}\left|\pi\frac{\nu((x-s,x+s))}{s}-\nu(x)\right| &= \frac{1}{2s}\left|\int_{(x-s,x+s)}\nu(z)-\nu(x)\,dz\right| \\ &\le \frac{A}{2s}\int_{(x-s,x+s)}|z-x|^a\,dz \\ &= \frac{As^a}{1+a} \\ &\le As^a.\end{aligned}$$

By considering the Taylor's expansion with remainder at the first order of the mapping $s \mapsto \exp(\beta[U_1(x)-U_1(x-s)])$ at $s=0$ and by taking into account (3.13), we get for any $x \in \mathbb{T}$ and $s \in (0,\pi/(2\beta))$,

$$\begin{aligned}\left|I_2(x,s)\right| &\le 2\frac{\|\nu\|_\infty}{2\pi}\exp\bigl(\beta\|U_1'\|_\infty s\bigr)\beta\|U_1'\|_\infty s \\ &\le \frac{\|\nu\|_\infty}{\pi}\exp(\beta s)\beta s \\ &\le 2\frac{1+\pi A}{\pi}\exp(\pi/2)\beta s.\end{aligned}$$

The term $I_3(x,s)$ is bounded in a similar manner, rather expanding at the second order the previous mapping and using (3.14) to see that $\|U_1''\|_\infty \le A$. □

We finish this subsection with the a variant of Lemma 3.4.

Lemma 3.7. *There exists a universal constant $k>0$, such that for any $s>0$ and $\beta \ge 1$ with $\beta s \le 1$, we have, for any $f \in \mathcal{C}^1(\mathbb{T})$,*

$$\int_{B(y,\pi-s)}\bigl(T_{y,s}^*[\widetilde{g}_y](x)-g_y(x)\bigr)^2\mu_\beta(dx) \le ks^2\beta^2\left(\int(\partial f)^2\,d\mu_\beta+\int f^2\,d\mu_\beta\right),$$

where $T_{y,s}^*$ is the adjoint operator of $T_{y,s}$ in $\mathbb{L}^2(\mu_\beta)$ and where for any fixed $y \in \mathbb{T}$,

$$\forall x \in \mathbb{T}\setminus\{y'\}, \qquad \begin{cases} g_y(x) := f(x)\dot{\gamma}(x,y,0), \\ \widetilde{g}_y(x) := \mathbb{1}_{(y-\pi,y-s)\sqcup(y+s,y+\pi)}(x)g_y(x). \end{cases}$$

Proof. As remarked at the beginning of the proof of Lemma 3.4, it is sufficient to deal with the case $y=0$. To simplify the notation, we remove $y=0$ from the indices, in particular we consider the mappings g and $\widetilde{g}$ defined by $g(x)=-\operatorname{sign}(x)f(x)$ and $\widetilde{g}(x)=\mathbb{1}_{(-\pi,-s)\sqcup(s,\pi)}(x)g(x)$.

Taking into account that $\widetilde{g}$ vanishes on $(-s,s)$, we deduce from Lemmas 3.5 and 3.6 that for a.e. $x \in (-\pi+s,\pi-s)$,

$$T_s^*[\widetilde{g}](x) = \exp\bigl(\beta U_2(x)\bigr)T_{-s}\bigl[\exp(-\beta U_2)\widetilde{g}\bigr](x).$$

This observation leads us to consider the upper bound

$$\int_{-\pi+s}^{\pi-s} \big(T_s^*[\widetilde{g}](x) - g(x)\big)^2 \mu_\beta(dx) \le 2J_1 + 2J_2,$$

where

$$J_1 := \int_{-\pi+s}^{\pi-s} \big(\exp\big(\beta\big[U_2(x) - U_2\big(x + \mathrm{sign}(x)s\big)\big]\big) - 1\big)^2 \big(T_{-s}[\widetilde{g}]\big)^2 \mu_\beta(dx),$$
$$J_2 := \int_{-\pi+s}^{\pi-s} \big(T_{-s}[\widetilde{g}] - g\big)^2 \, d\mu_\beta.$$

The arguments used in the proof of Lemma 3.4 to deal with J_1 and J_2 can now be easily adapted (even simplified) to obtain the wanted bounds. For instance, one would have noted that

$$J_2 = \int_{-\pi+s}^{0} \big(g(x-s) - g(x)\big)^2 \mu_\beta(dx) + \int_0^{\pi-s} \big(g(x+s) - g(x)\big)^2 \mu_\beta(dx). \qquad \square$$

3.3. Estimate of $L^*_{\alpha,\beta}[\mathbb{1}]$ in the cases $1 < p < 2$

In this situation, for any fixed $y \in \mathbb{T}$ and $s \ge 0$, the definition (3.2) must be replaced by

$$\forall x \in \mathbb{T}, \qquad T_{y,s} f(x) := f\big(\gamma\big(x, y, s d^{p-1}(x, y)\big)\big). \tag{3.15}$$

It leads us to introduce the function z defined on $(y - \pi, y + \pi)$ by

$$z(x) := \begin{cases} x - s(x-y)^{p-1}, & \text{if } x \in [y, y+\pi), \\ x + s(y-x)^{p-1}, & \text{if } x \in (y-\pi, y]. \end{cases} \tag{3.16}$$

To study the variations of this function, by symmetry, it is sufficient to consider its restriction to $(y, y+\pi)$. We need the following definitions, all of them depending on $y \in \mathbb{T}$, $s \ge 0$ and $p \in (1, 2)$:

$$u_+ := y + (p-1)^{1/(2-p)} s^{1/(2-p)},$$
$$\widetilde{u}_+ := y + s^{1/(2-p)},$$
$$v_+ := y - \big((p-1)^{(p-1)/(2-p)} - (p-1)^{1/(2-p)}\big) s^{1/(2-p)},$$
$$w_+ := y + \pi - \pi^{p-1} s.$$

Let $\sigma(p)$ be the largest positive real number in $(0, 1/2)$ such that for $s \in (0, \sigma(p))$, we have $u_+ < y+\pi$, $v_+ > y - \pi$ and $w_+ - y > y - v_+$. One checks that for $s \in (0, \sigma(p))$, the function z is decreasing on (y, u_+) and increasing on $(u_+, y+\pi)$. Furthermore $v_+ = z(u_+)$, $w_+ = z(y+\pi)$

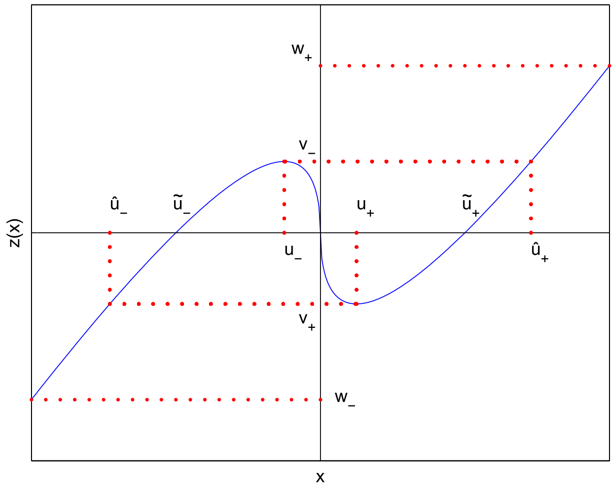

Figure 3. The function z.

and $\widetilde{u}_+$ is the unique point in $(u_+, y+\pi)$ such that $z(\widetilde{u}_+) = y$. Let us also introduce $\widehat{u}_+$ the unique point in $(\widetilde{u}_+, y+\pi)$ such that and $z(\widehat{u}_+) = -v_+$. All these definitions, as well as the symmetric notions with respect to (y, y), where the indices $+$ are replaced by $-$, are summarized in the following picture (see Figure 3).

Thus for $s \in (0, \sigma(p))$, we can consider $\varphi_+ : [v_+, y] \to [y, u_+]$ and $\psi_+ : [v_+, w_+] \to [u_+, y+\pi]$ the inverses of z, respectively, restricted to $[y, u_+]$ and $[u_+, y+\pi]$. The mappings φ_- and ψ_- are defined in a symmetrical manner on $[y, v_-]$ and $[w_-, v_-]$. These quantities were necessary to compute the adjoint $T^\dagger_{y,s}$ of $T_{y,s}$ in $\mathbb{L}^2(\lambda)$, for any fixed $y \in \mathbb{T}$ and $s > 0$ small enough.

Lemma 3.8. *Assume that $s \in (0, \sigma(p))$. Then for any bounded and measurable function g, we have, for almost every $x \in \mathbb{T}$ (identified with its representative in $(y-\pi, y+\pi)$),*

$$\begin{aligned} T^\dagger_{y,s}[g](x) &= \mathbb{1}_{(w_-,v_+)}(x)\psi'_-(x)g\big(\psi_-(x)\big) + \mathbb{1}_{(v_-,w_+)}(x)\psi'_+(x)g\big(\psi_+(x)\big) \\ &\quad + \mathbb{1}_{(v_+,y)}(x)\big[\psi'_-(x)g\big(\psi_-(x)\big) + \psi'_+(x)g\big(\psi_+(x)\big) + \big|\varphi'_+(x)\big|g\big(\varphi_+(x)\big)\big] \\ &\quad + \mathbb{1}_{(y,v_-)}(x)\big[\psi'_-(x)g\big(\psi_-(x)\big) + \psi'_+(x)g\big(\psi_+(x)\big) + \big|\varphi'_-(x)\big|g\big(\varphi_-(x)\big)\big]. \end{aligned}$$

Proof. The above formula is based on straightforward applications of the change of variable formula. For instance one can write for any bounded and measurable functions f, g defined on $(y-\pi, y+\pi)$,

$$\int_{(y,u_+)} g(x) f\big(T_{y,s}(x)\big)\,dx = \int_{(v_+,y)} f(z) g\big(\varphi_+(z)\big)\big|\varphi'_+(z)\big|\,dz. \qquad \square$$

Since we are more interested in adjoint operators in $\mathbb{L}^2(\mu_\beta)$, let us define for any fixed $y \in \mathbb{T}$, $s \in (0, \sigma(p))$ and any bounded and measurable function f defined on $(y - \pi, y + \pi)$,

$$T^*_{y,s}[f] := \exp(\beta U_p) T^\dagger_{y,s}\big[\exp(-\beta U_p) f\big]. \tag{3.17}$$

Then we get the equivalent of Lemmas 3.3 and 3.6.

Lemma 3.9. *For $\alpha > 0$ and $\beta > 0$ such that $s := p\alpha\beta/2 \in (0, \sigma(p))$, the domain of the maximal extension of $L_{\alpha,\beta}$ on $\mathbb{L}^2(\mu_\beta)$ is $\mathcal{D}$. Furthermore, the domain of its dual operator $L^*_{\alpha,\beta}$ in $\mathbb{L}^2(\mu_\beta)$ is $\mathcal{D}^*$ and we have for any $f \in \mathcal{D}^*$,*

$$L^*_{\alpha,\beta} f = \frac{1}{2} \exp(\beta U_p) \partial^2 \big[\exp(-\beta U_p) f\big] + \frac{1}{\alpha} \int T^*_{y,s}[f] \nu(dy) - \frac{f}{\alpha}.$$

In particular, if ν admits a continuous density, then $\mathcal{D}^* = \mathcal{D}$ and the above formula holds for any $f \in \mathcal{D}$.

Once again, the assumption that ν admits a continuous density enables us to consider $L^*_{\alpha,\beta}\mathbb{1}$, which is given, under the conditions of the previous lemma, for almost every $x \in \mathbb{T}$, by

$$L^*_{\alpha,\beta}\mathbb{1}(x) = \frac{\beta^2}{2}\big(U'_p(x)\big)^2 - \frac{\beta}{2} U''_p(x) + \frac{1}{\alpha}\left(\int T^*_{y,(p\alpha\beta)/2}[\mathbb{1}](x)\nu(dy) - 1\right). \tag{3.18}$$

We deduce the following.

Proposition 3.3. *Assume that ν admits a density with respect to λ satisfying (3.5). Then there exists a constant $C(A, p) > 0$, only depending on $A > 0$ and $p \in (1, 2)$, such that for any $\beta \geq 1$ and $\alpha \in (0, \sigma(p)/\beta^2)$, we have*

$$\big\| L^*_{\alpha,\beta}\mathbb{1}\big\|_\infty \leq C(A, p) \max\big(\alpha\beta^4, \alpha^{p-1}\beta^{1+p}, \alpha^a \beta^{1+a}\big).$$

Proof. We first keep in mind that from (2.5) and Lemma 3.1, we have for all $x \in \mathbb{T}$,

$$U'_p(x) = p\left(\int_{x-\pi}^{x} (x-y)^{p-1}\nu(dy) - \int_{x}^{x+\pi} (y-x)^{p-1}\nu(dy)\right), \tag{3.19}$$

$$U''_p(x) = p(p-1)\int_{\mathbb{T}} d^{p-2}(y, x)\nu(dy) - p\pi^{p-2}\nu\big(x'\big). \tag{3.20}$$

Taking into account (3.18), our goal is to see how the terms $\beta(U'_p(x))^2$ and $-U''_p(x)$ cancel with some parts of the integral

$$\frac{p}{s}\int T^*_{y,s}[\mathbb{1}](x) - 1\nu(dy),$$

where $s := p\alpha\beta/2 \in (0, \sigma(p)/\beta) \subset (0, \sigma(p))$, and to bound what remains by a quantity of the form $C'(A, p)(\beta^2 s + \beta s^{p-1} + s^a)$, for another constant $C'(A, p) > 0$, only depending on $A > 0$ and $p \in (1, 2)$.

We decompose the domain of integration of $\nu(dy)$ into six essential parts (with the convention that $-\pi \le y - x < \pi$ and remember that the points w_-, v_+, v_- and w_+ depend on y):

$$
\begin{aligned}
J_1 &:= \{y \in \mathbb{T} : y - \pi < x < w_-\},\\
J_2 &:= \{y \in \mathbb{T} : w_- < x < v_+\},\\
J_3 &:= \{y \in \mathbb{T} : v_+ < x < y\},\\
J_4 &:= \{y \in \mathbb{T} : y < x < v_-\},\\
J_5 &:= \{y \in \mathbb{T} : v_- < x < w_+\},\\
J_6 &:= \{y \in \mathbb{T} : w_+ < x < y + \pi\}.
\end{aligned}
$$

The cases of J_1 and J_6 are the simplest to treat. For instance, for J_6, we write that

$$
\begin{aligned}
\frac{p}{s}\int_{J_6} T^*_{y,s}[\mathbb{1}](x) - 1\nu(dy) &= -\frac{p}{s}\int_{x'}^{x'+\pi^{p-1}s} \mathbb{1}\nu(dy)\\
&= -\frac{p}{s}\int_{x'}^{x'+\pi^{p-1}s} \nu(y)\frac{dy}{2\pi}\\
&= -\frac{p\pi^{p-2}}{2}\nu\bigl(x'\bigr) - \frac{p}{2\pi s}\int_{x'}^{x'+\pi^{p-1}s} \nu(y) - \nu\bigl(x'\bigr)\,dy.
\end{aligned}
$$

A similar computation for J_1 and the use of assumption (3.5) lead to the bound

$$
\begin{aligned}
\left|\frac{p}{s}\int_{J_1\sqcup J_6} T^*_{y,s}[\mathbb{1}](x) - 1\nu(dy) + p\pi^{p-2}\nu\bigl(x'\bigr)\right| &\le Ap\frac{\pi^{(1+a)(p-1)-1}}{1+a}s^a\\
&\le 2\pi As^a.
\end{aligned}
\tag{3.21}
$$

The most important parts correspond to J_2 and J_5. For example, considering J_5, which can be written down as the segment (x_-, x_+), with

$$
\begin{aligned}
x_- &:= x - \pi + \pi^{p-1}s,\\
x_+ &:= x - \bigl((p-1)^{(p-1)/(2-p)} - (p-1)^{1/(2-p)}\bigr)s^{1/(2-p)},
\end{aligned}
$$

we have to evaluate the integral

$$
\frac{p}{s}\int_{x_-}^{x_+} \psi'_+(x)\exp\bigl(\beta\bigl[U_p(x) - U_p\bigl(\psi_+(x)\bigr)\bigr]\bigr) - 1\nu(dy) \tag{3.22}
$$

(y is present in the integrand through $\psi_+(x)$ and $\psi'_+(x)$). Indeed, in view of (3.19) and (3.20), we would like to compare it to

$$
-\beta U'_p(x)\int_{x_-}^{x_+}(x-y)^{p-1}\nu(dy) + p(p-1)\int_{x_-}^{x_+}(x-y)^{p-2}\nu(dy). \tag{3.23}
$$

To do so, we will expand the terms $\psi'_+(x)$ and $\exp(\beta[U_p(x) - U_p(\psi_+(x))])$ as functions of the (hidden) parameter $s > 0$. Fix $y \in J_5$ and recall that it amounts to $x \in (v_-, w_+)$. Due to (3.16) and to the definition of ψ_+, we have for such x,

$$\psi'_+(x) = \frac{1}{1 - s(p-1)(\psi_+(x) - y)^{p-2}}. \tag{3.24}$$

Let us begin by working heuristically, to outline why the quantities (3.22) and (3.23) should be close. From the above expression, we get

$$\psi'_+(x) \simeq 1 + s(p-1)\big(\psi_+(x) - y\big)^{p-2}.$$

By definition of ψ_+, we have

$$\begin{aligned} x - y &= \psi_+(x) - y - s\big(\psi_+(x) - y\big)^{p-1} \\ &= \big(\psi_+(x) - y\big)\big(1 - s\big(\psi_+(x) - y\big)^{p-2}\big), \end{aligned} \tag{3.25}$$

so that $x - y \simeq \psi_+(x) - y$ and

$$\psi'_+(x) \simeq 1 + s(p-1)(x - y)^{p-2}.$$

On the other hand,

$$\begin{aligned} \exp\big(\beta\big[U_p(x) - U_p\big(\psi_+(x)\big)\big]\big) &\simeq 1 + \beta\big[U_p(x) - U_p\big(\psi_+(x)\big)\big] \\ &\simeq 1 + \beta U'_p(x)\big(x - \psi_+(x)\big) \\ &= 1 - s\beta U'_p(x)\big(\psi_+(x) - y\big)^{p-1} \\ &\simeq 1 - s\beta U'_p(x)(x - y)^{p-1}. \end{aligned}$$

Putting together these approximations, we end up with

$$\psi'_+(x)\exp\big(\beta\big[U_p(x) - U_p\big(\psi_+(x)\big)\big]\big) - 1 \simeq s\big[(p-1)(x-y)^{p-2} - \beta U'_p(x)(x-y)^{p-1}\big],$$

suggesting the proximity of (3.22) and (3.23), after integration with respect to $\nu(dy)$ on (x_-, x_+).

To justify and quantify these computations, we start by remarking that $\psi_+(x) - y$ is bounded below by $\widehat{u}_+ - y$, itself bounded below by $\widetilde{u}_+ - y = s^{1/(2-p)}$. But this lower bound will not be sufficient in (3.25), so let us improve it a little. By definition of $\widehat{u}_+$, we have

$$v_- - y = \widehat{u} - y - s(\widehat{u} - y)^{p-1},$$

so that $\widehat{u}_+ - y = k_p s^{1/(2-p)}$ where k_p is the unique solution larger than 1 of the equation

$$k_p - k_p^{p-1} = (p-1)^{(p-1)/(2-p)} - (p-1)^{1/(2-p)}. \tag{3.26}$$

It follows that for any $y \in J_5$,

$$\begin{aligned}1 &\le \frac{1}{1-s(\psi_+(x)-y)^{p-2}} \le \frac{1}{1-s(\widehat{u}_+-y)^{p-2}}\\ &= \frac{\widehat{u}_+-y}{v_--y}\\ &= K_p,\end{aligned} \tag{3.27}$$

where the latter quantity only depends on $p \in (1,2)$ and is given by

$$K_p := \frac{k_p}{(p-1)^{(p-1)/(2-p)}-(p-1)^{1/(2-p)}}.$$

In particular, coming back to (3.24) and taking into account (3.25), we get that for $y \in J_5'$,

$$\begin{aligned}\left|\psi_+'(x)-1-s(p-1)\big(\psi_+(x)-y\big)^{p-2}\right| &= \frac{(s(p-1)(\psi_+(x)-y)^{p-2})^2}{1-s(p-1)(\psi_+(x)-y)^{p-2}}\\ &\le (p-1)^2s^2\frac{(\psi_+(x)-y)^{2(p-2)}}{1-s(\psi_+(x)-y)^{p-2}}\\ &= (p-1)^2s^2\frac{(x-y)^{2(p-2)}}{(1-s(\psi_+(x)-y)^{p-2})^{1+2(p-2)}}\\ &\le (p-1)^2K_p^{(2p-3)_+}s^2(x-y)^{2(p-2)}.\end{aligned}$$

To complete this estimate, we note that in a similar way, still for $y \in J_5$,

$$\begin{aligned}\left|\big(\psi_+(x)-y\big)^{p-2}-(x-y)^{p-2}\right| &= (x-y)^{p-2}\left|1-\big(1-s\big(\psi_+(x)-y\big)^{p-2}\big)^{2-p}\right|\\ &\le (x-y)^{p-2}\left|1-\big(1-s\big(\psi_+(x)-y\big)^{p-2}\big)\right|\\ &= s(x-y)^{p-2}\big(\psi_+(x)-y\big)^{p-2}\\ &= s(x-y)^{2(p-2)}\big(1-s\big(\psi_+(x)-y\big)^{p-2}\big)^{2-p}\\ &\le s(x-y)^{2(p-2)},\end{aligned}$$

so that in the end,

$$\left|\psi_+'(x)-1-s(p-1)(x-y)^{p-2}\right| \le \big[(p-1)^2K_p^{(2p-3)_+}+p-1\big]s^2(x-y)^{2(p-2)}. \tag{3.28}$$

We now come to the term $\exp(\beta[U_p(x)-U_p(\psi_+(x))])$. First we remark that

$$\begin{aligned}\left|U_p(x)-U_p\big(\psi_+(x)\big)\right| &\le \|U_p'\|_\infty\left|x-\psi_+(x)\right|\\ &\le p\pi^{p-1}s\big(\psi_+(x)-y\big)^{p-1}\end{aligned}$$

$$\leq p\pi^{2(p-1)}s$$
$$\leq 2\pi^2 s.$$

It follows, recalling our assumption $\beta s \leq \sigma(p)$, that

$$\begin{aligned}&\left|\exp\big(\beta\big[U_p(x)-U_p\big(\psi_+(x)\big)\big]\big)-1-\beta\big[U_p(x)-U_p\big(\psi_+(x)\big)\big]\right|\\&\quad\leq\frac{\beta^2[U_p(x)-U_p(\psi_+(x))]^2}{2}\exp\big(2\pi^2\beta s\big)\\&\quad\leq 2\pi^4\beta^2\exp\big(2\pi^2\sigma(p)\big)s^2.\end{aligned}$$

In addition, we have

$$\left|U_p(x)-U_p\big(\psi_+(x)\big)-U_p'(x)\big(x-\psi_+(x)\big)\right|\leq\frac{\|U_p''\|_\infty}{2}\big(x-\psi_+(x)\big)^2.$$

In view of (3.20) and taking into account that $\int U_p''\,d\lambda=0$, we have

$$\begin{aligned}\left\|U_p''\right\|_\infty&\leq 2p(p-1)\|\nu\|_\infty\int_0^\pi u^{p-2}\frac{du}{2\pi}\\&=2p\pi^{p-1}(1+\pi A).\end{aligned}$$

So we get

$$\begin{aligned}\left|U_p(x)-U_p\big(\psi_+(x)\big)-U_p'(x)\big(x-\psi_+(x)\big)\right|&\leq 2\pi(1+\pi A)\big(x-\psi_+(x)\big)^2\\&\leq 2\pi(1+\pi A)s^2\big(\psi_+(x)-y\big)^{2(p-1)}\\&\leq 2\pi^3(1+\pi A)s^2,\end{aligned}$$

namely

$$\left|U_p(x)-U_p\big(\psi_+(x)\big)+sU_p'(x)\big(\psi_+(x)-y\big)^{p-1}\right|\leq 2\pi^3 s^2.$$

Finally, using the inequality

$$\forall u,v\geq 0,\forall p\in(1,2),\qquad\left|u^{p-1}-v^{p-1}\right|\leq|u-v|^{p-1},$$

it appears that

$$\begin{aligned}\left|\big(\psi_+(x)-y\big)^{p-1}-(x-y)^{p-1}\right|&\leq\left|\psi_+(x)-x\right|^{p-1}\\&=\left|\psi_+(x)-y\right|^{(p-1)^2}s^{p-1}\\&\leq\pi^{(p-1)^2}s^{p-1},\end{aligned}\tag{3.29}$$

so we can deduce that

$$\begin{aligned}&\left|\exp\big(\beta\big[U_p(x)-U_p\big(\psi_+(x)\big)\big]\big)-1+\beta sU_p'(x)(x-y)^{p-1}\right|\\&\quad\le p\pi^pK_p\beta s^p+2\pi^3\beta\big(1+\pi A+\pi\exp\big(2\pi^2\sigma(p)\big)\beta\big)s^2.\end{aligned}$$

From the latter bound and (3.28), we obtain a constant $K(p,A)>0$ depending only on $p\in(1,2)$ and $A>0$, such that

$$\begin{aligned}&\frac{p}{s}\bigg|\int_{x_-}^{x_+}\psi_+'(x)\exp\big(\beta\big[U_p(x)-U_p\big(\psi_+(x)\big)\big]\big)\\&\qquad-\left(1+\frac{s(p-1)}{(x-y)^{2-p}}\right)\big(1-\beta sU_p'(x)(x-y)^{p-1}\big)\nu(dy)\bigg|\\&\quad\le K(p,A)\left(\beta s^{p-1}+\beta^2 s+s\int_{x_-}^{x_+}(x-y)^{2(p-2)}\nu(dy)\right).\end{aligned}\tag{3.30}$$

This leads us to upper bound

$$\begin{aligned}\int_{x_-}^{x_+}(x-y)^{2(p-2)}\nu(dy)&\le\frac{\|\nu\|_\infty}{2\pi}\int_{x_-}^{x_+}(x-y)^{2(p-2)}\,dy\\&\le\frac{1+A\pi}{2\pi}\int_{\kappa_p s^{1/(2-p)}}^{\pi-\pi^{p-1}s}y^{2(p-2)}\,dy,\end{aligned}$$

with

$$\kappa_p:=(p-1)^{(p-1)/(2-p)}-(p-1)^{1/(2-p)}.\tag{3.31}$$

An immediate computation gives, for $p\in(1,2)$, a constant $\kappa_p'>0$ such that for any $s\in(0,\sigma(p))$,

$$\int_{\kappa_p s^{1/(2-p)}}^{\pi-\pi^{p-1}s}y^{2(p-2)}\,dy\le\kappa_p'\begin{cases}1, & \text{if } p>3/2,\\ \ln\big(\big(1+\sigma(p)\big)/s\big), & \text{if } p=3/2,\\ s^{(2p-3)/(2-p)}, & \text{if } p<3/2.\end{cases}\tag{3.32}$$

Since $1+\frac{2p-3}{2-p}>p-1$, $\beta\ge1$ and $s\in(0,\sigma(p))$, we can find another constant $K'(p,A)>0$ such that the right-hand side of (3.30) can be replaced by $K'(p,A)(\beta s^{p-1}+\beta^2 s)$. It is now easy to see that such an expression, up to a new change of the factor $K'(p,A)$, bounds the difference between (3.22) and (3.23). Indeed, just use that

$$\int_{\kappa_p s^{1/(2-p)}}^{\pi-\pi^{p-1}s}y^{2p-3}\,dy\le\pi\int_{\kappa_p s^{1/(2-p)}}^{\pi-\pi^{p-1}s}y^{2(p-2)}\,dy$$

and resort to (3.32).

There is no more difficulty in checking that the cost of replacing x_- and x_+, respectively, by $x-\pi$ and x in (3.23) is also bounded by $K''(p,A)(\beta s^{p/(2-p)}+s^{(p-1)/(2-p)})\le 2K''(p,A)\beta s^{p-1}$, for an appropriate choice of the factor $K''(p,A)$ depending on $p\in(1,2)$ and $A>0$.

Symmetrical computations for J_2 and remembering (3.21) lead to the existence of a constant $K'''(p,A)>0$, depending only on $p\in(1,2)$ and $A>0$, such that for $\beta\ge 1$ and $s\in(0,\sigma(p)/\beta)$, we have

$$\begin{aligned}&\left|\beta\bigl(U_p'(x)\bigr)^2-U_p''(x)+\frac{p}{s}\left(\int_{J_1\sqcup J_2\sqcup J_5\sqcup J_6}T^*_{y,s}[\mathbb{1}](x)\nu(dy)-1\right)\right|\\&\quad\le K'''(p,A)\bigl(s^a+\beta s^{p-1}+\beta^2 s\bigr).\end{aligned}$$

It remains to treat the segments J_3 and J_4 and again by symmetry, let us deal with J_4 only: it is sufficient to exhibit a constant $K^{(4)}(p,A)>0$, depending on $p\in(1,2)$ and $A>0$, such that for $\beta\ge 1$ and $s\in(0,\sigma(p)/\beta)$,

$$\frac{p}{s}\left|\int_{J_4}T^*_{y,s}[\mathbb{1}](x)-1\nu(dy)\right|\le K^{(4)}(p,A)s^{(p-1)/(2-p)}$$

(since the right-hand side is itself bounded by $K^{(4)}(p,A)(\sigma(p))^{(p-1)^2/(2-p)}s^{p-1}$), or equivalently

$$\left|\int_{J_4}T^*_{y,s}[\mathbb{1}](x)-1\nu(dy)\right|\le\frac{K^{(4)}(p,A)}{p}s^{1/(2-p)}. \tag{3.33}$$

The constant part is immediate to bound:

$$\begin{aligned}\int_{J_4}1\nu(dy)&\le\frac{\|\nu\|_\infty}{2\pi}\int_{J_4}1\,dy\\&\le\frac{1+\pi A}{2\pi}\int_{x-\kappa_p s^{1/(2-p)}}^{x}1\,dy\\&=\frac{(1+\pi A)\kappa_p}{2\pi}s^{1/(2-p)}.\end{aligned}$$

For the other part, we first remark that for $y\in J_4$, we have

$$\begin{aligned}y&<x<y+\kappa_p s^{1/(2-p)},\\y+s^{1/(2-p)}&<\psi_+(x)<y+k_p s^{1/(2-p)},\\y-(p-1)^{1/(2-p)}s^{1/(2-p)}&<\varphi_-(x)<y,\\y-s^{1/(2-p)}&<\psi_-(x)<y-(p-1)^{1/(2-p)}s^{1/(2-p)}\end{aligned}$$

(recall that $\widehat{u}_+ = y + k_p s^{1/(2-p)}$ with k_p defined in (3.26)). It follows that we can find a constant $\kappa''_p > 0$, depending only on $p \in (1,2)$, such that for $s \in (0, \sigma(p))$,

$$\begin{aligned}
&\max\bigl(\bigl|U_p(x) - U_p\bigl(\psi_+(x)\bigr)\bigr|, \bigl|U_p(x) - U_p\bigl(\psi_-(x)\bigr)\bigr|, \bigl|U_p(x) - U_p\bigl(\varphi_-(x)\bigr)\bigr|\bigr) \\
&\quad \le \kappa''_p s^{1/(2-p)} \\
&\quad \le \kappa''_p \bigl(\sigma(p)\bigr)^{(p-1)/(2-p)} s.
\end{aligned}$$

In particular, we can find another constant $\kappa'''_p > 0$, such that under the conditions that $\beta \ge 1$ and $\beta s \in (0, \sigma(p))$,

$$\exp\bigl(\beta \max\bigl(\bigl|U_p(x) - U_p\bigl(\psi_+(x)\bigr)\bigr|, \bigl|U_p(x) - U_p\bigl(\psi_-(x)\bigr)\bigr|, \bigl|U_p(x) - U_p\bigl(\varphi_-(x)\bigr)\bigr|\bigr)\bigr) \le \kappa'''_p.$$

Thus, denoting ψ one of the functions ψ_+, φ_- or ψ_-, and remembering the bound $\|\nu\|_\infty \le 1 + \pi A$, it is sufficient to exhibit another constant $\kappa^{(4)}_p > 0$ such that

$$\int_{J_4} \bigl|\psi'(x)\bigr|\, dy \le \kappa^{(4)}_p s^{1/(2-p)}. \tag{3.34}$$

Let us consider the case $\psi = \psi_+$, the other functions admit a similar treatment. We begin by making the dependence of $\psi_+(x)$ more explicit by writing it $\psi_+(x, y)$. From the definition of this quantity (see the first line of (3.25)) and from (3.24), we get

$$\begin{aligned}
\partial_y \psi_+(x, y) &= -\frac{s(p-1)(\psi_+(x, y) - y)^{p-2}}{1 - s(p-1)(\psi_+(x, y) - y)^{p-2}} \\
&= -s(p-1)\bigl(\psi_+(x, y) - y\bigr)^{p-2} \partial_x \psi_+(x, y),
\end{aligned}$$

so that the left-hand side of (3.34) can be rewritten

$$\begin{aligned}
&\frac{1}{s(p-1)} \int_{J_4} \bigl|\bigl(\psi_+(x, y) - y\bigr)^{2-p} \partial_y \psi_+(x, y)\bigr|\, dy \\
&\quad \le \frac{1}{s(p-1)} \int_{J_4} \bigl(k_p s^{1/(2-p)}\bigr)^{2-p} \bigl|\partial_y \psi_+(x, y)\bigr|\, dy \\
&\quad \le \frac{k_p^{2-p}}{(p-1)} \int_{J_4} \bigl|\partial_y \psi_+(x, y)\bigr|\, dy.
\end{aligned}$$

Checking that $J_4 = (x - \kappa_p s^{1/(2-p)}, x)$, the last integral is equal to $|\psi_+(x, x) - \psi_+(x, x - \kappa_p s^{1/(2-p)})|$. By definition of ψ_+, we have $\psi_+(x, x) = x$ and it appears that the quantity $\zeta := \psi_+(x, x - \kappa_p s^{1/(2-p)}) - x$ is a positive solution to the equation

$$\zeta = s\bigl(\zeta + \kappa_p s^{1/(2-p)}\bigr)^{p-1}.$$

It follows that $\zeta = k'_p s^{1/(2-p)}$ where k'_p is the unique positive solution of $k'_p = (k'_p + \kappa_p)^{p-1}$.

Thus, (3.34) is proven and we can conclude to the validity of (3.33). □

To finish this subsection, here is a version of Lemma 3.7 for $p \in (1,2)$, which is a little weaker, since we need a preliminary integration with respect to $\nu(y)$.

Lemma 3.10. *Under the assumption (3.5), there exists a universal constant $k(p,A) > 0$, depending only on $p \in (1,2)$ and $A > 0$, such that for any $s > 0$ and $\beta \geq 1$ with $\beta s \leq \sigma(p)$, we have, for any $f \in \mathcal{C}^1(\mathbb{T})$,*

$$\begin{aligned}&\int_{\mathbb{T}} \nu(dy) \int_{B(y,\pi-\pi^{p-1}s)} \left(T^*_{y,s}[\widetilde{g}_y](x) - g_y(x)\right)^2 \mu_\beta(dx) \\ &\quad \leq k(p,A)\left(s^{2(p-1)} + \beta^2 s^2\right)\left(\int (\partial f)^2 \, d\mu_\beta + \int f^2 \, d\mu_\beta\right),\end{aligned} \tag{3.35}$$

where $T^*_{y,s}$ is the adjoint operator of $T_{y,s}$ in $\mathbb{L}^2(\mu_\beta)$ and where for any fixed $y \in \mathbb{T}$,

$$\forall x \in \mathbb{T} \setminus \{y'\}, \qquad \begin{cases} g_y(x) := f(x) d^{p-1}(x,y)\dot{\gamma}(x,y,0), \\ \widetilde{g}_y(x) := \mathbb{1}_{(y-\pi, y-s^{1/(2-p)}) \sqcup (y+s^{1/(2-p)}, y+\pi)}(x) g_y(x). \end{cases}$$

Proof. We begin by fixing $y \in \mathbb{T}$ and by remembering the notation of the proof of Proposition 3.3 (see Figure 3). Due to fact that $\widetilde{g}_y$ vanishes on $(\widetilde{u}_-, \widetilde{u}_+) = (y - s^{1/(2-p)}, y + s^{1/(2-p)})$, we deduce from Lemma 3.8 and (3.17) that for a.e. $x \in (y - \pi + \pi^{p-1}s, y + \pi - \pi^{p-1}s)$,

$$T^*_{y,s}[\widetilde{g}_y](x) = \psi'_\varepsilon(x) \exp\left(\beta\left[U_p(x) - U_p\left(\psi_\varepsilon(x)\right)\right]\right) \widetilde{g}_y\left(\psi_\varepsilon(x)\right),$$

where $\varepsilon \in \{-,+\}$ stands for the sign of $x - y$ with the conventions of the proof of Proposition 3.3. Thus, we are led to the decomposition

$$\int_{B(y,\pi-\pi^{p-1}s)} \left(T^*_{y,s}[\widetilde{g}_y](x) - g_y(x)\right)^2 \mu_\beta(dx) \leq 3J_1(y) + 3J_2(y) + 3J_3(y),$$

where

$$\begin{aligned} J_1(y) &:= \int_{B(y,\pi-\pi^{p-1}s)} \left(\exp\left(\beta\left[U_p(x) - U_p\left(\psi_\varepsilon(x)\right)\right]\right) - 1\right)^2 \left(\psi'_\varepsilon(x) \widetilde{g}_y\left(\psi_\varepsilon(x)\right)\right)^2 \mu_\beta(dx), \\ J_2(y) &:= \int_{B(y,\pi-\pi^{p-1}s)} \left(\psi'_\varepsilon(x)\right)^2 \left(\widetilde{g}_y\left(\psi_\varepsilon(x)\right) - g_y(x)\right)^2 \mu_\beta(dx), \\ J_3(y) &:= \int_{B(y,\pi-\pi^{p-1}s)} \left(\psi'_\varepsilon(x) - 1\right)^2 g_y^2(x) \mu_\beta(dx). \end{aligned}$$

We begin by dealing with $J_1(y)$, or rather with just half of it, by symmetry and to avoid the consideration of ε:

$$\int_y^{y+\pi-\pi^{p-1}s} \left(\exp\left(\beta\left[U_p(x) - U_p\left(\psi_+(x)\right)\right]\right) - 1\right)^2 \left(\psi'_+(x) \widetilde{g}_y\left(\psi_+(x)\right)\right)^2 \mu_\beta(dx).$$

Let us recall that $x = \psi_+(x) - s(\psi_+(x) - y)^{p-1}$ and that $\psi_+(x) - y \geq s^{1/(2-p)}$. From (3.24), we deduce that for $x \in (y, y + \pi - \pi^{p-1}s)$, $1 \leq \psi_+(x) \leq 1/(2-p)$. Thus, it is sufficient to bound

$$\int_y^{y+\pi-\pi^{p-1}s} \big(\exp\big(\beta\big[U_p(x) - U_p\big(\psi_+(x)\big)\big]\big) - 1\big)^2 \big(\widetilde{g}_y\big(\psi_+(x)\big)\big)^2 \mu_\beta(dx).$$

Furthermore, for $x \in (y, y + \pi - \pi^{p-1}s)$, we have

$$\big|x - \psi_+(x)\big| \leq s\pi^{p-1}, \tag{3.36}$$

so under the assumption that $s\beta \in (0, 1/2)$, we can bound $(\exp(\beta[U_p(x) - U_p(\psi_+(x))]) - 1)^2$ by a term of the form $k\beta^2 s^2$ for a universal constant $k > 0$. It remains to use $\widetilde{g}_y^2(x) \leq \pi^2 f^2(x)$ to get an upper bound going in the direction of (3.35).

We now come to $J_2(y)$ and again only to half of it:

$$\int_y^{y+\pi-\pi^{p-1}s} \big(\psi_+'(x)\big)^2 \big(\widetilde{g}_y\big(\psi_+(x)\big) - g_y(x)\big)^2 \mu_\beta(dx).$$

Due to the upper bound on ψ_+ seen just above, it is sufficient to deal with

$$\int_y^{y+\pi-\pi^{p-1}s} \big(\widetilde{g}_y\big(\psi_+(x)\big) - g_y(x)\big)^2 \mu_\beta(dx).$$

But for $x \in (y, y + \pi - \pi^{p-1}s)$, we have $\psi_+(x) \in (y + s^{1/(2-p)}, y + \pi)$, so that $\widetilde{g}_y(\psi_+(x)) = g_y(\psi_+(x))$ and the above expression is equal to

$$\int_y^{y+\pi-\pi^{p-1}s} \big(g_y\big(\psi_+(x)\big) - g_y(x)\big)^2 \mu_\beta(dx).$$

Coming back to the definition of g_y, it appears that for $x \in (y, y + \pi - \pi^{p-1}s)$, both $\psi_+(x)$ and x belong to the same hemicircle obtained by cutting $\mathbb{T}$ at y and y', so

$$\begin{aligned}
&\big(g_y\big(\psi_+(x)\big) - g_y(x)\big)^2 \\
&\quad = \big(d^{p-1}\big(y, \psi_+(x)\big) f\big(\psi_+(x)\big) - d^{p-1}(y, x) f(x)\big)^2 \\
&\quad \leq 2d^{2(p-1)}\big(y, \psi_+(x)\big)\big(f\big(\psi_+(x)\big) - f(x)\big)^2 + 2f^2(x)\big(d^{p-1}\big(y, \psi_+(x)\big) - d^{p-1}(y, x)\big)^2 \\
&\quad \leq 2\pi^{2(p-1)}\big(f\big(\psi_+(x)\big) - f(x)\big)^2 + 2\pi^{2(p-1)^2} s^{2(p-1)} f^2(x),
\end{aligned}$$

where we have used (3.29) to majorize the last term. From (3.36), we deduce that

$$\big(f\big(\psi_+(x)\big) - f(x)\big)^2 \leq 2s\pi^{p-1} \int_{x-s\pi^{p-1}}^{x+s\pi^{p-1}} \big(f'(z)\big)^2 \, dz.$$

As usual, the assumption $0 < s\beta \le 1/2$ enables to find a universal constant $k > 0$ such that for any $z \in (x - s\pi^{p-1}, x + s\pi^{p-1})$, we have $\mu_\beta(x) \le k\mu_\beta(z)$. From the above computations, it follows there exists another universal constant $k' > 0$ such that for any $y \in \mathbb{T}$,

$$\begin{aligned}
J_2(y) &\le k'\left(s^{2(p-1)} \int f^2 \, d\mu_\beta + s^2 \int (f')^2 \, d\mu_\beta\right) \\
&\le k' s^{2(p-1)}\left(\int f^2 \, d\mu_\beta + \int (f')^2 \, d\mu_\beta\right).
\end{aligned}$$

Finally, we come to $J_3(y)$, which will need to be integrated with respect to $\nu(dy)$. From (3.24), we first get that

$$\begin{aligned}
J_3(y) &= \int_{B(y,\pi-\pi^{p-1}s)} \left(\frac{s(p-1)d^{p-2}(\psi_\varepsilon(x), y)}{1 - s(p-1)d^{p-2}(\psi_\varepsilon(x), y)}\right)^2 g_y^2(x)\mu_\beta(dx) \\
&\le \frac{(p-1)^2}{(2-p)^2} s^2 \int_{B(y,\pi-\pi^{p-1}s)} d^{2(p-2)}\big(\psi_\varepsilon(x), y\big) g_y^2(x)\mu_\beta(dx) \\
&\le \frac{\pi^{2(p-1)}(p-1)^2}{(2-p)^2} s^2 \int_{B(y,\pi-\pi^{p-1}s)} d^{2(p-2)}\big(\psi_\varepsilon(x), y\big) f^2(x)\mu_\beta(dx).
\end{aligned}$$

Next, recalling that $\|\nu\|_\infty \le 1 + \pi A$ and that $d(\psi_\varepsilon(x), y) \ge s^{1/(2-p)}$ for any $x \in B(y, \pi - \pi^{p-1}s)$, it appears that

$$\begin{aligned}
&\int_{\mathbb{T}} J_3(y)\nu(dy) \\
&\quad \le \frac{1+\pi A}{2\pi} \frac{\pi^{2(p-1)}(p-1)^2}{(2-p)^2} s^2 \int_{\mathbb{T}} dy \int_{B(y,\pi-\pi^{p-1}s)} d^{2(p-2)}\big(\psi_\varepsilon(x), y\big) f^2(x)\mu_\beta(dx) \\
&\quad \le \frac{1+\pi A}{2\pi} \frac{\pi^{2(p-1)}(p-1)^2}{(2-p)^2} s^2 \int_{\mathbb{T}} \mu_\beta(dx) f^2(x) \\
&\qquad \times \int_{\mathbb{T}} \mathbb{1}_{\{d(\psi_\varepsilon(x),y) \ge s^{1/(2-p)}\}} d^{2(p-2)}\big(\psi_\varepsilon(x), y\big)\, dy.
\end{aligned}$$

But for any fixed $z \in \mathbb{R}/(2\pi\mathbb{Z})$, we compute that

$$\begin{aligned}
\int_{\mathbb{T}} \mathbb{1}_{\{d(z,y) \ge s^{1/(2-p)}\}} d^{2(p-2)}(z, y)\, dy &= 2\int_{s^{1/(2-p)}}^{\pi} \frac{1}{y^{2(2-p)}}\, dy \\
&\le k''_p \begin{cases} 1, & \text{if } p > 3/2, \\ \ln(1/s), & \text{if } p = 3/2, \\ s^{(2p-3)/(2-p)}, & \text{if } p < 3/2, \end{cases}
\end{aligned}$$

for $s \in (0, 1/2)$ and for an appropriate constant $k''_p > 0$ depending only on $p \in (1, 2)$. It is not difficult to check that as $s \to 0_+$, we have

$$s^{2(p-1)} \gg \begin{cases} s^2, & \text{if } p > 3/2, \\ s^2 \ln(1/s), & \text{if } p = 3/2, \\ s^2 s^{(2p-3)/(2-p)}, & \text{if } p < 3/2. \end{cases}$$

It follows that for any $p \in (1, 2)$, we can find a constant $k'(p, A) > 0$, depending only on $p \in (1, 2)$ and $A > 0$, such that

$$\int_{\mathbb{T}} J_3(y)\nu(dy) \le k'(p, A)s^{2(p-1)} \int_{\mathbb{T}} f^2(x)\mu_\beta(dx).$$

This ends the proof of the estimate (3.35). □

3.4. Estimate of $L^*_{\alpha,\beta}[\mathbb{1}]$ in the cases $p > 2$

This situation is simpler than the one treated in the previous subsection and is similar to the case $p = 2$, because for $y \in T$ fixed and $s \ge 0$ small enough, the mapping z defined in (3.16) is injective when $p > 2$. Again for any fixed $y \in \mathbb{T}$ and $s \ge 0$, the definition (3.2) has to be replaced by (3.15), namely,

$$\forall x \in \mathbb{T}, \qquad T_{y,s} f(x) := f\big(z(x)\big). \tag{3.37}$$

With the previous subsections in mind, the computations are quite straightforward, so we will just outline them.

The first task is to determine the adjoint $T^\dagger_{y,s}$ of $T_{y,s}$ in $\mathbb{L}^2(\lambda)$. An immediate change of variable gives that for any $s \in (0, \sigma)$, for any bounded and measurable function g, we have, for almost every $x \in \mathbb{T}$ (identified with its representative in $(y - \pi, y + \pi)$),

$$T^\dagger_{y,s}[g](x) = \mathbb{1}_{(y,z(y))}(x)\psi'(x)g\big(\psi(x)\big),$$

where $\sigma := \pi^{2-p}/(p - 1)$ and $\psi : (z(y - \pi), z(y + \pi)) \to (y - \pi, y + \pi)$ is the inverse mapping of z (with the slight abuses of notation: $z(y - \pi) := x - \pi + \pi^{p-1}s$, $z(y + \pi) := x + \pi - \pi^{p-1}s$). The adjoint $T^*_{y,s}$ of $T_{y,s}$ in $\mathbb{L}^2(\mu_\beta)$ is still given by (3.17). As in the previous subsections, this operator is bounded in $\mathbb{L}^2(\mu_\beta)$. It follows, if ν admits a continuous density with respect to λ and at least for $\alpha > 0$ and $\beta \ge 0$ such that $s := (p/2)\alpha\beta \in [0, \sigma)$, that the adjoint $L^*_{\alpha,\beta}$ of $L_{\alpha,\beta}$ in $\mathbb{L}^2(\mu_\beta)$ is defined on $\mathcal{D}$. In particular, we can consider $L^*_{\alpha,\beta}\mathbb{1}$, which is given, for almost every $x \in \mathbb{T}$, by

$$L^*_{\alpha,\beta}\mathbb{1}(x) = \frac{\beta^2}{2}\big(U'_p(x)\big)^2 - \frac{\beta}{2}U''_p(x) + \frac{p\beta}{2s}\bigg(\int T^*_{y,s}[\mathbb{1}](x)\nu(dy) - 1\bigg). \tag{3.38}$$

From this formula, we deduce the following.

Proposition 3.4. *Assume that ν admits a density with respect to λ satisfying* (3.5). *Then there exists a constant $C(A,p)>0$, only depending on $A>0$ and $p>2$, such that for any $\beta\geq 1$ and $\alpha\in(0,\sigma/(p\beta^2))$, we have*

$$\left\|L^*_{\alpha,\beta}\mathbb{1}\right\|_\infty\leq C(A,p)\max\left(\alpha\beta^4,\alpha^a\beta^{1+a}\right).$$

Proof. The arguments are similar to those of the case J_5 in the proof of Proposition 3.3, but are less involved, because the omnipresent term $1-s(p-1)(\psi(x)-y)^{p-2}$ is now easy to bound: for any $s\in[0,\sigma/2]$, we have for any $y\in\mathbb{T}$ and $x\in(z(y-\pi),z(y+\pi))$,

$$\tfrac{1}{2}\leq 1-(p-1)\left|\psi(x)-y\right|^{p-2}s\leq 1.$$

In particular, we have under these conditions,

$$\psi'(x)=\frac{1}{1-(p-1)|\psi(x)-y|^{p-2}s}\in[1,2].$$

Following the arguments of the previous subsection, one finds a constant $K(p,A)$, depending only on $p>2$ and $A>0$, such that for any $\beta\geq 1$, $s\in[0,\sigma/(2\beta)]$ and $x\in(z(y-\pi),z(y+\pi))$,

$$\begin{aligned}\left|\psi'_+(x)-1-(p-1)\left|\psi_+(x)-y\right|^{p-2}s\right|&\leq K(p,A)s^2,\\ \left|\exp\left(\beta\left[U_p(x)-U_p\left(\psi_+(x)\right)\right]\right)-1+\beta\operatorname{sign}(x-y)U'_p(x)|x-y|^{p-1}s\right|&\leq K(p,A)\beta^2s^2.\end{aligned}$$

This bound enables us to approximate $T^*_{\alpha,\beta}\mathbb{1}(x)-1$ up to a term $\mathcal{O}_{p,A}(\beta^2s^2)$ (recall that this designates a quantity which is bounded by an expression of the form $K'(p,A)\beta^2s^2$ for a constant $K'(p,A)>0$ depending on $p>2$ and $A>0$), by

$$\left((p-1)\left|\psi_+(x)-y\right|^{p-2}-\beta\operatorname{sign}(x-y)U'_p(x)|x-y|^{p-1}\right)s.$$

Next, we consider

$$\begin{aligned}J&:=\left\{y\in\mathbb{T}:x\in\left(z(y-\pi),z(y+\pi)\right)\right\}\\&=\mathbb{T}\setminus\left[x'-s\pi^{p-1},x'+s\pi^{p-1}\right],\end{aligned}\tag{3.39}$$

in order to decompose

$$\begin{aligned}&\frac{p\beta}{2s}\int_{\mathbb{T}}T^*_{y,s}[\mathbb{1}](x)-1\nu(dy)\\&\quad=\frac{p\beta}{2s}\int_J T^*_{y,s}[\mathbb{1}](x)-1\nu(dy)-\frac{p\beta}{2s}\nu\left(\left[x'-s\pi^{p-1},x'+s\pi^{p-1}\right]\right).\end{aligned}\tag{3.40}$$

According to the previous estimate, up to a term $\mathcal{O}_{p,A}(\beta^3s^2)$ the first integral is equal to

$$\frac{p(p-1)\beta}{2}\int_J d^{p-2}(y,x)\nu(dy)-\frac{p\beta^2}{2}U'_p(x)\int_J\operatorname{sign}(x-y)d^{p-1}(x,y)\nu(dy).$$

In view of (3.39), up to an additional term $\mathcal{O}_{p,A}(\beta^2 s)$, we can replace J in the above integrals by $\mathbb{T}$. Thus putting together (3.38) and (3.40) with (3.19) and (3.20) (which are also valid here), it remains to estimate

$$\frac{p\beta}{2}\left|\pi^{p-2}\nu(x') - \frac{1}{s}\nu\left[x' - s\pi^{p-1}, x' + s\pi^{p-1}\right]\right|$$

and this is easily done through the assumption (3.5). □

We finish this subsection with the equivalent of Lemma 3.4.

Lemma 3.11. *For $p > 2$, there exists a constant $k(p) > 0$, depending only on $p > 2$, such that for any $s \in (0, \sigma)$, with $\sigma := \pi^{2-p}/(p-1)$, and $\beta \geq 1$ with $\beta s \leq 1$, we have, for any $y \in \mathbb{T}$ and $f \in \mathcal{C}^1(\mathbb{T})$,*

$$\int_{B(y,\pi - s\pi^{p-1})} \left(T^*_{y,s}[g_y](x) - g_y(x)\right)^2 \mu_\beta(dx) \leq k(p) s^2 \beta^2 \left(\int (\partial f)^2\, d\mu_\beta + \int f^2\, d\mu_\beta\right),$$

where $T^*_{y,s}$ is the adjoint operator of $T_{y,s}$ in $\mathbb{L}^2(\mu_\beta)$ and where for any fixed $y \in \mathbb{T}$,

$$\forall x \in \mathbb{T} \setminus \{y'\}, \qquad g_y(x) := f(x) d^{p-1}(x, y)\dot{\gamma}(x, y, 0).$$

Proof. We only sketch the arguments, which are just an adaptation of those of the proof of Lemma 3.4. Again it is sufficient to deal with the case $y = 0$, which is removed from the notation, and consequently with the function $g(x) = -\operatorname{sign}(x)|x|^{p-1} f(x)$. As seen previously in this subsection, we have for $s \in (0, \sigma)$ and $x \in (-\pi, \pi)$,

$$T^*_s[g](x) = \mathbb{1}_{(-\pi + s\pi^{p-1}, \pi - s\pi^{p-1})}(x) \exp\left(\beta\left[U_p(x) - U_p\left(\psi(x)\right)\right]\right)\psi'(x) g\left(\psi(x)\right),$$

where ψ is the inverse mapping of $(-\pi, \pi) \ni x \mapsto x - \operatorname{sign}(x)|x|^{p-1}$. Recall that for $x \in (-\pi + s\pi^{p-1}, \pi - s\pi^{p-1})$,

$$\psi'(x) = \frac{1}{1 - (p-1)|\psi(x)|^{p-2} s} \in [1, 2]. \tag{3.41}$$

Considering the decomposition

$$\begin{aligned} T^*_s[g](x) - g(x) &= \left(\exp\left(\beta\left[U_p(x) - U_p\left(\psi(x)\right)\right]\right) - 1\right)\psi'(x) g\left(\psi(x)\right) \\ &\quad + \psi'(x)\left(g\left(\psi(x)\right) - g(x)\right) + \left(\psi'(x) - 1\right) g(x), \end{aligned}$$

we are led, after integration with respect to $\mathbb{1}_{(-\pi + s\pi^{p-1}, \pi - s\pi^{p-1})}(x)\mu_\beta(dx)$, to computations similar to those of Sections 3.1 and 3.3, and indeed simpler than in the latter one, due to the boundedness property described in (3.41). □

Let us summarize the Propositions 3.1, 3.2, 3.3 and 3.4 of the previous subsections into the statement.

Proposition 3.5. *Assume that* (3.5) *is satisfied and for* $p \geq 1$, *consider the constant* $a(p) > 0$ *defined in* (1.5). *Then there exists two constants* $\sigma(p) \in (0, 1/2)$ *and* $C(A, p) > 0$, *depending only on the quantities inside the parentheses, such that for any* $\alpha > 0$ *and* $\beta > 1$ *such that* $\alpha\beta < \sigma(p)$, *we have*

$$\sqrt{\mu_\beta\big[\big(L^*_{\alpha,\beta}\mathbb{1}\big)^2\big]} \leq C(A, p)\alpha^{a(p)}\beta^4.$$

Despite this bound is very rough, since we have replaced an essential norm by a $\mathbb{L}^2$ norm, it will be sufficient in the next section, when $\alpha^{a(p)}\beta^4$ is small, as a measure of the discrepancy between μ_β and the invariant measure for $L_{\alpha,\beta}$.

4. Proof of convergence

This is the main part of the paper: we are going to prove Theorem 1.1 by the investigation of the evolution of a $\mathbb{L}^2$ type functional.

On $\mathbb{T}$ consider the algorithm $X := (X_t)_{t\geq 0}$ described in the Introduction. We require that the underlying probability measure ν admits a density with respect to λ which is Hölder continuous: $a \in (0, 1]$ and $A > 0$ are constants such that (3.5) is satisfied. For the time being, the schemes $\alpha : \mathbb{R}_+ \to \mathbb{R}^*_+$ and $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ are assumed to be, respectively, continuous and continuously differentiable. Only later on, in Proposition 4.3, will we present the conditions insuring the wanted convergence (1.4). On the initial distribution m_0, the last ingredient necessary to specify the law of X, no hypothesis is made. We also denote m_t the law of X_t, for any $t > 0$. From the lemmas given in the Appendix, we have that m_t admits a $\mathcal{C}^1$ density with respect to λ, which is equally written m_t. As it was mentioned in the previous section, we want to compare these temporal marginal laws with the corresponding instantaneous Gibbs measures, which were defined in (2.2) with respect to the potential U_p given in (1.1). A convenient way to quantify this discrepancy is to consider the variance of the density of m_t with respect to μ_{β_t} under the probability measure μ_{β_t}:

$$\forall t > 0, \qquad I_t := \int \left(\frac{m_t}{\mu_{\beta_t}} - 1\right)^2 d\mu_{\beta_t}. \tag{4.1}$$

Our goal here is to derive a differential inequality satisfied by this quantity, which implies its convergence to zero under appropriate conditions on the schemes α and β. More precisely, our purpose is to obtain the following.

Proposition 4.1. *There exists two constants* $c_1(p, A), c_2(p, A) > 0$, *depending on* $p \geq 1$ *and* $A > 0$, *and a constant* $\varsigma(p) \in (0, 1/2)$, *depending on* $p \geq 1$, *such that for any* $t > 0$ *with* $\beta_t \geq 1$ *and* $0 < \alpha_t\beta_t^2 \leq \varsigma(p)$, *we have*

$$I'_t \leq -c_1(p, A)\big(\beta_t^{-3}\exp\big(-b(U_p)\beta_t\big) - \alpha_t^{\widetilde{a}(p)}\beta_t^3 - \big|\beta'_t\big|\big)I_t + c_2(p, A)\big(\alpha_t^{a(p)}\beta_t^4 + \big|\beta'_t\big|\big)\sqrt{I_t},$$

where $b(U_p)$ *was defined in* (1.6), $a(p)$ *in Proposition* 3.5 *and*

$$\widetilde{a}(p) := \begin{cases} 1, & \text{if } p = 1 \text{ or } p \geq 3/2, \\ 2(p-1), & \text{if } p \in (1, 3/2). \end{cases}$$

At least formally, there is no difficulty to differentiate the quantity I_t with respect to the time $t > 0$. But we postpone the rigorous justification of the following computations to the end of the Appendix, where the regularity of the temporal marginal laws is discussed in detail. Thus, we get at any time $t > 0$,

$$\begin{aligned} I_t' &= 2\int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)\frac{\partial_t m_t}{\mu_{\beta_t}}\, d\mu_{\beta_t} - 2\int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)\frac{m_t}{\mu_{\beta_t}}\partial_t \ln(\mu_{\beta_t})\, d\mu_{\beta_t} \\ &\quad + \int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)^2 \partial_t \ln(\mu_{\beta_t})\, d\mu_{\beta_t} \\ &= 2\int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)\partial_t m_t \, d\lambda - \int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)^2 \partial_t \ln(\mu_{\beta_t})\, d\mu_{\beta_t} - 2\int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)\partial_t \ln(\mu_{\beta_t})\, d\mu_{\beta_t} \\ &\leq 2\int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)\partial_t m_t \, d\lambda + \left\|\partial_t \ln(\mu_{\beta_t})\right\|_\infty\left(\int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)^2 d\mu_{\beta_t} + 2\int\left|\frac{m_t}{\mu_{\beta_t}} - 1\right| d\mu_{\beta_t}\right) \\ &\leq 2\int\left(\frac{m_t}{\mu_{\beta_t}} - 1\right)\partial_t m_t \, d\lambda + \left\|\partial_t \ln(\mu_{\beta_t})\right\|_\infty\left(I_t + 2\sqrt{I_t}\right), \end{aligned}$$

where we used the Cauchy–Schwarz inequality. The last term is easy to deal with.

Lemma 4.1. *For any* $t \geq 0$, *we have*

$$\left\|\partial_t \ln(\mu_{\beta_t})\right\|_\infty \leq \pi^p \left|\beta_t'\right|.$$

Proof. Since for any $t \geq 0$, we have

$$\forall x \in \mathbb{T}, \qquad \ln(\mu_{\beta_t}) = -\beta_t U_p(x) - \ln\left(\int \exp\left(-\beta_t U_p(y)\right)\lambda(dy)\right),$$

it appears that

$$\forall x \in \mathbb{T}, \qquad \partial_t \ln(\mu_{\beta_t}) = \beta_t' \int U_p(y) - U_p(x)\mu_{\beta_t}(dy),$$

so that

$$\left\|\partial_t \ln(\mu_{\beta_t})\right\|_\infty \leq \operatorname{osc}(U_p)\left|\beta_t'\right|.$$

The bound $\operatorname{osc}(U_p) \leq \pi^p$ is an immediate consequence of the definition (1.1) of U_p and of the fact that the (intrinsic) diameter of $\mathbb{T}$ is π. □

Denote for any $t>0$, $f_t := m_t/\mu_{\beta_t}$. If this function was to be $\mathcal{C}^2$, we would get, by the martingale problem satisfied by the law of X, that

$$\begin{aligned}\int\left(\frac{m_t}{\mu_{\beta_t}}-1\right)\partial_t m_t\,d\lambda &= \int L_{\alpha_t,\beta_t}[f_t-1]\,dm_t\\ &= \int L_{\alpha_t,\beta_t}[f_t-1]f_t\,d\mu_{\beta_t},\end{aligned}$$

where L_{α_t,β_t}, described in the previous section, is the instantaneous generator at time $t\geq 0$ of X. The interest of the estimate of Proposition 3.5 comes from the decomposition of the previous term into

$$\begin{aligned}&\int L_{\alpha_t,\beta_t}[f_t-1](f_t-1)\,d\mu_{\beta_t}+\int L_{\alpha_t,\beta_t}[f_t-1]\,d\mu_{\beta_t}\\ &\quad=\int L_{\alpha_t,\beta_t}[f_t-1](f_t-1)\,d\mu_{\beta_t}+\int (f_t-1)L^*_{\alpha_t,\beta_t}[\mathbb{1}]\,d\mu_{\beta_t}\\ &\quad\leq\int L_{\alpha_t,\beta_t}[f_t-1](f_t-1)\,d\mu_{\beta_t}+\sqrt{I_t}\sqrt{\mu_{\beta_t}\left[\left(L^*_{\alpha_t,\beta_t}[\mathbb{1}]\right)^2\right]}.\end{aligned}$$

It follows that to prove Proposition 4.1, it remains to treat the first term in the above right-hand side. A first step is the following.

Lemma 4.2. *There exist a constant $c_3(p,A)>0$, depending on $p\geq 1$ and $A>0$ and a constant $\widetilde{\sigma}(p)\in(0,1/2)$, such that for any $\alpha>0$ and $\beta\geq 1$ such that $\alpha\beta^2\leq\widetilde{\sigma}(p)$, we have, for any $f\in\mathcal{C}^2(\mathbb{T})$,*

$$\begin{aligned}&\int L_{\alpha,\beta}[f-1](f-1)\,d\mu_\beta\\ &\quad\leq-\left(\frac{1}{2}-c_3(p,A)\alpha^{\widetilde{a}(p)}\beta^3\right)\int(\partial f)^2\,d\mu_\beta+c_3(p,A)\alpha^{\widetilde{a}(p)}\beta^3\int(f-1)^2\,d\mu_\beta,\end{aligned}$$

where $\widetilde{a}(p)$ is defined in Proposition 4.1.

Proof. For any $\alpha>0$ and $\beta\geq 0$, we begin by decomposing the generator $L_{\alpha,\beta}$ into

$$L_{\alpha,\beta}=L_\beta+R_{\alpha,\beta},\tag{4.2}$$

where $L_\beta:=(\partial^2-\beta U_p'\partial)/2$ was defined in (2.1) (recall that U_p' is well-defined, since ν has no atom) and where $R_{\alpha,\beta}$ is the remaining operator. An immediate integration by parts leads to

$$\begin{aligned}\int L_\beta[f-1](f-1)\,d\mu_\beta &= -\frac{1}{2}\int\left(\partial(f-1)\right)^2\,d\mu_\beta\\ &= -\frac{1}{2}\int(\partial f)^2\,d\mu_\beta.\end{aligned}$$

Thus, our main task is to find constants $c_3(p, A) > 0$ and $\widetilde{\sigma}(p) \in (0, 1/2)$ such that for any $\alpha > 0$ and $\beta \geq 1$ with $\alpha\beta^2 \leq \widetilde{\sigma}(p)$, we have, for any $f \in \mathcal{C}^2(\mathbb{T})$,

$$\left|\int R_{\alpha,\beta}[f-1](f-1)\,d\mu_\beta\right| \leq c_3(p, A)\alpha^{\widetilde{a}(p)}\beta^3\left(\int (\partial f)^2\,d\mu_\beta + \int (f-1)^2\,d\mu_\beta\right). \quad (4.3)$$

By definition, we have for any $f \in \mathcal{C}^2(\mathbb{T})$ (but what follows is valid for $f \in \mathcal{C}^1(\mathbb{T})$),

$$R_{\alpha,\beta}[f](x) = \frac{1}{\alpha}\int f\big(\gamma\big(x, y, (p/2)\alpha\beta d^{p-1}(x, y)\big)\big) - f(x)\nu(dy) + \frac{\beta}{2}U_p'(x)f'(x) \qquad \forall x \in \mathbb{T}.$$

To evaluate this quantity, on one hand, recall that we have for any $x \in \mathbb{T}$,

$$U_p'(x) = -p\int_{\mathbb{T}} d^{p-1}(x, y)\dot{\gamma}(x, y, 0)\nu(dy)$$

and on the other hand, write that for any $x \in \mathbb{T}$ and $y \in \mathbb{T} \setminus \{x\}$,

$$\begin{aligned} &f\big(\gamma\big(x, y, (p/2)\alpha\beta d^{p-1}(x, y)\big)\big) - f(x) \\ &\quad = \frac{p}{2}\alpha\beta\int_0^1 f'\big(\gamma\big(x, y, (p/2)\alpha\beta d(x, y)u\big)\big)d^{p-1}(x, y)\dot{\gamma}(x, y, 0)\,du. \end{aligned}$$

Writing $s := (p/2)\alpha\beta$ and considering again the operators introduced in (3.15) (now for any $p \geq 1$), it follows that

$$\begin{aligned} &\int R_{\alpha,\beta}[f-1](f-1)\,d\mu_\beta \\ &\quad = \frac{p\beta}{2}\int_0^1 du\int \nu(dy)\int \mu_\beta(dx)\big(T_{y,su}[f'](x) - f'(x)\big)\big(f(x)-1\big)d^{p-1}(x, y)\dot{\gamma}(x, y, 0) \\ &\quad = \frac{p\beta}{2}\int_0^1 du\int \nu(dy)\int \mu_\beta(dx)\big(T_{y,su}[f'](x) - f'(x)\big)g_y(x), \end{aligned}$$

where for any fixed $y \in \mathbb{T}$,

$$\forall x \in \mathbb{T} \setminus \{y\}, \qquad g_y(x) := \big(f(x) - 1\big)d^{p-1}(x, y)\dot{\gamma}(x, y, 0) \quad (4.4)$$

(with, e.g., the convention that $g_y(y') := 0$). Let us also fix the variable $u \in [0, 1]$ for a while.

We begin by considering the case where $p \geq 2$. By definition of $T^*_{y,su}$ (discussed in Section 3), we have

$$\begin{aligned} \int \big(T_{y,su}[f'](x) - f'(x)\big)g_y(x)\mu_\beta(dx) &= \int f'(x)\big(T^*_{y,su}[g_y](x) - g_y(x)\big)\mu_\beta(dx) \\ &= I_1(y, u) + I_2(y, u), \end{aligned} \quad (4.5)$$

where for any $y \in \mathbb{T}$,

$$\begin{aligned} I_1(y,u) &:= \int_{B(y,\pi - su\pi^{p-1})} f'(x)\big(T^*_{y,su}[g_y](x) - g_y(x)\big)\mu_\beta(dx),\\ I_2(y,u) &:= -\int_{B(y',su\pi^{p-1})} f'(x) g_y(x)\mu_\beta(dx) \end{aligned} \tag{4.6}$$

(recall from Sections 3.1 and 3.4 that for any measurable function g, $T^*_{y,s}[g]$ vanishes on $B(y', su\pi^{p-1})$). The first integral is treated through the Cauchy–Schwarz inequality,

$$\big|I_1(y,u)\big| \le \sqrt{\int \big(f'\big)^2\, d\mu_\beta} \sqrt{\int_{B(y,\pi - su\pi^{p-1})} \big(T^*_{y,su}[g_y] - g_y\big)^2 \mu_\beta}$$

and Lemmas 3.4 and 3.11, at least if $s\beta > 0$ is smaller than a certain constant $\widetilde{\sigma}(p) \in (0, /12)$. It follows that for a universal constant $k > 0$, we have

$$\begin{aligned} \int_{\mathbb{T}\times[0,1]} \big|I_1(y,u)\big|\nu(dy)\, du &\le ks^2\beta^2\bigg(\int (\partial f)^2\, d\mu_\beta + \int (f-1)^2\, d\mu_\beta\bigg)\int_0^1 u^2\, du\\ &= \frac{k}{2}s^2\beta^2\bigg(\int (\partial f)^2\, d\mu_\beta + \int f^2\, d\mu_\beta\bigg)\\ &\le \frac{k}{4}s\beta\bigg(\int (\partial f)^2\, d\mu_\beta + \int f^2\, d\mu_\beta\bigg), \end{aligned}$$

bound going in the direction of (4.3).

Next, we turn to the integral $I_2(y,u)$. We cannot deal with it uniformly over $y \in \mathbb{T}$ but we get a convenient bound by integrating it with respect to $\nu(dy)$. Recalling that under the assumption (3.5) the density of ν with respect to λ is bounded by $1 + A\pi$, it appears that

$$\begin{aligned} \int \big|I_2(y,u)\big|\nu(dy) &\le \frac{1 + A\pi}{2\pi}\int_{-\pi}^{\pi} \big|I_2(y,u)\big|\, dy\\ &\le \frac{1 + A\pi}{2\pi}\int_{\mathbb{T}} dy \int_{B(y',su\pi^{p-1})} \big|f'(x)\big|\big|g_y(x)\big|\mu_\beta(dx)\\ &\le \frac{1 + A\pi}{2}\pi^{p-2}\int_{\mathbb{T}} \mu_\beta(dx)\big|f'(x)\big|\big|f(x) - 1\big|\int_{B(x',su\pi^{p-1})} \mathbb{1}\, dy\\ &= (1 + A\pi)\pi^{2p-3} su \int_{\mathbb{T}} \big|f'\big||f - 1|\, d\mu_\beta. \end{aligned} \tag{4.7}$$

The Cauchy–Schwarz inequality and integration with respect to $\mathbb{1}_{[0,1]}(u)\, du$ lead again to a bound contributing to (4.3).

It is time to consider the cases where $p \in [1, 2)$. We will rather decompose the left-hand side of (4.5) into three parts. Let us extend the notation $\widetilde{u}_\pm := y \pm (su)^{1/(2-p)}$ from Section 3.3 to all $p \in$

$[1,2)$. Next, we modify the definition (4.4) by introducing $\widetilde{g}_y(x) := \mathbb{1}_{[y-\pi,\widetilde{u}_-]\sqcup[\widetilde{u}_+,y+\pi]}(x)g_y(x)$. Then we write

$$\int \big(T_{y,su}\big[f'\big](x) - f'(x)\big)g_y(x)\mu_\beta(dx) = \widetilde{I}_1(y,u) + I_2(y,u) + I_3(y,u),$$

where

$$\widetilde{I}_1(y,u) := \int_{B(y,\pi - su\pi^{p-1})} f'(x)\big(T^*_{y,su}[\widetilde{g}_y](x) - g_y(x)\big)\mu_\beta(dx),$$

$$I_2(y,u) := -\int_{B(y',su\pi^{p-1})} f'(x)g_y(x)\mu_\beta(dx),$$

$$I_3(y,u) := \int_{[\widetilde{u}_-,\widetilde{u}_+]} T_{y,su}\big[f'\big](x)g_y(x)\mu_\beta(dx).$$

The treatment of $\widetilde{I}_1(y,u)$ is similar to that of $I_1(y,u)$, with Lemmas 3.7 and 3.10 (where a preliminary integration with respect to $\nu(dy)$ was necessary) replacing Lemmas 3.4 and 3.11.

Concerning $I_2(y,u)$, it is bounded in the same manner as the corresponding quantity defined in (4.6).

It seems that the most convenient way to deal with $I_3(y,u)$ is to first integrate it with respect to $\mathbb{1}_{[0,1]}(u)\nu(dy)\,du$. Taking into account that $\|\nu\|_\infty \le (1 + A\pi)$ and using Cauchy–Schwarz inequality, we get

$$\begin{aligned}
&\int \big|I_3(y,u)\big|\mathbb{1}_{[0,1]}(u)\nu(y)\,du \\
&\quad\le \frac{1+A\pi}{2\pi}\int \big|I_3(y,u)\big|\mathbb{1}_{[0,1]}(u)\,dy\,du \\
&\quad\le \frac{1+A\pi}{2\pi}\sqrt{\int \mathbb{1}_{[\widetilde{u}_-,\widetilde{u}_+]}(x)\big(T_{y,su}\big[f'\big](x)\big)^2\mathbb{1}_{[0,1]}(u)\mu_\beta(dx)\,dy\,du} \\
&\qquad\times\sqrt{\int \mathbb{1}_{[\widetilde{u}_-,\widetilde{u}_+]}(x)g_y^2(x)\mathbb{1}_{[0,1]}(u)\mu_\beta(dx)\,dy\,du}.
\end{aligned}$$

The last factor can be rewritten under the form

$$\begin{aligned}
&\sqrt{\int \mu_\beta(dx)\int \mathbb{1}_{[x-s^{1/(2-p)},x+s^{1/(2-p)}]}(y)g_y^2(x)\,dy} \\
&\quad\le \pi^{p-1}\sqrt{\int \mu_\beta(dx)\big(f(x)-1\big)^2\int_{x-s^{1/(2-p)}}^{x+s^{1/(2-p)}} dy} \qquad (4.8)\\
&\quad= \pi\sqrt{2s^{1/(2-p)}}\sqrt{\int (f-1)^2\,d\mu_\beta}.
\end{aligned}$$

So it remains to consider the term

$$\begin{aligned}&\int \mathbb{1}_{[\widetilde{u}_-,\widetilde{u}_+]}(x)\big(T_{y,su}\big[f'\big](x)\big)^2\mathbb{1}_{[0,1]}(u)\mu_\beta(dx)\,dy\,du\\&\quad=\frac{1}{2\pi}\int \mathbb{1}_{[\widetilde{u}_-,\widetilde{u}_+]}(x)T_{y,su}\big[\big(f'\big)^2\big](x)\mu_\beta(x)\mathbb{1}_{[0,1]}(u)\,dy\,du\end{aligned}\tag{4.9}$$

(where as a function, μ_β stands for the density of the measure μ_β with respect to λ). Remember that for any measurable function h, we have $T_{y,su}[h](x) := h(x + sud^{p-1}(x, y) \times \dot\gamma(x, y, 0))$. For $x \in [\widetilde{u}_-, \widetilde{u}_+]$, we have $d(x, y) \le (su)^{1/(2-p)}$ and it follows that $d(x, x + sud^{p-1}(x, y)\dot\gamma(x, y, 0)) \le (su)^{(3-p)/(2-p)}$. Taking into account that $\|U_p'\|_\infty \le \pi^{p-1}$, we can then a universal constant $k > 0$ such that for $0 \le s\beta \le \widetilde\sigma(p)$ (for an appropriate constant $\widetilde\sigma(p) \in (0, 1/2))$ and $x \in \mathbb{T}$, we have $\mu_\beta(x)/\mu_\beta(x + sud^{p-1}(x, y)\dot\gamma(x, y, 0)) \le k$. This leads us to consider the function h defined by

$$\forall x \in \mathbb{T}, \qquad h(x) := \big(f'(x)\big)^2\mu_\beta(x),\tag{4.10}$$

since up to a universal constant, we have to find an upper bound of

$$\begin{aligned}&\int \mathbb{1}_{[\widetilde{u}_-,\widetilde{u}_+]}(x)T_{y,su}[h](x)\mathbb{1}_{[0,1]}(u)\,dx\,dydu\\&\quad\le\int_{-\pi}^{\pi}dx\int_{x-s^{1/(2-p)}}^{x+s^{1/(2-p)}}dy\int_{x-sd^{p-1}(x,y)}^{x+sd^{p-1}(x,y)}h(v)\frac{dv}{sd^{p-1}(x,y)}\\&\quad=\int_{\mathbb{T}}H(v)h(v)\,dv,\end{aligned}$$

where for any fixed $v \in \mathbb{T}$,

$$H(v) := \frac{1}{s}\int_{\mathbb{T}^2}\mathbb{1}_{\{d(x,y)\le s^{1/(2-p)},d(v,x)\le sd^{p-1}(x,y)\}}\frac{dx\,dy}{d^{p-1}(x,y)}.$$

Let us furthermore fix $x \in \mathbb{T}$,

$$\frac{1}{s}\int_{\mathbb{T}}\mathbb{1}_{\{(d(v,x)/s)^{1/(p-1)}\le d(x,y)\le s^{1/(2-p)}\}}\frac{dy}{d^{p-1}(x,y)}=\frac{2}{(2-p)s}\left(s-\left(\frac{d(v,x)}{s}\right)^{(2-p)/(p-1)}\right)_+.$$

The integration of the last right-hand side with respect to dx is bounded above by

$$\frac{2}{2-p}\int_0^{(s^{1/(2-p)})^{p-1}s}dx=\frac{2}{2-p}s^{1/(2-p)}.$$

Thus, we have found a constant $k(p) > 0$ depending on $p \in [1, 2)$ such that (4.9) is bounded above by $k(p)s^{1/(2-p)}$ under our conditions on $s > 0$ and $\beta \ge 1$. In conjunction with (4.8) and definition (4.10), it enables to conclude to the existence of a constant $k(p, A) > 0$, depending on $p \in [1, 2)$ and $A > 0$, such that

$$\int \left|I_3(y, u)\right| \mathbb{1}_{[0,1]}(u)\nu(y)\, du \le k(p, A)s^{1/(2-p)} \sqrt{\int (f-1)^2\, d\mu_\beta} \sqrt{\int \left(f'\right)^2 d\mu_\beta}.$$

Putting together all these estimates and taking into account that $\beta \ge 1$, $0 < s\beta \le \widetilde{\sigma}(p)$ and $s^{2(p-1)} \ge s^{1/(2-p)}$, it appears that

$$\begin{aligned} &\left|\int_{\mathbb{T}\times[0,1]} \widetilde{I}_1(y, u) + I_2(y, u) + I_3(y, u)\nu(dy)\, du\right| \\ &\quad \le k'(p, A) \begin{cases} \beta s, & \text{if } p = 1 \text{ or } p \ge 2, \\ \beta s + s^{2(p-1)}, & \text{if } p \in (1, 2) \end{cases} \\ &\quad \le 2k'(p, A) \begin{cases} \beta s, & \text{if } p = 1 \text{ or } p \ge 3/2, \\ \beta s + s^{2(p-1)}, & \text{if } p \in (1, 3/2), \end{cases} \end{aligned}$$

for another constant $k'(p, A) > 0$, depending on $p \in [1, 2)$ and $A > 0$. This finishes the proof of (4.3). □

To conclude the proof of Proposition 4.1, we must be able to compare, for any $\beta \ge 0$ and any $f \in \mathcal{C}^1(\mathbb{T})$, the energy $\mu_\beta[(\partial f)^2]$ and the variance $\operatorname{Var}(f, \mu_\beta)$. This task was already done by [14], let us recall their result.

Proposition 4.2. *Let U_p be a $\mathcal{C}^1$ function on a compact Riemannian manifold M of dimension $m \ge 1$. Let $b(U_p) \ge 0$ be the associated constant as in* (1.6). *For any $\beta \ge 0$, consider the Gibbs measure μ_β given in* (2.2). *Then there exists a constant $C_M > 0$, depending only on M, such that the following Poincaré inequalities are satisfied*:

$$\forall \beta \ge 0, \forall f \in \mathcal{C}^1(M), \qquad \operatorname{Var}(f, \mu_\beta) \le C_M\left[1 \vee \left(\beta \left\|U_p'\right\|_\infty\right)\right]^{5m-2} \exp\left(b(U_p)\beta\right)\mu_\beta\left[|\nabla f|^2\right].$$

We can now come back to the study of the evolution of the quantity $I_t = \operatorname{Var}(f_t, \mu_{\beta_t})$, for $t > 0$. Indeed applying Lemma 4.2 and Proposition 4.2 with $\alpha = \alpha_t$, $\beta = \beta_t$ and $f = f_t$, we get at any time $t > 0$ such that $\beta_t \ge 1$ and $\alpha_t \beta_t^2 \le \varsigma(p)$,

$$\begin{aligned} &\int L_{\alpha_t,\beta_t}[f_t - 1](f_t - 1)\, d\mu_{\beta_t} \\ &\quad \le -c_4\beta_t^{-3} \exp\left(-b(U_p)\beta_t\right)\left(1 - 2c_3(p, A)\alpha_t^{\widetilde{a}(p)}\beta_t^3\right)I_t + c_3(p, A)\alpha_t^{\widetilde{a}(p)}\beta_t^3 I_t \\ &\quad \le -\left(c_4\beta_t^{-3} \exp\left(-b(U_p)\beta_t\right) - c_5(p, A)\alpha_t^{\widetilde{a}(p)}\beta_t^3\right)I_t, \end{aligned}$$

where $c_4 := (16\pi^3 C_{\mathbb{T}})^{-1}$ and $c_5(p, A) := c_3(p, A)(1 + 2c_4)$.

Taking into account Lemma 4.1, the computations preceding Lemma 4.2 and Proposition 3.5, one can find constants $c_1(p, A), c_2(p, A) > 0$ and $\varsigma(p) \in (0, 1/2)$ such that Proposition 4.1 is satisfied.

This result leads immediately to conditions insuring the convergence toward 0 of the quantity I_t for large times $t > 0$.

Proposition 4.3. *Let* $\alpha : \mathbb{R}_+ \to \mathbb{R}_+^*$ *and* $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ *be schemes as at the beginning of this section and assume*

$$\lim_{t\to+\infty} \beta_t = +\infty,$$

$$\int_0^{+\infty} (1 \vee \beta_t)^{-3} \exp\bigl(-b(U_p)\beta_t\bigr)\, dt = +\infty$$

and that for large times $t > 0$,

$$\max\bigl\{\alpha_t^{a(p)}\beta_t^4, \alpha_t^{\widetilde{a}(p)}\beta_t^3, \bigl|\beta_t'\bigr|\bigr\} \ll \exp\bigl(-b(U_p)\beta_t\bigr)$$

(*where* $a(p) > 0$ *and* $\widetilde{a}(p) > 0$ *are defined in Propositions* 3.5 *and* 4.1). *Then we are assured of*

$$\lim_{t\to+\infty} I_t = 0.$$

Proof. The differential equation of Proposition 4.1 can be rewritten under the form

$$F_t' \le -\eta_t F_t + \epsilon_t, \tag{4.11}$$

where for any $t > 0$,

$$\begin{aligned}
F_t &:= \sqrt{I_t},\\
\eta_t &:= c_1(p, A)\bigl(\beta_t^{-3}\exp\bigl(-b(U_p)\beta_t\bigr) - \alpha_t^{\widetilde{a}(p)}\beta_t^3 - \bigl|\beta_t'\bigr|\bigr)/2,\\
\epsilon_t &:= c_2(p, A)\bigl(\alpha_t^{a(p)}\beta_t^4 + \bigl|\beta_t'\bigr|\bigr)/2.
\end{aligned}$$

The assumptions of the above proposition imply that for $t \ge 0$ large enough, $\beta_t \ge 1$ and $\alpha_t\beta_t^2 \le \varsigma(p)$, where $\varsigma(p) \in (0, 1/2)$ is as in Proposition 4.1. This ensures that there exists $T > 0$ such that (4.11) is satisfied for any $t \ge T$ (and also $F_T < +\infty$). We deduce that for any $t \ge T$,

$$F_t \le F_T \exp\biggl(-\int_T^t \eta_s\, ds\biggr) + \int_T^t \epsilon_s \exp\biggl(-\int_s^t \eta_u\, du\biggr)\, ds. \tag{4.12}$$

It appears that $\lim_{t\to+\infty} F_t = 0$ as soon as

$$\int_T^{+\infty} \eta_s\, ds = +\infty,$$

$$\lim_{t\to+\infty} \epsilon_t/\eta_t = 0.$$

The above assumptions were chosen to ensure these properties. □

In particular, remarking that $a(p) \leq \widetilde{a}(p)$ for any $p \geq 1$, the schemes given in (1.3) satisfy the hypotheses of the previous proposition, so that under the conditions of Theorem 1.1, we get

$$\lim_{t\to+\infty} I_t = 0.$$

Let us deduce (1.4) for any neighborhood $\mathcal{N}$ of the set $\mathcal{M}_p$ of the global minima of U_p. From Cauchy–Schwarz inequality we have for any $t > 0$,

$$\begin{aligned}\|m_t - \mu_{\beta_t}\|_{\mathrm{tv}} &= \int |f_t - 1|\mu_{\beta_t} \\ &\leq \sqrt{I_t}.\end{aligned}$$

An equivalent definition of the total variation norm states that

$$\|m_t - \mu_{\beta_t}\|_{\mathrm{tv}} = 2\max_{A\in\mathcal{T}}\bigl|m_t(A) - \mu_{\beta_t}(A)\bigr|,$$

where $\mathcal{T}$ is the Borelian σ-algebra of $\mathbb{T}$. It follows that (1.4) reduces to

$$\lim_{\beta\to+\infty} \mu_\beta(\mathcal{N}) = 1,$$

for any neighborhood $\mathcal{N}$ of $\mathcal{M}_p$, property which is immediate from the definition (2.2) of the Gibbs measures μ_β for $\beta \geq 0$. This finishes the proof of Theorem 1.1.

Remark 4.1. Under mild conditions, the results of [17] enable to go further, because he identifies the weak limit μ_∞ of the Gibbs measures μ_β as β goes to $+\infty$. Thus, if one knows, as above, that

$$\lim_{t\to+\infty} \|m_t - \mu_{\beta_t}\|_{\mathrm{tv}} = 0,$$

then one gets that m_t also weakly converges toward μ_∞ for large times $t > 0$. The weight given by μ_∞ to a point $x \in \mathcal{M}_p$ is inversely related to the value of $\sqrt{U_p''(x)}$ and in this respect Lemma 3.1 is useful (still assuming that ν admits a continuous density).

First note that for any $x \in \mathcal{M}_p$, we have $U_p''(x) \geq 0$, since x is a global minima of U_p, and by consequence $\nu(x') \leq 1$. Next, assume that we have for any $x \in \mathcal{M}_p$, $\nu(x') < 1$. It follows that $\mathcal{M}_p$ is discrete and by consequence finite, since $\mathbb{T}$ is compact. This property was already noted by [15], among other features of intrinsic means on the circle. Then we deduce from [17] that

$$\mu_\infty = \frac{1}{Z}\sum_{x\in\mathcal{M}_p} \frac{1}{\sqrt{1-\nu(x')}}\delta_x,$$

where $Z := \sum_{x\in\mathcal{M}_p}(1-\nu(x'))^{-1/2}$ is the normalizing factor.

In this situation, $\mathcal{L}(X_t)$ concentrates for large times $t > 0$ on all the p-means of ν. Thus, to find all of them with an important probability, one should sample independently several trajectories of X, for example, starting from a fixed point $X_0 \in \mathbb{T}$.

Remark 4.2. Similarly to the approach presented, for instance, in [22,25], we could have studied the evolution of $(E_t)_{t>0}$, which are the relative entropies of the time marginal laws with respect to the corresponding instantaneous Gibbs measures, namely

$$\forall t > 0, \qquad E_t := \int \ln\left(\frac{m_t}{\mu_{\beta_t}}\right) dm_t.$$

To get a differential inequality satisfied by these functionals, the spectral gap estimate of [14] recalled in Proposition 4.2 must be replaced by the corresponding logarithmic Sobolev constant estimate, which is proven in the same article of [14].

5. Extension to all probability measures ν

Our main task here is to adapt the computations of the two previous sections in order to prove Theorem 1.2. As in the statement of this result, it is better for simplicity of the exposition to restrict ourselves to the important and illustrative case $p = 2$; the general situation will be alluded to in the last remark of this section.

We begin by remarking that the algorithm Z described in the Introduction evolves similarly to the process X, if we allow the probability measure ν to depend on time. More precisely, for any $\kappa > 0$, consider the probability measure ν_κ given by

$$\forall z \in M, \qquad \nu_\kappa(dz) := \int \nu(dy) K_{y,\kappa}(dz), \tag{5.1}$$

where the kernel on M, $(y, dz) \mapsto K_{y,\kappa}(dz)$ was defined before the statement of Theorem 1.2. For $\alpha > 0$, $\beta \geq 0$ and $\kappa > 0$, let us denote by $L_{\alpha,\beta,\kappa}$ the generator defined in (2.4), where ν is replaced by ν_κ. Then the law of Z is solution of the time-inhomogeneous martingale problem associated to the family of generators $(L_{\alpha_t,\beta_t,\kappa_t})_{t\geq 0}$. This observation leads us to introduce the potentials

$$\forall \kappa > 0, \forall x \in M, \qquad U_{2,\kappa}(x) := \int d^2(x, y)\nu_\kappa(dy),$$

as well as the associated Gibbs measures:

$$\forall \beta \geq 0, \forall \kappa > 0, \qquad \mu_{\beta,\kappa}(dx) := Z_{\beta,\kappa}^{-1} \exp\bigl(-\beta U_{2,\kappa}(x)\bigr)\lambda(dx),$$

where $Z_{\beta,\kappa}$ is the renormalization constant.

Denote by m_t the law of Z_t for any $t \geq 0$. The proof of Theorem 1.2 is then similar to that of Theorem 1.1 and relies on the investigation of the evolution of

$$\forall t > 0, \qquad \mathcal{I}_t := \int \left(\frac{m_t}{\mu_{\beta_t,\kappa_t}} - 1\right)^2 d\mu_{\beta_t,\kappa_t}, \tag{5.2}$$

which play the role of the quantities defined in (4.1).

While the above program was presented for a general compact Riemannian manifold M, we again restrict ourselves to the situation $M = \mathbb{T}$.

We first need some estimates on the probability measures ν_κ, for $\kappa > 0$.

Lemma 5.1. *For any $\kappa > 0$, ν_κ admits a density with respect to λ, still denoted ν_κ. Furthermore we have, for any $\kappa > 1/\pi$,*

$$\|\nu_\kappa\|_\infty \le 2\pi\kappa,$$
$$\|\partial\nu_\kappa\|_\infty \le 2\pi\kappa^2,$$

where $\partial\nu_\kappa$ stands for the weak derivative (so that the last norm $\|\cdot\|_\infty$ is the essential supremum norm with respect to λ).

Proof. When $M = \mathbb{T}$, for any $\kappa > 0$, the kernel $K_{\cdot,\kappa}(\cdot)$ corresponds to the rolling around $\mathbb{T}$ of the kernel defined on $\mathbb{R}$ by $(y, dz) \mapsto \kappa(1 - \kappa|z - y|)_+ \, dz$. In particular for any $y \in \mathbb{T}$, $K_{y,\kappa}(\cdot)$ is absolutely continuous with respect to λ and (5.1) shows that the same is true for ν_κ. If furthermore $\kappa > 1/\pi$, from this definition we can write for any $z \in \mathbb{T}$,

$$\nu_\kappa(dz) = \kappa\left(\int_{z-1/\kappa}^{z+1/\kappa} \big(1 - \kappa d(y,z)\big)_+ \nu(dy)\right) dz,$$

namely, almost everywhere with respect to $\lambda(dz)$,

$$\begin{aligned}
\nu_\kappa(z) &= 2\pi\kappa \int_{z-1/\kappa}^{z+1/\kappa} \big(1 - \kappa d(y,z)\big)_+ \nu(dy) \\
&\le 2\pi\kappa \int_{z-1/\kappa}^{z+1/\kappa} \nu(dy) \\
&\le 2\pi\kappa.
\end{aligned}$$

Next, for almost every $x, y \in \mathbb{T}$, we have

$$\begin{aligned}
\big|\nu_\kappa(x) - \nu_\kappa(y)\big| &\le 2\pi\kappa \int_{\mathbb{T}} \big|\big(1 - \kappa d(x,z)\big)_+ - \big(1 - \kappa d(y,z)\big)_+\big| \nu(dz) \\
&\le 2\pi\kappa \int_{\mathbb{T}} \big|1 - \kappa d(x,z) - 1 + \kappa d(y,z)\big| \nu(dz) \\
&\le 2\pi\kappa^2 \int_{\mathbb{T}} \big|d(x,z) - d(y,z)\big| \nu(dz) \\
&\le 2\pi\kappa^2 d(x,y).
\end{aligned}$$

This proves the second bound. □

An immediate consequence of the last bound is that for any $x \in \mathbb{T}$, the map $(1/\pi, +\infty) \ni \kappa \mapsto U_{2,\kappa}(x)$ is weakly differentiable and for almost every $\kappa > 1/\pi$, $|\partial_\kappa U_{2,\kappa}(x)| \leq 2\pi^4\kappa^2$; but one can do better.

Lemma 5.2. *For any $x \in \mathbb{T}$ and any $\kappa > 1/\pi$, we have*

$$\left|\partial_\kappa U_{2,\kappa}(x)\right| \leq \frac{3\pi^3}{\kappa}.$$

Proof. It is better to come back to the definition of ν_κ, to get, for $x \in \mathbb{T}$ and $\kappa > 1/\pi$ (where ∂_κ stands for weak derivative):

$$\begin{aligned}\partial_\kappa U_{2,\kappa}(x) &= \partial_\kappa\left(2\pi\kappa \int \lambda(dy)d^2(x,y)\int_{\mathbb{T}}\bigl(1-\kappa d(y,z)\bigr)_+\nu(dz)\right)\\ &= 2\pi \int \lambda(dy)d^2(x,y)\int_{\mathbb{T}}\nu(dz)\bigl(1-\kappa d(y,z)\bigr)_+\\ &\quad - 2\pi\kappa\int \lambda(dy)d^2(x,y)\int_{y-1/\kappa}^{y-1/\kappa}\nu(dz)d(y,z).\end{aligned}$$

The first term of the right-hand side is equal to $U_{2,\kappa}(x)/\kappa$ and is bounded by $\|U_{2,\kappa}\|_\infty/\kappa \leq \pi^2/\kappa$. In absolute value, the second term can be written under the form

$$\begin{aligned}2\pi\kappa\int\nu(dz)\int_{z-1/\kappa}^{z-1/\kappa}\lambda(dy)d^2(x,y)d(y,z) &\leq 2\pi^3\kappa\int\nu(dz)\int_{z-1/\kappa}^{z-1/\kappa}\lambda(dy)|y-z|\\ &= \frac{2\pi^3}{\kappa}.\end{aligned}$$

□

The improvement of the estimate of the previous lemma with respect to the one given before its statement is important for us, since it enables to obtain that if $(\beta_t)_{t\geq 0}$ and $(\kappa_t)_{t\geq 0}$ are $\mathcal{C}^1$ schemes, then we have

$$\forall t \geq 0, \qquad \bigl\|\partial_t \ln(\mu_{\beta_t,\kappa_t})\bigr\|_\infty \leq \pi^2\bigl|\beta_t'\bigr| + 3\pi^3\beta_t\bigl|\bigl(\ln(\kappa_t)\bigr)'\bigr|. \tag{5.3}$$

This bound replaces that of Lemma 4.1 in the present context. Note that for the schemes we have in mind and up to mild logarithmic corrections, we recover a bound of order $1/(1+t)$ for $\|\partial_t \ln(\mu_{\beta_t,\kappa_t})\|_\infty$, which is compatible with our purposes.

In the same spirit, even if this cannot be deduced directly from Lemma 5.2, we have the following.

Lemma 5.3. *As κ goes to infinity, $U_{2,\kappa}$ converges uniformly toward U_2. In particular, if $b(\cdot)$ is the functional defined in (1.6), then we have*

$$\lim_{\kappa\to+\infty} b(U_{2,\kappa}) = b(U_2).$$

Proof. Since $\|\partial U_{2,\kappa}\|_\infty \leq 2\pi$, for any $\kappa > 0$, it appears that $(U_{2,\kappa})_{\kappa>0}$ is an equicontinuous family of mappings. It is besides clear that ν_κ weakly converges toward ν as κ goes to infinity, so that $U_{2,\kappa}(x)$ converges toward $U_2(x)$ for any fixed $x \in \mathbb{T}$. Compactness of $\mathbb{T}$ and the Arzelà–Ascoli theorem then enable to conclude to the uniform of $U_{2,\kappa}$ toward U_2 as κ goes to infinity. The second assertion of the lemma is an immediate consequence of this convergence. □

Consider for the evolution of the inverse temperature the scheme

$$\forall t \geq 0, \qquad \beta_t := b^{-1}\ln(1+t),$$

where $b > b(U_2)$ and denote $\rho := (1 + b(U_2)/b)/2 < 1$. Assume that the scheme $(\kappa_t)_{t\geq 0}$ is such that $\lim_{t\to+\infty}\kappa_t = +\infty$. Then from the above lemma and Proposition 4.2 (recall that $\|\partial U_{2,\kappa}\|_\infty \leq 2\pi$, for any $\kappa > 0$), there exists a time $T > 0$ such that for any $t \geq T$,

$$\forall f \in \mathcal{C}^1(\mathbb{T}), \qquad \frac{2}{(1+t)^\rho}\operatorname{Var}(f, \mu_{\beta_t,\kappa_t}) \leq \mu_{\beta_t,\kappa_t}\big[(\partial f)^2\big]. \tag{5.4}$$

Like (5.3), this crucial estimate for the investigation of the evolution of the quantities (5.2) still does not explain the requirement that $k \in (0, 1/2)$ in Theorem 1.2. Its justification comes from the next result, which replaces Proposition 3.1 in the present situation.

Proposition 5.1. *For $\alpha > 0$, $\beta \geq 0$ and $\kappa > 0$, let $L^*_{\alpha,\beta,\kappa}$ be the adjoint operator of $L_{\alpha,\beta,\kappa}$ in $\mathbb{L}^2(\mu_{\beta,\kappa})$. There exists a constant $C_1 > 0$ such that for any $\beta \geq 1$, $\kappa \geq 1$ and $\alpha \in (0, (2\beta)^{-1} \wedge (\beta^3(\beta+\kappa))^{-1/2})$, we have*

$$\big\|L^*_{\alpha,\beta,\kappa}\mathbb{1}\big\|_\infty \leq C_1\alpha\beta^2\big(\beta^2+\kappa^2\big).$$

Proof. It is sufficient to replace U_2 by $U_{2,\kappa}$ in the proofs of Section 3, in particular note that (3.4) still holds. From Lemma 3.1 and the first part of Lemma 5.1, it appears that (3.6) has to be replaced by

$$\forall \kappa \geq 1, \qquad \big\|U''_{2,\kappa}\big\|_\infty \leq 4\pi\kappa.$$

Instead of (3.7), we deduce that for any $x, y \in \mathbb{T}$ and α, β and κ as in the statement of the proposition,

$$\begin{aligned}&\exp\left(\beta\left[U_{2,\kappa}(x) - U_{2,\kappa}\left(x - \frac{\alpha\beta}{1-\alpha\beta}(y-x)\right)\right]\right)\\ &\quad= 1 + \frac{\alpha\beta^2}{1-\alpha\beta}U'_{2,\kappa}(x)(y-x) + \mathcal{O}\big(\alpha^2\beta^3(\beta+\kappa)\big).\end{aligned}$$

Keeping following the computations of the same proof, we end up with

$$L^*_{\alpha,\beta,\kappa}\mathbb{1}(x) = \frac{\beta}{1-\alpha\beta}\frac{1}{2\pi\alpha\beta}\int_{x'-\alpha\beta\pi}^{x'+\alpha\beta\pi}\nu_\kappa\big(x'\big) - \nu_\kappa(y)\,dy + \mathcal{O}\big(\alpha\beta^3(\beta+\kappa)\big).$$

To estimate the last integral, we resort to the second part of Lemma 5.1: we get

$$\left|\int_{x'-\alpha\beta\pi}^{x'+\alpha\beta\pi} \nu_\kappa(x') - \nu_\kappa(y)\,dy\right| \leq 2\pi\kappa^2 \int_{x'-\alpha\beta\pi}^{x'+\alpha\beta\pi} |x'-y|\,dy = 2\pi\kappa^2(\alpha\beta\pi)^2.$$

This leads to the announced bound. □

Similar arguments transform Lemma 4.2 into the following.

Lemma 5.4. *There exists a constant* $C_2 > 0$, *such that for any* $\alpha > 0$, $\beta \geq 1$ *and* $\kappa \geq 1$ *with* $\alpha\beta^2 \leq 1/2$, *we have, for any* $f \in \mathcal{C}^2(\mathbb{T})$,

$$\int L_{\alpha,\beta,\kappa}[f-1](f-1)\,d\mu_\beta \leq -\left(\frac{1}{2} - C_2\alpha\beta^2(\beta+\kappa)\right)\int (\partial f)^2\,d\mu_\beta + C_2\alpha\beta^2(\beta+\kappa)\int (f-1)^2\,d\mu_\beta.$$

Proof. The modifications with respect to the proof of Lemma 4.2 are very limited: one just needs to take into account the bounds $\|U'_{p,\kappa}\|_\infty \leq 2\pi$ and $\|\nu_\kappa\|_\infty \leq 2\pi\kappa$ for $\kappa \geq 1$. Indeed, there are two main changes:

- in (4.2), where the remaining operator has to be defined by

$$R_{\alpha,\beta,\kappa} := L_{\alpha,\beta,\kappa} - \tfrac{1}{2}\left(\partial^2 - \beta U'_{p,\kappa}\partial\right),$$

- in (4.7), the factor $1 + A\pi$ must be replaced by $2\pi\kappa$, by virtue of the first estimate of Lemma 5.1. It leads to the supplementary term $\alpha\beta^2\kappa$ in the bound of the above lemma. □

All the ingredients are collected together to get a differential inequality satisfied by $(\mathcal{I}_t)_{\geq 0}$. More precisely, under the requirement that (5.4) is true for $t \geq T > 0$, as well as $\beta_t \geq 1$, $\kappa_t \geq 1$ and $\alpha_t\beta_t^2\sqrt{\kappa_t} \leq 1/2$, we get that there exists a constant $C_3 > 0$ such that

$$\forall t \geq T, \qquad \mathcal{I}'_t \leq -\eta_t\mathcal{I}_t + \epsilon_t\sqrt{\mathcal{I}_t},$$

where for any $t \geq T$,

$$\eta_t := \frac{1}{(1+t)^\rho} - C_3\left(\alpha_t\beta_t^2(\beta_t+\kappa_t) + |\beta'_t| + \beta_t\left|(\ln(\kappa_t))'\right|\right),$$
$$\epsilon_t := C_3\left(\alpha_t\beta_t^2\left(\beta_t^2+\kappa_t^2\right) + |\beta'_t| + \beta_t\left|(\ln(\kappa_t))'\right|\right).$$

Under the assumptions of Theorem 1.2 (already partially used to ensure the validity of (5.4) for some $\rho \in (0,1)$), it appears that as t goes to infinity,

$$\eta_t \sim \frac{1}{(1+t)^\rho},$$
$$\epsilon_t = \mathcal{O}\left(\frac{1}{1+t}\right)$$

and this is sufficient to ensure that

$$\lim_{t\to+\infty} \mathcal{I}_t = 0.$$

The proof of Theorem 1.2 finishes by the arguments given at the end of Section 4.

Remark 5.1. As it was mentioned at the end of the Introduction, if one does not want to waste rapidly the sample $(Y_n)_{n\in\mathbb{N}}$ (especially if it is not infinite...), one should take the exponent c the smallest possible. From our assumptions, we necessarily have $c > 1$. But the limit case $c = 1$ can be attained: the above proof shows that the convergence of Theorem 1.2 is also valid for the schemes:

$$\forall t \geq 0, \qquad \begin{cases} \alpha_t := (1+t)^{-1}, \\ \beta_t := b^{-1}\ln(1+t), \\ \kappa_t := \ln(2+t). \end{cases}$$

The drawback is that ν is not rapidly approached by ν_{κ_t} as t goes to infinity and this may slow down the convergence of the algorithm toward $\mathcal{N}$. Indeed, from the previous computations, it appears that the law of Z_t is rather close to the set of global minima of U_{2,κ_t}.

Remark 5.2. The cases $p = 1$ and $p \geq 2$ can be treated in the same manner, but for $p \in (1, 2)$, one must follow the dependence on A of the constants in the proof of Lemma 3.10. In the end it only leads to supplementary factors of κ, so that Theorem 1.2 is satisfied with a sufficiently large constant c, depending on $p \geq 1$ and on the exponent k entering in the definition of the scheme $(\kappa_t)_{t\geq 0}$. But before going further in the direction of this generalization, it would be more rewarding to first check if the dependence on p of a_p in Theorem 1.1 is just technical or really necessary.

Appendix: Regularity of temporal marginal laws

Our goal is to see that at positive times, the marginal laws of the considered algorithms are absolutely continuous and that if furthermore $\nu \ll \lambda$, then the corresponding densities belong to $\mathcal{C}^1(\mathbb{T})$. We will also check that this is sufficient to justify the computations made in Section 4.

Let X be the process described in the Introduction, for simplicity on $\mathbb{T}$, but the following arguments could be extended to general connected and compact Riemannian manifolds. We are going to use the probabilistic construction of X to obtain regularity results on m_t, which as usual stands for the law of X_t, for any $t \geq 0$. So for fixed $t > 0$, let T_t be the largest jump time of $N^{(\alpha)}$ in the interval $[0, t]$, with the convention that $T_t = 0$ if there is no jump time in this interval. Denote by ξ_t the law of (T_t, X_{T_t}) on $[0, t] \times \mathbb{T}$. Furthermore, let $P_s(x, dy)$ be the law at time $s \geq 0$ of the Brownian motion on $\mathbb{T}$, starting at $x \in \mathbb{T}$. From the construction given in the Introduction, we have for any $t > 0$,

$$m_t(dx) = \int_{[0,t]\times\mathbb{T}} \xi_t(ds, dz) P_{t-s}(z, dx). \tag{A.1}$$

An immediate consequence is the following.

Lemma A.1. *Let $t > 0$ be fixed. About the measurable evolutions $\alpha : \mathbb{R}_+ \to \mathbb{R}_+^*$ and $\beta : \mathbb{R}_+ \to \mathbb{R}_+$, only assume that $\inf_{s\in[0,t]} \alpha_s > 0$. Then, whatever the probability measure ν entering in the definition of X, we have that m_t is absolutely continuous.*

Proof. By the hypothesis on α, 0 is the unique atom of $\xi(\cdot, \mathbb{T})$, the distribution of T_t (its mass is $\xi_t(\{0\}, \mathbb{T}) = \exp(-\int_0^t 1/\alpha_s\, ds)$) and $\xi(\cdot, \mathbb{T})$ admits a bounded density on $(0, t]$. Since furthermore for any $s > 0$ and $z \in \mathbb{T}$, $P_s(z, \cdot)$ is absolutely continuous, the same is true for m_t due to (A.1). □

To go further, we need to strengthen the assumption on ν.

Lemma A.2. *In addition to the hypotheses of the previous lemma, assume that ν admits a bounded density and that $\inf_{s\in[0,t]} \beta_s > 0$. Then for any $t > 0$, the density of m_t belongs to $\mathcal{C}^1(\mathbb{T})$.*

Proof. We begin by recalling a few bounds on the heat kernels $P_s(x, dy)$, for $s > 0$ and $x \in \mathbb{T}$. We have already mentioned they admit a density, namely they can be written under the form $p_s(x, y)\, dy$. Since the Brownian motion on $\mathbb{T}$ is just the rolling up of the usual Brownian motion on $\mathbb{R}$, we have for any $x \in \mathbb{T}$,

$$\forall y \in (x - \pi, x + \pi], \qquad p_s(x, y) = \sum_{n\in\mathbb{Z}} \frac{\exp(-(y - x + 2\pi n)^2/(2s))}{\sqrt{2\pi s}}. \tag{A.2}$$

From a general bound due to [16], we deduce that there exists a constant $C_0 > 0$ such that for any $s > 0$ and $y \in (x - \pi, x + \pi]$, we have

$$\left|\partial_y p_s(x, y)\right| \le C_0\left(\frac{d(x, y)}{s} + \frac{1}{\sqrt{s}}\right) p_s(x, y).$$

To get an upper bound on $p_s(x, y) = p_s(0, y - x)$, consider separately in (A.2) the sums of $n \in \mathbb{Z}_\sigma$ and $n \in \mathbb{Z}_{-\sigma} \setminus \{0\}$, where $\sigma \in \{-, +\}$ is the sign of $y - x$. It appears that for $s \in (0, t]$,

$$\begin{aligned} p_s(x, y) &\le 2 \sum_{n\in\mathbb{Z}_\sigma} \frac{\exp(-(y - x + 2\pi n)^2/(2s))}{\sqrt{2\pi s}} \\ &\le 2 \frac{\exp(-(y - x)^2/(2s))}{\sqrt{2\pi s}} \sum_{n\in\mathbb{Z}_+} \exp\bigl(-(2\pi n)^2/(2s)\bigr) \\ &\le C_1(t) \frac{\exp(-d^2(x, y)/(2s))}{\sqrt{2\pi s}}, \end{aligned}$$

where $C_1(t) := \sum_{n\in\mathbb{Z}_+} \exp(-2(\pi n)^2/t)$. Taking into account (A.1) and Lemma A.1, if we were allowed to differentiate under the sign integral, we would get for any $x \in \mathbb{T}$,

$$\partial_x m_t(x) = \int_{[0,t]\times\mathbb{T}} \xi_t(ds, dz)\partial_x p_{t-s}(z, x) \tag{A.3}$$

(where the left-hand side stands for the density of m_t with respect to $2\pi\lambda$). Unfortunately, the usual conditions do not apply here, so it is better to consider the approximation of the density m_t by $m_{\epsilon,t}$, where for $\epsilon \in (0, t)$,

$$\forall x \in \mathbb{T}, \qquad m_{t,\epsilon}(x) := \int_{[0,t-\epsilon]\times\mathbb{T}} \xi_t(ds, dz) p_{t-s}(z, x).$$

There is no difficulty in differentiating this expression under the sign sum and in the end it appears to be smooth in x. So to get the announced result, it is sufficient to see that $\partial_x m_{\epsilon,t}(x)$ converges to the right-hand side of (A.3), uniformly in $x \in \mathbb{T}$ as ϵ goes to 0_+. Let us prove the stronger convergence

$$\lim_{\epsilon\to 0_+} \sup_{x\in\mathbb{T}} \int_{[t-\epsilon,t]\times\mathbb{T}} \xi_t(ds, dz)\left|\partial_x p_{t-s}(z, x)\right| = 0.$$

The assumptions that $\inf_{s\in[0,t]} \alpha_s\beta_s > 0$ and that ν admits a bounded density imply that the latter is equally true for $\xi_t(s, \cdot)$, the regular conditional law of X_{T_t} knowing that $T_t = s$, for any $s > 0$. We can even find $C_2(t) > 0$ such that $\xi_t(s, dz) \le C_2(t)\, dz$, uniformly over $s \in (0, t]$ (but a priori $C_2(t)$ may depend on $t > 0$ through $\inf_{s\in[0,t]} \alpha_s\beta_s$). In the proof of Lemma A.1, we have already noticed that there exists $C_3(t) > 0$ such that $\xi_t(ds, \mathbb{T}) \le C_3(t)\, ds$, for $s \ne 0$. It follows that for $\epsilon \in (0, t)$,

$$\begin{aligned}
&\int_{[t-\epsilon,t]\times\mathbb{T}} \xi_t(ds, dz)\left|\partial_x p_{t-s}(z, x)\right| \\
&\quad\le C_0C_1(t)C_2(t)C_3(t) \int_{[t-\epsilon,t]} ds \int_{\mathbb{T}} dz\left(\frac{d(z, x)}{(t-s)^{3/2}} + \frac{1}{t-s}\right)\frac{\exp(-d^2(z, x)/(2(t-s)))}{\sqrt{2\pi}} \\
&\quad= 2C_0C_1(t)C_2(t)C_3(t) \int_0^{\pi} dz \int_0^{\epsilon} ds\left(\frac{z}{s^{3/2}} + \frac{1}{s}\right)\frac{\exp(-z^2/(2s))}{\sqrt{2\pi}}.
\end{aligned}$$

This bound no longer depends on x and to compute the latter integral, consider the change of variable $u = z^2/s$, z being fixed:

$$\int_0^{\pi} dz \int_0^{\epsilon} ds\left(\frac{z}{s^{3/2}} + \frac{1}{s}\right)\exp\bigl(-z^2/(2s)\bigr) = \int_0^{\pi} dz \int_{z^2/\epsilon}^{+\infty} du\left(\frac{1}{\sqrt{u}} + u\right)\exp(-u/2).$$

We conclude by remarking that by the dominated convergence theorem, the latter term goes to zero with ϵ. □

Remark A.3. More generally, but still under the assumption that ν admits a bounded density, the density m_t is $\mathcal{C}^1$ at some time $t > 0$, if we can find $\epsilon \in (0, t)$ such that $\inf_{s\in[t-\epsilon,t]} \alpha_s > 0$ and $\inf_{s\in[t-\epsilon,t]} \beta_s > 0$. This comes from the above proof or can be deduced directly from Lemma A.2 and the Markov property of X.

The same arguments cannot be used to prove that for $t > 0$, the density of m_t belongs to $\mathcal{C}^2(\mathbb{T})$. A priori, this is annoying, since in Section 4, to study the evolution of the quantity I_t defined in (4.1), we had to differentiate it with respect to $t > 0$ and the computations were justified only if the densities m_t were $\mathcal{C}^2$. The classical way go around this apparent difficulty is to use a mollifier.

Let ρ be a smooth non-negative function on $\mathbb{R}$ whose support is included in $[-1, 1]$ and satisfying $\int_{\mathbb{R}} \rho(y)\, dy = 1$. For any $\delta \in (0, 1)$, define

$$\forall t \geq 0, \forall x \in \mathbb{T}, \qquad m_t^{(\delta)}(x) := \frac{1}{\delta} \int_{\mathbb{R}} m_t(x + y)\rho\left(\frac{y}{\delta}\right) dy$$

(where functions on $\mathbb{T}$ are naturally identified with 2π-periodic functions on $\mathbb{R}$). These functions are smooth and what is even more important for Section 4, the mapping $\mathbb{R}_+^* \times \mathbb{T} \ni (t, x) \mapsto \partial_x^2 m_t^{(\delta)}(x)$ is continuous. Furthermore, the $m_t^{(\delta)}$ are densities of probability measures on $\mathbb{T}$. More precisely, for any $t \geq 0$, $m_t^{(\delta)}$ is the density of $\mathcal{L}(X_t)$ when $\mathcal{L}(X_0) = m_0^{(\delta)}$, as a consequence of the linearity of the underlying evolution equation (i.e., $\forall t \geq 0$, $\partial_t m_t = m_t L_{\alpha_t,\beta_t}$, in the sense of distributions). Thus, the computations of Section 4 are justified if we replace there $(m_t)_{t>0}$ by $(m_t^{(\delta)})_{t>0}$, for any fixed $\delta \in (0, 1)$. In particular, the inequality (4.12) is satisfied for $(m_t^{(\delta)})_{t>0}$ instead of $(m_t)_{t>0}$. It remains to let δ go to 0_+ to see that the same bound is true for the flow $(m_t)_{t>0}$. This proves Theorem 1.1 for general initial distributions m_0, for instance, Dirac masses. In fact, one could pass to the limit $\delta \to 0_+$ before (4.12), for instance, already in Proposition 4.1, to see that it is also valid.

Acknowledgements

We would like to express our gratitude to our colleague Sébastien Gadat, who has encoded and drawn all the figures of the present work. We are also thankful to the referees whose suggestions have helped us to improve the presentation of the paper. Finally, we are grateful to the support of the Laboratoire de Mathématiques et Applications (UMR 7348) of the Université de Poitiers, where much of this work took place.

References

[1] Afsari, B., Tron, R. and Vidal, R. (2013). On the convergence of gradient descent for finding the Riemannian center of mass. *SIAM J. Control Optim.* **51** 2230–2260. MR3057324

[2] Arnaudon, M., Dombry, C., Phan, A. and Yang, L. (2012). Stochastic algorithms for computing means of probability measures. *Stochastic Process. Appl.* **122** 1437–1455. MR2914758

[3] Arnaudon, M. and Miclo, L. (2014). Means in complete manifolds: Uniqueness and approximation. *ESAIM Probab. Stat.* **18** 185–206. MR3230874

[4] Arnaudon, M. and Miclo, L. (2014). A stochastic algorithm finding generalized means on compact manifolds. *Stochastic Process. Appl.* **124** 3463–3479. MR3231628

[5] Arnaudon, M. and Nielsen, F. (2012). Medians and means in Finsler geometry. *LMS J. Comput. Math.* **15** 23–37. MR2891143

[6] Bădoiu, M. and Clarkson, K.L. (2003). Smaller core-sets for balls. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (*Baltimore*, *MD*, 2003) 801–802. New York: ACM. MR1974995

[7] Bonnabel, S. (2013). Stochastic gradient descent on Riemannian manifolds. *IEEE Trans. Automat. Control* **58** 2217–2229. MR3101606

[8] Cardot, H., Cénac, P. and Zitt, P.-A. (2013). Efficient and fast estimation of the geometric median in Hilbert spaces with an averaged stochastic gradient algorithm. *Bernoulli* **19** 18–43. MR3019484

[9] Catoni, O. (1999). Simulated annealing algorithms and Markov chains with rare transitions. In *Séminaire de Probabilités*, *XXXIII*. *Lecture Notes in Math.* **1709** 69–119. Berlin: Springer. MR1767994

[10] Charlier, B. (2013). Necessary and sufficient condition for the existence of a Fréchet mean on the circle. *ESAIM Probab. Stat.* **17** 635–649. MR3126155

[11] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes*: *Characterization and Convergence*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0838085

[12] Groisser, D. (2004). Newton's method, zeroes of vector fields, and the Riemannian center of mass. *Adv. in Appl. Math.* **33** 95–135. MR2064359

[13] Groisser, D. (2005). On the convergence of some Procrustean averaging algorithms. *Stochastics* **77** 31–60. MR2138772

[14] Holley, R.A., Kusuoka, S. and Stroock, D.W. (1989). Asymptotics of the spectral gap with applications to the theory of simulated annealing. *J. Funct. Anal.* **83** 333–347. MR0995752

[15] Hotz, T. and Huckemann, S. (2015). Intrinsic means on the circle: Uniqueness, locus and asymptotics. *Ann. Inst. Statist. Math.* **67** 177–193. MR3297863

[16] Hsu, E.P. (1999). Estimates of derivatives of the heat kernel on a compact Riemannian manifold. *Proc. Amer. Math. Soc.* **127** 3739–3744. MR1618694

[17] Hwang, C.-R. (1980). Laplace's method revisited: Weak convergence of probability measures. *Ann. Probab.* **8** 1177–1182. MR0602391

[18] Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. Amsterdam: North-Holland. MR1011252

[19] Jost, J. (2011). *Riemannian Geometry and Geometric Analysis*, 6th ed. *Universitext*. Heidelberg: Springer. MR2829653

[20] Le, H. (2004). Estimation of Riemannian barycentres. *LMS J. Comput. Math.* **7** 193–200. MR2085875

[21] McKilliam, R.G., Quinn, B.G. and Clarkson, I.V.L. (2012). Direction estimation by minimum squared arc length. *IEEE Trans. Signal Process.* **60** 2115–2124. MR2954196

[22] Miclo, L. (1992). Recuit simulé sans potentiel sur une variété riemannienne compacte. *Stoch. Stoch. Rep.* **41** 23–56. MR1275365

[23] Miclo, L. (1995). Remarques sur l'ergodicité des algorithmes de recuit simulé sur un graphe. *Stochastic Process. Appl.* **58** 329–360. MR1348382

[24] Miclo, L. (1995). Une étude des algorithmes de recuit simulé sous-admissibles. *Ann. Fac. Sci. Toulouse Math.* (6) **4** 819–877. MR1623480

[25] Miclo, L. (1996). Recuit simulé partiel. *Stochastic Process. Appl.* **65** 281–298. MR1425361

[26] Pennec, X. (2006). Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J. Math. Imaging Vision* **25** 127–154. MR2254442

[27] Sturm, K.-T. (2003). Probability measures on metric spaces of nonpositive curvature. In *Heat Kernels and Analysis on Manifolds*, *Graphs*, *and Metric Spaces* (*Paris*, 2002). *Contemp*. *Math*. **338** 357–390. Providence, RI: Amer. Math. Soc. MR2039961

[28] Yang, L. (2010). Riemannian median and its estimation. *LMS J*. *Comput*. *Math*. **13** 461–479. MR2748393

Received September 2013 and revised March 2015