Geometric Analysis of Metropolis Algorithm on Bounded Domain

L. Michel
(joint work with P. Diaconis and G. Lebeau)

Laboratoire J.-A. Dieudonné
Université de Nice

Let $\mu = \rho(x)dx$ be a probability measure on $[a, b]$ and let f be a regular function on $[a, b]$. We want to compute numerically the quantity $I = \frac{1}{b-a}\int_a^b f(x)d\mu(x)$.

- Standard "deterministic" method consist to divide $[a, b]$ into N interval and to approximate I by $\sum_{k=1}^N A_k$ where A_k is the area corresponding to the kth interval.
- Probabilist approach: let $(x_n)_{n\in\mathbb{N}}$ be a sequence of numbers in $[a, b]$ such that x_n is choosen at random with respect to μ. Then, the quantity $\frac{1}{N}\sum_{n=1}^N f(x_n)$ provides a good approximation of I.
- A priori, "choose a point at random with respect to μ" is not a simpler problem than "compute I".
  The Metropolis Algorithm provides an efficient procedure to sample from μ.

The problem of hard spheres

Consider a fixed box in $\mathbb{R}^d$, $B =]-1,1[^d$. We consider the problem of placement of N balls of radius $\epsilon > 0$ with centers in B under the condition that the balls do not overlap. We denote $\mathcal{O}_{N,\epsilon} \subset B^N$ the set of all possible configurations. We endowe $\mathcal{O}_{N,\epsilon}$ with the Lebesgue measure dL.

Problem:

Build a sample of points $X^1, \dots, X^r \in \mathcal{O}_{N,\epsilon}$ distributed uniformly with respect to dL.

- This problem occurs in statistical physics in phase transition studies.
- It can be formulated in a more abstract setting.

Metropolis and al (50's) proposed the following algorithm to solve this problem. Let $h > 0$ being fixed and $X^0 \in \mathcal{O}_{N,\epsilon}$.

- Starting from $X^0 = (x_1^0, \dots, x_N^0)$, move one of the ball say x_k^0 uniformly at random in the ball $B(x_k^0, h)$, it results in a new position x_k^1. Denote $X^1 = (x_1^0, \dots, x_k^1, \dots, x_N^0)$ the new configuration. If $X^1 \in \mathcal{O}_{N,\epsilon}$, keep X^1.
- If $X^1 \notin \mathcal{O}_{N,\epsilon}$, throw away X^1 and restart the procedure from X^0.
- Once, X^1 is constructed, define X^2 by the same procedure starting from X^1, etc.

As r goes to infinity, the distribution of $X^0, \dots, X^r$ in $\mathcal{O}_{N,\epsilon}$ is close to the uniform distribution.

Abstract probabilistic setting

Let (X, d) be a metric space and $\mathcal{B}$ the Borel σ-algebra on X. Let $K(x, dy)$ be a Markov kernel on X, i.e.

- for all $x \in X$, $K(x, dy)$ is a probability measure on $(X, \mathcal{B})$.
- for all $B \in \mathcal{B}$, $x \mapsto K(x, B)$ is continuous (to simplify).

For $n \in \mathbb{N}^*$ we define the iterated kernel $K^n(x, dy)$ by

$$K^{m+n}(x, B) = \int K^m(y, B) K^n(x, dy), \ \forall B \in \mathcal{B}$$

The kernel K induces an operator on continuous functions by

$$Kf(x) = \int_X f(y) K(x, dy)$$

and its transpose acts on Borel measure on X.

Definition

A stationary distribution is a probability measure $\pi(dx)$ on X such that ${}^tK(\pi) = \pi$. In other words:

$$\forall B \in \mathcal{B},\ \pi(B) = \int K(x, B)\pi(dx)$$

example

Suppose that X is a finite space and let $n = \sharp X$. Then a Markov kernel is a matrix $(K(x,y))_{1\leq x,y\leq n}$ with non-negative coefficients and such that for any $x \in X, \Sigma_{y\in X} K(x,y) = 1$. Hence , a stationary distribution is an eigenvector of ${}^tK$ associated to the eigenvalue 1 and with non-negative coordinates.

Theorem

Suppose that $K(x, dy)$ is a *strictly positive, regular* Markov kernel and that $\pi(dx)$ is stationary for K. Then,

$$\forall x \in X, \forall B \in \mathcal{B}, \lim_{n\to\infty} K^n(x, B) = \pi(B)$$

A Markov kernel is strictly positive if $K(x, A) > 0$ for any open subset A. We do not define the notion of regular Markov kernel. Think it as a density $k(x, y)dy$ on an open subset of $\mathbb{R}^d$, with k continuous w.r.t. (x, y) (enough to apply Ascoli's theorem).

Question

What can we say about the speed of convergence?

Given a probability distribution π on X we may be interested in sampling π. From the preceding theorem, it is clear that if $K(x, dy)$ is a Markov kernel for which π is stationary, we can build a sample by the following process:

- Start from $x^0 \in X$ and buid $x^1 \in X$ at random with the probability $K(x^0, dy)$.
- Knowing $x^0, \dots, x^n \in X$ build x^{n+1} at random with the probability $K(x^n, dy)$.

Since $K^n(x, dy)$ converges to π, the distribution of the point $x^0, \dots, x^n$ "looks like" it was choosen according to π.

Question

Given a probability π, how can we construct a Markov kernel $K(x, dy)$ such that π is stationary for K?

The Metropolis Algorithm on Lipschitz domain

Our framework is the following:

- Ω denotes a bounded connected open subset of $\mathbb{R}^d$ s.t. $\partial\Omega$ has Lipschitz regularity.
- ρ is a measurable function on $\overline{\Omega}$ such that
  - * there exists $m, M > 0$, s.t. $m \leq \rho(x) \leq M$, $\forall x \in \Omega$.
  - * $\int_\Omega \rho(x)dx = 1$
- B_1 denotes the unit ball in $\mathbb{R}^d$ and $|B_1|$ its volume.

We are willing to define a Markov kernel which permit to sample from $\rho(x)dx$.

Introduce the following kernel on Ω:

$$K_{h,\rho}(x,y) = \frac{1}{h^d|B_1|} 1_{|x-y|<h} \min\left(\frac{\rho(y)}{\rho(x)}, 1\right)$$

The Metropolis kernel is given by

$$T_{h,\rho}(x,dy) = m_{h,\rho}(x)\delta_x + K_{h,\rho}(x,y)dy.$$

with

$$m_{h,\rho}(x) = 1 - \int_\Omega K_{h,\rho}(x,y)dy$$

The Metropolis operator associated to this kernel is

$$T_{h,\rho}u(x) = m_{h,\rho}(x)u(x) + \int_\Omega u(y)K_{h,\rho}(x,y)dy$$

Basic properties

- The Metropolis kernel $T_{h,\rho}(x, dy)$ is a Markov kernel $(T_{h,\rho}(1) = 1)$.
- The operator $T_{h,\rho}$ is self-adjoint on $L^2(\Omega, \rho(x)dx)$ and $\| T_{h,\rho} \|_{L^2 \to L^2} = 1$.
- The probability measure $\rho(x)dx$ is stationary for $T_{h,\rho}$.
- $Spec(T_h)$ is discrete near 1 (use this).

Definition

We define the spectral gap of the Metropolis operator $T_{h,\rho}$ as $g(h, \rho) = dist(1, spect(T_h) \setminus \{1\})$. This is the largest constant such that

$$\|u\|^2_{L^2(\rho)} - \langle u, 1\rangle^2_{L^2(\rho)} \leq \frac{1}{g(h, \rho)} \langle u - T_{h,\rho} u, u\rangle_{L^2(\rho)}$$

Theorem 1

Let Ω be an open, connected, bounded, Lipschitz subset of $\mathbb{R}^d$. There exists $h_0 > 0$, $\delta_0 \in]0, 1/2[$ and constants $C_i > 0$ such that for $h \in]0, h_0]$, the following holds true:

- $Spec(T_{h,\rho}) \subset [-1 + \delta_0, 1]$
- 1 is a simple eigenvalue of $T_{h,\rho}$
- The spectral gap $g(h, \rho)$ satisfies

$$C_2 h^2 \leq g(h, \rho) \leq C_3 h^2$$

- $\forall \lambda \in [0, \delta_0]$,

$$\sharp(Spect(T_{h,\rho}) \cap [1 - \lambda, 1]) \leq C(1 + \lambda h^{-2})^{d/2}$$

Total variation estimate

The total variation distance between two probability measures μ, ν is defined by

$$\|\mu - \nu\|_{TV} = \sup_{A \text{ measurable}} |\mu(A) - \nu(A)| = \frac{1}{2} \sup_{f \in L^\infty, |f| \leq 1} |\int f d\mu - \int f d\nu|$$

Theorem 2

Under the same assumption as above, the following estimate holds true for all $n \in \mathbb{N}$:

$$C_4 e^{-ng(h,\rho)} \leq sup_{x \in \Omega} \| T^n_{h,\rho}(x, dy) - \rho(y) dy \|_{TV} \leq C_5 e^{-ng(h,\rho)}.$$

Some references

- Diaconis-Lebeau (08): consider the case of the Metropolis kernel on $X = [0, 1]$ and use semiclassical analysis.
- Lebeau-Michel (09) consider the case of a random walk operator on a Riemannian manifold.
- Lebeau : Cours à l'école d'été du GDR MOAD, aout 2009.
- For an introduction to this topics, see: Diaconis, *The Markov chain Monte Carlo Revolution*, Bull. Amer. Math. Soc. (N.S.) 46 (2009).

Variational approach

Since, $m \leq \rho(x) \leq M$ on Ω, we can easily suppose that $\rho = 1$ (and we denote T_h instead of $T_{h,\rho}$). The spectral gap is the largest constant such that

$$\|u\|_{L^2}^2 - \langle u, 1\rangle_{L^2}^2 \leq \frac{1}{g(h,\rho)} \langle u - T_h u, u\rangle_{L^2}$$

A standard computation shows that

$$\|u\|_{L^2}^2 - \langle u, 1\rangle_{L^2}^2 = \frac{1}{2} \int_{\Omega\times\Omega} |u(x) - u(y)|^2 dxdy := Var(u)$$

$$\langle u - T_h u, u\rangle_{L^2} = \frac{h^{-d}}{2} \int_{\Omega\times\Omega} 1_{|x-y|<h} |u(x) - u(y)|^2 dxdy := \mathcal{E}_h(u).$$

Hence, the spectral gap is the largest constant s.t.

$$Var(u) \leq \frac{1}{g(h)} \mathcal{E}_h(u)$$

The following properties are easy to prove:

- 1 is a simple eigenvalue (use this)
- $g(h,\rho) \leq Ch^2$ (take $u \in C_0^\infty(\Omega)$ such that $\int_\Omega u(x)dx = 0$, $\| u \|_{L^2} = 1$, make a Taylor expansion and use again this)

Lower bound for the spectral gap

Let us show the lower bound on the spectral gap when Ω is convex. For any $u \in L^2(\Omega)$, we have

$$\int_{\Omega\times\Omega} |u(x)-u(y)|^2 dxdy \leq$$

$$Ch^{-1}\sum_{k=0}^{K(h)-1}\int_{\Omega\times\Omega} |u\left(x+k\hbar(y-x)\right)-u\left(x+(k+1)\hbar(y-x)\right)|^2 dxdy,$$

where $K(h)$ is the greatest integer $\leq h^{-1}$ and $K(h)\hbar = 1$.

With the new variables $x' = x + k\hbar(y-x)$, $y' = x + (k+1)\hbar(y-x)$, one has $dx'dy' = \hbar^d dxdy$ and we get

$$\int_{\Omega\times\Omega} |u(x)-u(y)|^2 dxdy \leq$$

$$Ch^{-d-1}K(h)\int_{\Omega\times\Omega} 1_{|x'-y'|<\hbar \mathrm{diam}(\Omega)}|u(x')-u(y')|^2 dx'dy',$$

This yields to

$$Var(u) \leq C' h^{-2}\mathcal{E}_h(u)$$

and proves the lower bound.

Proof of total variation estimates

Let Π_0 be the orthogonal projector in $L^2(\Omega)$ on the space of constant functions

$$\Pi_0(u)(x) = 1_\Omega(x) \int_\Omega u(y)dy. \tag{1}$$

Then, by definition

$$2 \sup_{x_0 \in \Omega} \| T_h^n(x_0, dy) - dy \|_{TV} = \| T_h^n - \Pi_0 \|_{L^\infty \to L^\infty}. \tag{2}$$

Thus, we have to prove that for $h > 0$ small and any n, one has

$$\| T_h^n - \Pi_0 \|_{L^\infty \to L^\infty} \leq C_0 e^{-ng(h,\rho)}. \tag{3}$$

Since $g(h, \rho) = O(h^2)$, we can suppose that $nh^2 >> 1$.

Denote $\lambda_{j,h}$ the eigenvalues of T_h and Π_j the associated spectral projector. We fix $\alpha > 0$ small and use the spectral decomposition $T_h - \Pi_0 = T_{h,1} + T_{h,2}$ with

$$T_{h,1} = \sum_{1-h^{2-\alpha}<\lambda_{j,h}<1} \lambda_{j,h}\Pi j$$

and $T_{h,2}$ spectrally localized in $[-1+\delta_0, 1-h^{2-\alpha}]$. It is easy to see that

$$\|T_h^n - \Pi_0\|_{L^2\to L^2} \leq Ce^{-ng(h,\rho)}.$$

Since, we deal with $L^\infty \to L^\infty$ norm, we need:

- to control $\|\Pi_j\|_{L^2\to L^\infty}$
- a bound on the number of eigenvalues in any interval $[\alpha_h, 1]$ with $1-\delta_0 < \alpha_h < 1-Ch^2$.

For this purpose, we compare our operator with a more simple one.

Comparaison with the random walk on the torus

Since Ω is bounded, it is contained in a large box $]-A,A[^d$. We denote $\Pi=(\mathbb{R}/2A\mathbb{Z})^d$. Since Ω is Lipschitz, using local coordinates, we can define an extension map

$$P: L^2(\Omega)\rightarrow L^2(\Pi)$$

which is also bounded from $H^1(\Omega)$ into $H^1(\Pi)$.
Any function $v\in L^2(\Pi)$ can be extended in Fourier series $v(x)=\sum_{k\in\mathbb{Z}^d} c_k(v)e^{2ik\pi x/A}$. The L^2 and H^1 norm on Π can be expressed as follows

- $\|v\|^2_{L^2(\Pi)}=(2A)^d\sum_k |c_k|^2$.
- $\|v\|^2_{H^1(\Pi)}=(2A)^d\sum_k (1+\frac{4\pi^2k^2}{A^2})|c_k|^2$.

Recall that for $u \in L^2(\Omega)$,

$$\mathcal{E}_h(u) = \langle u - T_h u, u \rangle_{L^2(\Omega)} = \frac{h^{-d}}{2} \int_{\Omega\times\Omega} 1_{|x-y|<h} |u(x) - u(y)|^2 dxdy.$$

For $v \in L^2(\Pi)$, we define

$$\widetilde{\mathcal{E}}_h(v) = \langle u - \tilde{T}_h u, u \rangle_{L^2(\Pi)} = \frac{h^{-d}}{2} \int_{\Pi\times\Pi} 1_{|x-y|<h} |v(x) - v(y)|^2 dxdy.$$

where $\tilde{T}_h$ is the metropolis operator on the torus.

Remark

A simple calculus using the Fourier expansion, shows that $\tilde{T}_h = \Gamma(-h^2\Delta)$ where Γ is a smooth function decreasing to 0 at infinity.

Lemma 1

There exist $C_0, C_1, h_0 > 0$ such that the following holds true for any $h \in]0, h_0]$ and any $u \in L^2(\Omega)$.

$$\mathcal{E}_h(u)/C_0 \leq \widetilde{\mathcal{E}}_h(P(u)) \leq C_0\left(\mathcal{E}_h(u) + h^2\|u\|_{L^2}^2\right). \tag{4}$$

As a by-product, any $u \in L^2(\Omega)$ such that

$$\|u\|_{L^2(\rho)}^2 + h^{-2}\langle(1 - T_h)u, u\rangle_{L^2(\rho)} \leq 1$$

admits a decomposition $u = u_L + u_H$ with $u_L \in H^1(\Omega)$, $\|u_L\|_{H^1} \leq C_1$, and $\|u_H\|_{L^2} \leq C_1 h$.

Proof.

- The first inequality is trivial. The second one is obtained by working in local coordinates for which the boundary is an half-space.
- We observe that (thanks to Parseval identity)

$$\widetilde{\mathcal{E}}_h(v) = \frac{(2A)^d}{2} \sum_k |c_k|^2 \theta(hk/A),$$

$$\theta(\xi) = \int_{|z|\leq 1} |e^{2i\pi\xi z} - 1|^2 dz.$$

The by-product is obtained by projecting the extension $v = P(u)$ on low frequencies $h|k| \leq \delta$ and high frequencies $h|k| > \delta$ for some fixed $\delta > 0$. Hence, it suffices to use the fact that the function θ is quadratic near 0 and has a positive lower bound for $|\xi| \geq \delta$. □

Control of small eigenvalues

Using the preceding Lemma, we show that there exists $\delta_0 > 0$ s.t.

- for any $0 \leq \lambda \leq \delta_0/h^2$,

$$\sharp(Spec(T_h) \cap [1 - h^2\lambda, 1]) \leq C(1 + \lambda)^{d/2}$$

- any eigenfuntion $T_h(u) = \lambda u$ with $\lambda \in [1 - \delta_0, 1]$ satisfies the bound

$$\|u\|_{L^\infty} \leq C_2 h^{-d/2} \|u\|_{L^2}.$$

Using these estimates we get easily:

$$\|T_{2,h}^n\|_{L^\infty \to L^\infty} \leq Ch^{-3d/2} e^{-nh^{2-\alpha}} << e^{-ng(h,\rho)}$$

since $g(h, \rho) \sim h^2$.

Nash inequality

Let $E_\alpha = span(e_{j,h}, 1 - h^{2-\alpha} < \lambda_{j,h} < 1)$. We have the following **Nash inequality**:

Lemma 2

There exists $C, D, \alpha > 0$, s.t. any funtion $u \in E_\alpha$ satisfies:

$$\|u\|_{L^2}^{2+1/D} \leq Ch^{-2}(\|u\|_{L^2}^2 - \|T_h u\|_{L^2}^2 + h^2\|u\|_{L^2}^2)\|u\|_{L^1}^{1/D}.$$

Proof.

- Use Lemma 1 to show that there exists $p > 2$ such that any function $u \in E_\alpha$ satisfies

$$\|u\|_{L^p}^2 \leq Ch^{-2}(\|u\|_{L^2}^2 - \|T_h u\|_{L^2}^2 + h^2\|u\|_{L^2}^2)$$

- Interpolate between L^p and L^1 to get the L^2 estimate. □

Control of $T_{h,1}$

We want to control the norm $\| T_{h,1}^n \|_{L^2 \to L^\infty} = \| T_{h,1}^n \|_{L^1 \to L^2}$.

- Take $g \in L^2$ s.t. $\|g\|_{L^1} = 1$ and denote $c_n = \| T_{h,1}^n g \|_{L^2}^2$. Thanks to the preceding Lemma:

$$c_n^{1+2D} \leq Ch^{-2}(c_n - c_{n+1} + h^2 c_n)$$

  Hence, for $0 \leq n \leq h^{-2}$, $c_n \leq (h^{-2}/(1+n))^{2D}$.
- This permit to show that for some large $n \simeq h^{-2}$,

$$\| T_{h,1}^n \|_{L^2 \to L^\infty} = \| T_{h,1}^n \|_{L^1 \to L^2} = O(1)$$

Combined with $\| T_h^p \|_{L^2 \to L^2} \leq Ce^{-pg(h,\rho)}$, this completes the proof.

Case of a smooth density

If the density ρ is smooth on $\overline{\Omega}$ we can give a more precise description of the spectrum of $T_{h,\rho}$. For simplicity, we assume in this section that $\partial\Omega$ is smooth. Let us introduce the unbounded operator acting on $L^2(\Omega, \rho(x)dx)$, defined by

$$L_\rho(u) = \frac{-\alpha_d}{2}(\triangle u + \frac{\nabla\rho}{\rho}.\nabla u)$$
$$D(L_\rho) = \{u \in H^2(\Omega), \partial_n u|_{\partial\Omega} = 0\}$$

where

$$\alpha_d = \frac{1}{vol(B_1)}\int_{B_1} z_1^2 dz = \frac{1}{d+2}$$

- L_ρ is the self-adjoint realization of the Dirichlet form

$$\frac{\alpha_d}{2}\int_\Omega |\nabla u(x)|^2 \rho(x)dx. \tag{5}$$

- L_ρ has compact resolvant (thanks to Sobolev embeddings).
- We denote

$$Spec(L_\rho) = \{\nu_0 = 0 < \nu_1 < \nu_2 < \dots\}$$

and by $m_j = multiplicity(\nu_j)$. Observe that $m_0 = 1$ since $Ker(L_\rho)$ is spanned by the constant function equal to 1.

Theorem 3

Let Ω be an open, connected, bounded and smooth subset of $\mathbb{R}^d$. Assume that the density ρ is smooth on $\overline{\Omega}$, then for any $R > 0$ and $\varepsilon > 0$ such that $\nu_{j+1} - \nu_j > 2\varepsilon$ for $\nu_{j+2} < R$, there exists $h_1 > 0$ such that one has for all $h \in]0, h_1]$,

$$Spec\left(\frac{1 - T_{h,\rho}}{h^2}\right) \cap]0, R] \subset \cup_{j\geq 1}[\nu_j - \varepsilon, \nu_j + \varepsilon], \tag{6}$$

and the number of eigenvalues of $\frac{1-T_{h,\rho}}{h^2}$ in the interval $[\nu_j - \varepsilon, \nu_j + \varepsilon]$ is equal to m_j.

A simple quasimode calculus

Assume $\rho = 1$ and $\partial\Omega$ is smooth. Let $\lambda > 0$ and $u \in C^\infty(\overline{\Omega})$ satisfy

$$(-\frac{\alpha_d}{2}\Delta - \lambda)u = 0 \text{ in } \Omega \quad \text{and} \quad \partial_n u_{|\partial\Omega} = 0.$$

- For $x \in \Omega$ s.t. $dist(x, \partial\Omega) > h$, Taylor expansion shows that

$$T_h u(x) - u(x) = \int_{|z|<1, x+hz\in\Omega} (u(x+hz) - u(x))dz$$

$$= h\sum_{j=1}^{d} \partial_{x_j} u(x) \int_{|z|<1} z_j dz + \alpha_d h^2 \Delta u(x) + O_{L^\infty}(h^4)$$

$$= \frac{\alpha_d}{2} h^2 \Delta u(x) + O_{L^\infty}(h^4)$$

where the term of order h and h^3 vanish for parity reason.

- For $x \in \Omega$ s.t. $dist(x, \partial\Omega) < h$, we use local coordinates such that $\Omega = \{(x_1, x') \in \mathbb{R}^d, x_1 > 0\}$. Taylor expansion shows that

$$T_h u(x) - u(x) = \int_{|z|<1, x_1+hz_1>0} (u(x+hz) - u(x))dz$$

$$= h \sum_{j=1}^{d} \partial_{x_j} u(x) \int_{|z|<1, x_1+hz_1>0} z_j dz + O_{L^\infty}(h^2)$$

  - Parity argument $\Longrightarrow$ term of index $j \geq 2$ vanish.
  - $\partial_n u_{|\partial\Omega} = 0$ and $dist(x, \partial\Omega) < h \Longrightarrow$ term of index $j = 1$ is $O_{L^\infty}(h^2)$.

  Since $meas(\{dist(x, \partial\Omega) < h\}) = O(h)$, it follows that

$$1_{dist(x,\partial\Omega)<h}(T_h u - u) = O_{L^2}(h^{\frac{5}{2}}).$$

Combining the two estimates, we get

$$T_h u - (1 - h^2\lambda)u = O(h^{\frac{5}{2}}).$$

Application to Random Placement of Non-Overlapping Balls

We consider the initial problem that motivated the works of Metropolis et al. Given an open set $\Omega \subset \mathbb{R}^d$ and $N \in \mathbb{N}$ we consider the set of all possible positions in Ω for N non-overlapping balls of radius $\epsilon > 0$. This can be identified to the possible locations for their centers

$$\mathcal{O}_{N,\epsilon} = \left\{ x = (x_1, \dots, x_N) \in \Omega^N, \forall\, 1 \leq i < j \leq N, |x_i - x_j| > \epsilon \right\}.$$

The problem we adress is to sample from the uniform distribution, according with the following Metropolis algorithm:
Starting from a configuration $(X_1, \dots, X_N)$ we choose a ball at random and move it uniformly at random in a small ball of radius $h > 0$. If it results in an admissible configuration, "we keep" the move. Otherwise we don't move and try again.

This is associated to the following Markov kernel (where $\varphi = 1_{B_{\mathbb{R}^d}(0,1)}$)

$$K_h(x, dy) = \frac{1}{N}\sum_{j=1}^{N} \delta_{x_1}\otimes\cdots\otimes\delta_{x_{j-1}}\otimes h^{-d}\varphi\left(\frac{x_j - y_j}{h}\right) dy_j\otimes\delta_{x_{j+1}}\otimes\cdots\otimes\delta_{x_N},$$

and the associated Metropolis operator on $L^2(\mathcal{O}_{N,\epsilon})$

$$T_h(u)(x) = m_h(x)u(x) + \int_{\mathcal{O}_{N,\epsilon}} u(y)K_h(x, dy),$$

with

$$m_h(x) = 1 - \int_{\mathcal{O}_{N,\epsilon}} K_h(x, dy).$$

Proposition

There exists $\alpha > 0$ such that for $N\epsilon \leq \alpha$, the set $\mathcal{O}_{N,\epsilon}$ is connected, Lipschitz and quasi-regular.

Proof. The proof is rather technical. The quasiregularity is notion used to replace "smooth" by "Lipschitz".
To prove the "Lipschitz boundary" use the following caraterisation:
A domain $\mathcal{O} \subset \mathbb{R}^p$ has Lipschitz boundary iff it satisfies the following cone property:
$\forall a \in \partial\mathcal{O}, \exists \delta > 0, \exists \nu_a \in S^{p-1}, \forall b \in B(a,\delta) \cap \partial\mathcal{O}$ we have

$$b + \Gamma_+(\nu_a, \delta) \subset \mathcal{O} \quad \text{and} \quad b + \Gamma_-(\nu_a, \delta) \subset \mathbb{R}^p \setminus \overline{\mathcal{O}}.$$

where for $\nu \in S^p$,

$$\Gamma_+(\nu_a, \delta) = \{\xi \in \mathbb{R}^p,\ \pm\langle \xi, \nu\rangle > (1-\delta)|\xi|,\ |\langle \xi, \nu\rangle| < \delta\}$$

Thanks to the preceding proposition, we can consider the Neumann Laplacian $|\Delta|_N$ on $\mathcal{O}_{N,\epsilon}$ defined by

$$|\Delta|_N = -\frac{\alpha_d}{2N}\Delta,$$
$$D(|\Delta|_N) = \left\{u \in H^1(\mathcal{O}_{N,\epsilon}),\ -\Delta u \in L^2(\mathcal{O}_{N,\epsilon}), \partial_n u|_{\partial\mathcal{O}_{N,\epsilon}} = 0\right\}.$$

We still denote $0 = \nu_0 < \nu_1 < \nu_2 < \dots$ the spectrum of $|\Delta|_N$ and m_j the multiplicity of ν_j.

Theorem (part 1)

Let $N \geq 2$ and $\epsilon > 0$ small be fixed. Let $R > 0$ be given and $\beta > 0$ small. Then, there exists $h_0 > 0$, $\delta_0 \in]0, 1/2[$ and constants $C_i > 0$ such that for any $h \in]0, h_0]$, the following hold true:

i) The spectrum of T_h is a subset of $[-1 + \delta_0, 1]$, 1 is a simple eigenvalue of T_h, and $Spec(T_h) \cap [1 - \delta_0, 1]$ is discrete. Moreover,

$$Spec\left(\frac{1 - T_h}{h^2}\right) \cap]0, R] \subset \cup_{j \geq 1} [\nu_j - \beta, \nu_j + \beta];$$

$$\sharp Spec\left(\frac{1 - T_h}{h^2}\right) \cap [\nu_j - \beta, \nu_j + \beta] = m_j \qquad \forall \nu_j \leq R;$$

and for any $0 \leq \lambda \leq \delta_0 h^{-2}$, the number of eigenvalues of T_h in $[1 - h^2\lambda, 1]$ (with multiplicity) is bounded by $C_1(1 + \lambda)^{dN/2}$.

Theorem (part 2)

ii) The spectral gap $g(h)$ satisfies

$$\lim_{h\to 0^+} h^{-2}g(h) = \nu_1$$

and the following estimate holds true for all $n \in \mathbb{N}$:

$$\sup_{x\in\mathcal{O}_{N,\epsilon}} \| T_h^n(x,dy) - \frac{dy}{vol(\mathcal{O}_{N,\epsilon})}\|_{TV} \leq C_4 e^{-ng(h)}.$$