Gibbs/Metropolis algorithms on a convex polytope

Persi Diaconis* Gilles Lebeau[†‡] Laurent Michel[§]

Departments of Mathematics and Statistics,
Stanford University
and
Département de Mathématiques,
Université de Nice Sophia-Antipolis

Abstract

This paper gives sharp rates of convergence for natural versions of the Metropolis algorithm for sampling from the uniform distribution on a convex polytope. The singular proposal distribution, based on a walk moving locally in one of a fixed, finite set of directions, needs some new tools. We get useful bounds on the spectrum and eigenfunctions using Nash and Weyl-type inequalities. The top eigenvalues of the Markov chain are closely related to the Neuman eigenvalues of the polytope for a novel Laplacian.

1 Introduction

1.1 Overview

The Metropolis algorithm and the Gibbs sampler (also known as Glauber dynamics) are often used together as one of the basic tools of scientific computation. We treat the following example: let Ω be a polyhedral convex set in d dimensions. To sample from the uniform distribution on Ω, from a point x in Ω, pick a direction e from a fixed finite collection. Set $y = x + ue$ where u is chosen uniformly in $[-h, h]$. If $y \in \Omega$, move to y. Else, stay at x. Under a mild generality condition on the set of directions in relation to Ω, this Markov chain converges to the uniform distribution on Ω. Our main result gives a sharp determination of the exponential rate of convergence of this algorithm. It is $ce^{-ng(h)}$ with $g(h)$ asymptotic to $h^2\nu$ for ν the first non zero eigenvalue of a novel Laplacian defined on Ω with Neumann condition on the boundary.

Sampling from a convex set is a practical problem. For example, choosing a uniformly distributed 100×100 doubly stochastic matrix [1] or a uniformly distributed 100×100 tri-diagonal doubly stochastic matrix [5]. It is also a basic problem of study in theoretical computer science [9, 10]. Many algorithms have been proposed and studied. A readable textbook description of the Gibbs sampler is in Liu [8]. See [6] for a review of rigorous results for the Metropolis algorithm in finite spaces. The popular *hit and run algorithm* [2, 11, 15] was introduced for this purpose.

*Supported in part by NSF grant 0804324
†Supported in part by ANR-06-BLAN-0250-03
‡Corresponding author: Parc Valrose 06108 Nice Cedex 02, France; `lebeau@unice.fr`
§Supported in part by ANR-06-BLAN-0250-03 and ANR-09-JCJC-0099-01

Hit and run makes long moves and will probably be preferred in practice to the local algorithms studied here.

Spectral techniques for analysis of the Metropolis algorithm on continuous spaces are developed in [3, 4, 7]. The proposal distributions there are "ball walks" choosing from the uniform distribution on the interior of a ball. The discrete set of directions studied here is widely used in practice and necessitates new ideas. Present problems can also be studied by Harris recurrence techniques [12, 14] and by the path techniques of Yuen [16]. These give useful results but do not get the sharp rates on the exponents derived here.

The remainder of this section gives a careful description of the Markov chain and the geometric connection between the underlying directions and the convex set Ω required for ergodicity. Section 2 gives bounds on the spectrum and eigenvectors using Nash inequalities and Weyl-type inequalities. Section 3 uses this spectral information to get rates of convergence. Section 4 proves that our operator (suitably rescaled) converges, in the strong resolvent sense, to a novel Laplace operator on Ω with Neumann boundary conditions. A similar convergence of the ball walk Metropolis operator to the usual Neumann Laplacian is a key ingredient of [4, 7]. The final section shows how to modify the argument to handle a continuous choice of direction.

1.2 Basic definitions

Let Ω be an open convex polytope in $\mathbb{R}^d$, $d \geq 2$. Thus there exists linear forms $\ell_j : \mathbb{R}^d \to \mathbb{R},\ j = 1, \ldots, m$ and real numbers b_j such that

$$\Omega = \Big\{x \in \mathbb{R}^d,\ \forall j = 1, \ldots, m,\ \ell_j(x) > b_j\Big\} \tag{1.1}$$

Assume also that Ω is bounded and non empty.

Consider $\mathfrak{E} = \{e_1, \ldots, e_p\}$ a family of vectors in $\mathbb{R}^d$. For any $j \in \{1, \ldots, p\}$ we introduce the operator acting on continuous functions $M_{j,h}f(x) = m_{j,h}(x)f(x) + K_{j,h}f(x)$, where

$$K_{j,h}(f)(x) = \frac{1}{2}\int_{t\in[-1,1]} 1_\Omega(x + hte_j)f(x + hte_j)\,dt \tag{1.2}$$

and $m_{j,h}(x) = 1 - K_{j,h}(1)(x)$.

The local Metropolis operator associated to the family $\mathfrak{E}$ is

$$M_h(f)(x) = \frac{1}{p}\sum_{j=1}^{p}\left(K_{j,h}f(x) + m_{j,h}(x)f(x)\right). \tag{1.3}$$

In the sequel, denote $m_h = \frac{1}{p}\sum_{j=1}^{p} m_{j,h}$ and $K_h = \frac{1}{p}\sum_{j=1}^{p} K_{j,h}$. Let $M_h(x, dy)$ be the Markov kernel associated to this operator. This defines a bounded self-adjoint operator on $L^2(\Omega)$. Moreover, since $M_h(1) = 1$, $\| M_h \|_{L^2\to L^2} = 1$. Thus the probability measure $\frac{dx}{\mathrm{vol}(\Omega)}$ on Ω is stationary. For $n \geq 1$, denote by $M_h^n(x, dy)$ the kernel of the iterated operator $(M_h)^n$. For any $x \in \Omega$, $M_h^n(x, dy)$ is a probability measure on Ω, and our main goal is to get some estimates on the rate of convergence, when $n \to +\infty$, of the probability $M_h^n(x, dy)$ toward the stationary probability $\frac{dy}{\mathrm{vol}(\Omega)}$.

A good example to keep in mind is the case where $\Omega = A_N$ is the set of $N \times N$ doubly stochastic matrices. In other words,

$$A_N = \left\{(a_{i,j})_{1\leq i,j\leq N},\ \forall i, j,\ a_{i,j} > 0,\ \sum_k a_{ik} = \sum_k a_{kj} = 1\right\}. \tag{1.4}$$

The set A_N can be viewed as convex open polytope in $A_N^0 = \{(a_{i,j})_{1\le i,j\le N},\ \sum_k a_{ik} = \sum_k a_{kj} = 1\}$. A good way to sample from A_N is to use the Metropolis strategy in the following manner. Starting from a matrix $A \in A_N$ choose two distinct rows R_{i_1}, R_{i_2} and two distinct columns C_{j_1}, C_{j_2} at random. Denote $\vec{i} = (i_1, i_2, j_1, j_2)$ and $F = F(\vec{i})$ the matrix such that $F_{i,j} = \delta_{i_1 j_1} - \delta_{i_1,j_2} - \delta_{i_2 j_1} + \delta_{i_2 j_2}$. For $h > 0$ given, build the family of matrices $(\tilde{A}(t) = A + tF(\vec{i}))_{t\in[-h,h]}$. For any $t \in \mathbb{R}$ the matrix $\tilde{A}(t)$ belongs to the set A_N^0. Taking $t \in [-h, h]$ at random and keeping the move $A \to \tilde{A}(t)$ only if it results in an element of A_N, we are exactly in the above situation with $\mathfrak{E} = \{F(\vec{i})\}$. This algorithm is used in [1] to study things like the distribution of typical entries or the eigenvalues of random doubly stochastic matrices.

Let us go back to the general problem. From the definition of Ω, a point $x \in \mathbb{R}^d$ belongs to $\partial\Omega$ iff there exists a partition $I \cup J = \{1, \dots, m\}$ such that $I \neq \emptyset$ and

$$\forall i \in I,\ \ell_i(x) = b_i \quad \text{and} \quad \forall j \in J,\ \ell_j(x) > b_j. \tag{1.5}$$

Define the following function $\mathfrak{c} : \mathbb{R}^d \to \mathbb{N} \cup \{+\infty\}$ by

$$\begin{aligned} \mathfrak{c}(x) &= 0 && \text{if } x \in \Omega \\ &= +\infty && \text{if } x \in \mathbb{R}^d \setminus \overline{\Omega} \\ &= \mathrm{card}(I) && \text{if } x \in \partial\Omega. \end{aligned} \tag{1.6}$$

To proceed, the following geometric condition is needed; it shows how the generating set $\mathfrak{E}$ must be related to the convex set Ω. Proposition 1.5 shows the condition is equivalent to M_h having a spectral gap.

Definition 1.1. The family $\mathfrak{E}$ is weakly incoming to the set Ω if for any point $x_0 \in \partial\Omega$ there exists $\epsilon > 0,\ \theta \in \{\pm 1\}$ and $e \in \mathfrak{E}$ such that, for $\mathfrak{c}$ defined in (1.6),

$$\mathfrak{c}(x_0 + \theta t e) < \mathfrak{c}(x_0) \qquad \forall t \in]0; \epsilon]. \tag{1.7}$$

The following observation is simple and fundamental. Suppose that $\mathfrak{E}$ is weakly incoming, then $\mathrm{span}(\mathfrak{E}) = \mathbb{R}^d$. Indeed, otherwise there is a hyperplane $H = (\mathbb{R}\nu)^\perp$ of $\mathbb{R}^d$ such that $\mathrm{span}(\mathfrak{E}) \subset H$. Since $\overline{\Omega}$ is compact, the function $x \in \overline{\Omega} \mapsto \langle x, \nu\rangle$ would have a global minimum in some $x_0 \in \partial\Omega$. Since Ω is open, $\Omega \subset x_0 + H^+$, where $H^+ = \{y \in \mathbb{R}^d, \langle y, \nu\rangle > 0\}$. As $\mathfrak{E}$ is weakly incoming, there is $u \in \mathrm{span}(\mathfrak{E})$ such that $\mathfrak{c}(x_0 + u) = 0$. In other words, $x_0 + u \in \Omega \cap (x_0 + H)$. This contradicts $\Omega \subset x_0 + H^+$.

Example 1.1. *Consider Ω the convex hull of on equilateral triangle (ABC) in $\mathbb{R}^2$ and $\mathfrak{E} = \{e_1, e_2\}$ like on Figure 1. For $\alpha \in]0, \pi/3]$, $\mathfrak{E}$ is weakly incoming to Ω whereas for $\alpha \in]\pi/3, \pi[$, condition (1.7) is satisfied in every point x_0 of the boundary excepted in point A.*

Remark 1.2. *In the above case of doubly stochastic matrices, the set $\mathfrak{E} = \{F(\vec{i})\}$ is weakly incoming. Indeed, if A is in the boundary of A_N, there exists i_1 and j_1 such that $A_{i_1 j_1} = 0$. Since A is doubly stochastic, there exists i_2, j_2 such that $A_{i_1 j_2} > 0$ and $A_{i_2 j_1} > 0$. Let $\epsilon = \min(A_{i_1 j_2}, A_{i_2 j_1})/2$, then for all $t \in]0, \epsilon]$, $\mathfrak{c}(A + tF(i_1, i_2, j_1, j_2)) < \mathfrak{c}(A)$.*

Denote $H_k = \ker(\ell_k)$ and let ν_k be the unit vector such that $\ell_k(\nu_k) > 0$ and $H_k^+ = \{y \in \mathbb{R}^d,\ \ell_k(y) > b_k\}$.

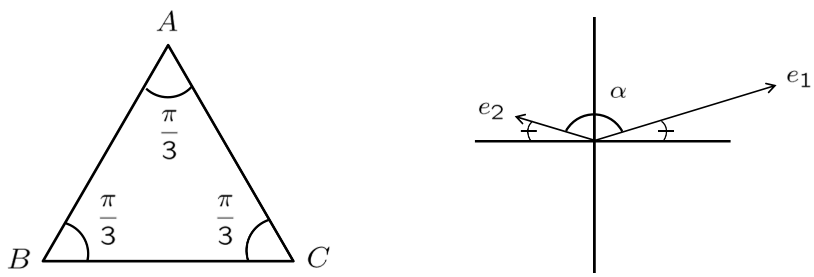

Figure 1: Weakly incoming condition in the case of an equilateral triangle.

Definition 1.2. Let $u \in \mathbb{R}^d \setminus \{0\}$. The vector u is incoming to H_k if $\langle u, \nu_k \rangle \geq 0$. Further, u is strictly incoming to H_k if $\langle u, \nu_k \rangle > 0$; u is strictly outgoing to H_k if $\langle u, \nu_k \rangle < 0$; u is parallel to H_k if $u \in H_k$.

Lemma 1.3. *Suppose that $\mathfrak{E}$ is weakly incoming to Ω and let $x_0 \in \overline{\Omega}$ and $k = \mathfrak{c}(x_0)$. There exists $r > 0$ and $I \subset \{1, \ldots, m\}$ such that $\sharp I = k$ and for $B(x_0, r)$ the open ball of radius r about x_0,*

$$\Omega \cap B(x_0, r) = B(x_0, r) \cap \left(\cap_{i \in I} H_i^+\right). \tag{1.8}$$

Further, there exists $\beta_1, \ldots, \beta_k, \in \{1, \ldots, p\}$, a family $(\theta_n)_{n=1,\ldots,k}$ of numbers in $\{\pm 1\}$ and a bijection $\{1, \ldots, k\} \ni n \mapsto i_n \in I$ such that for all $n \in \{1, \ldots, k\}$,

$$\begin{aligned} &\theta_n e_{\beta_n} \textit{ is strictly incoming to } H_{i_n}; \\ &\theta_n e_{\beta_n} \textit{ is incoming to } H_{i_m}, \quad \forall m > n. \end{aligned} \tag{1.9}$$

Proof. Proceed by induction on $k = \mathfrak{c}(x_0)$. When $k = 0$, there is nothing to prove.

Suppose that the property holds true at rank $k' \leq k-1$ and let $x_0 \in \partial\Omega$ be such that $\mathfrak{c}(x_0) = k$. By definition of Ω, there exists $r > 0$ and $I \subset \{1, \ldots, m\}$ with $\sharp I = k$ such that

$$\Omega \cap B(x_0, r) = B(x_0, r) \cap (\cap_{i \in I} H_i^+) \tag{1.10}$$

Since $\mathfrak{E}$ is weakly incoming to Ω, there exists $q \in \{1, \ldots, k\}$, $\theta_1 = \pm 1$, $\beta_1 \in \{1, \ldots, p\}$ and $i_1, \ldots, i_q \in I$ such that $\theta_1 e_{\beta_1}$ is strictly incoming to H_{i_n} for $n = 1, \ldots, q$ and $\theta_1 e_{\beta_1}$ is parallel to H_i for $i \in I' := I \setminus \{i_1, \ldots, i_q\}$. By definition of Ω there exists x_0' close to x_0 and $r' > 0$ such that $B(x_0', r') \cap \Omega = B(x_0', r') \cap (\cap_{i \in I'} H_i^+)$. From the induction hypothesis, there exists $\beta_{q+1}, \ldots, \beta_k \in \{1, \ldots, p\}$, $\theta_{q+1}, \ldots, \theta_k = \pm 1$ and a bijection $\{q+1, \ldots, k\} \ni n \mapsto i_n \in I'$ such that for all $n \geq q+1$,

$$\begin{aligned} &\theta_n e_{\beta_n} \text{ is strictly incoming to } H_{i_n}; \\ &\theta_n e_{\beta_n} \text{ is incoming to } H_{i_m} \qquad \forall m > n. \end{aligned} \tag{1.11}$$

and the proof is complete. ☐

Corollary 1.4. *Suppose that $\mathfrak{E}$ is weakly incoming to Ω. There exist $r > 0$ and $\epsilon \in]0, 1]$, such that for all $x_0 \in \overline{\Omega}$, there exists $q \in \{1, \ldots, p\}$ and $\theta_q = \pm 1$ such that*

$$x + t\theta_q e_q \in \Omega \qquad \forall x \in B(x_0, r) \cap \Omega, \ \forall t \in [0, \epsilon]. \tag{1.12}$$

Proof. The fact that $r, \epsilon > 0$ can be chosen uniformly with respect to x_0 follows easily from compactness of $\overline{\Omega}$. The statement is trivial when $x_0 \in \Omega$. Suppose that $x_0 \in \partial\Omega$. From Lemma 1.3, there exists $r > 0$ and $I = \{i_1, \ldots, i_k\} \subset \{1, \ldots, p\}$ such that

$$\Omega \cap B(x_0, 2r) = B(x_0, 2r) \cap \left(\cap_{i\in I} H_i^+\right) \tag{1.13}$$

and $\theta_1 = \pm 1,\ \beta_1 \in \{1, \ldots, p\}$ such that

$$\begin{aligned} &\theta_1 e_{\beta_1} \text{ is strictly incoming to } H_{i_1};\\ &\theta_1 e_{\beta_1} \text{ is incoming to } H_{i_q},\ \forall q > 1. \end{aligned} \tag{1.14}$$

Let $x \in B(x_0, r) \cap \Omega$ and $\epsilon \in]0, r[$. Then

$$\langle \nu_i, \theta_1 \beta_1 \rangle \geq 0,\ \forall i \in I \Longrightarrow x + t\theta_1 e_{\beta_1} \in \Omega. \tag{1.15}$$

Thanks to (1.14), the left hand side of the above property is satisfied and the proof is complete. □

Proposition 1.5. *The family $\mathfrak{E}$ is weakly incoming to Ω iff 1 is not in the essential spectrum of M_h.*

Proof. If $\mathfrak{E}$ is weakly incoming to Ω, 1 is not in the essential spectrum of M_h thanks to Proposition 2.2 of this paper and Theorem 1.1 in [4].

Suppose now that $\mathfrak{E}$ is not weakly incoming to Ω. This means that there exists $x_0 \in \partial\Omega$ such that (1.7) does not hold. Let $k = \mathfrak{c}(x_0)$. There exists a neighborhood V of x_0 and $I \subset \{1, \ldots, m\}$ with $\sharp I = k$ such that $V \cap \Omega = V \cap (\cap_{i\in I} H_i)$. Then, for any $\theta = \pm 1$ and any $j \in \{1, \ldots, p\}$, the following holds true:

If θe_j is strictly incoming to one of the $(H_i)_{i\in I}$,
then θe_j is strictly outgoing to one of the $(H_i)_{i\in I}$.

Otherwise, there is $j \in \{1, \ldots, p\}$ and $\theta = \pm 1$ such that θe_j is strictly incoming to one of the $(H_i)_{i\in I}$ and incoming to the other. Then for $t > 0$ small enough, $\mathfrak{c}(x_0 + \theta t e_j) < \mathfrak{c}(x_0)$.

Hence, assume that there exists $r \geq 1$ such that

- for any $j \in \{1, \ldots, r\}$, e_j and $-e_j$ are strictly outgoing to some of the $(H_i)_{i\in I}$;
- for any $j \in \{r+1, \ldots, p\}$, e_j is parallel to the $(H_i)_{i\in I}$.

Recall that ν_i denotes the unit incoming orthogonal vector to H_i. Let $W = \text{span}(\nu_i, i \in I)$ and near x_0 use the variable $x = x_0 + (x', x'')$ with $x' \in W$ and $x'' \in W^\perp$. Let $\chi(x') = 1\!\!1_{\frac{1}{2}<|x'|<1}$ and for $\lambda, h > 0$ denote $f_{\lambda h}(x) = (\lambda h)^{-dim(W)/2} \chi(\frac{x'}{\lambda h})$. Since any $v \in W^\perp$ is parallel to the $(H_i)_{i\in I}$, there exists $\lambda_0, c_0 > 0$ such that for all $h \in]0, 1]$ and $\lambda \in]0, \lambda_0]$, $\|f_{\lambda h}\|_{L^2(\Omega)} \geq c_0$.

For any $j \in \{r+1, \ldots, p\}$, e_j is parallel to the $(H_i)_{i\in I}$. Hence, the function $t \mapsto f_{\lambda h}(x + hte_j)$ is constant and $(M_{j,h} - 1) f_{\lambda h}(x) = 0$.

On the other hand, for any $j \in \{1, \ldots, r\}$ there exists $i_j, i'_j \in I$ such that e_j is strictly outgoing to H_{i_j} and $-e_j$ is strictly outgoing to $H_{i'_j}$. Consequently, there exists $\gamma_j, \delta_j > 0$ such that for $t > 0$,

$$\begin{aligned} &x \in \Omega \text{ and } x - te_j \in \Omega \Longrightarrow \text{dist}(x - te_j, H_{i'_j}) \leq \text{dist}(x, H_{i'_j}) - \gamma_j t\\ &x \in \Omega \text{ and } x + te_j \in \Omega \Longrightarrow \text{dist}(x + te_j, H_{i_j}) \leq \text{dist}(x, H_{i_j}) - \delta_j t. \end{aligned} \tag{1.16}$$

Let us compute the potential $m_{j,h}$ on the support of $f_{\lambda h}$. For $x \in \mathrm{supp}(f_{\lambda h})$, $|x_j| \leq \lambda h$ for all $j = 1, \ldots, r$. In particular $\mathrm{dist}(x, H_{i_j}) \leq \lambda h$ and $\mathrm{dist}(x, H_{i'_j}) \leq \lambda h$ and thanks to (1.16),

$$\begin{aligned} 1 - m_{j,h}(x) &= \int_0^1 1_\Omega(x + hte_j) + 1_\Omega(x - hte_j)\,dt \\ &\leq \int_{0 \leq t \leq \mathrm{dist}(x,H_{i_j})/(\delta_j h)} dt + \int_{0 \leq t \leq \mathrm{dist}(x,H_{i'_j})/(\gamma_j h)} dt \leq \lambda \left(\frac{1}{\gamma_j} + \frac{1}{\delta_j} \right). \end{aligned} \tag{1.17}$$

Finally,

$$\begin{aligned} \langle (1 - M_h) f_{\lambda h}, f_{\lambda h} \rangle_{L^2(\Omega)} &= \frac{1}{p} \sum_{j=1}^{r} \langle (1 - M_{j,h}) f_{\lambda h}, f_{\lambda h} \rangle_{L^2(\Omega)} \\ &\leq \frac{1}{p} \sum_{j=1}^{r} \int_\Omega (1 - m_{j,h}(x))\, |f_{\lambda h}(x)|^2\,dx \leq C\lambda \|f_{\lambda h}\|^2_{L^2(\Omega)}. \end{aligned} \tag{1.18}$$

Here we used the fact that for any non-negative fonction f, one has $\langle K_{j,h} f, f \rangle \geq 0$. Finally, we conclude by taking $\lambda = 2^{-n} \to 0$ as $n \to \infty$. Indeed, the functions $f_{2^{-n}h}$ are mutually orthogonal. Their norm is bounded uniformly from below and they satisfy $0 \leq \langle (1 - M_h) f_{2^{-n}h}, f_{2^{-n}h} \rangle \leq C2^{-n}$. □

2 Spectral Analysis of the Metropolis Operator

This section is devoted to the analysis of the spectral theory of the Metropolis operator. For this purpose, we introduce a Laplace operator associated to the family $\mathfrak{E}$ to be used as a model. For any $e \in \mathbb{R}^d \setminus \{0\}$ and any smooth function u, define $\partial_e u(x) = \frac{d}{dt}(u(x + te))_{|t=0}$. Then, consider the operator $\Delta_{\mathfrak{E}}$, defined by

$$\begin{aligned} \Delta_{\mathfrak{E}} u &= \frac{1}{6p} \sum_{j=1}^{p} \partial_{e_j}^2 u \\ D(\Delta_{\mathfrak{E}}) &= \left\{ u \in H^1(\Omega),\, \Delta_{\mathfrak{E}} u \in L^2, \partial_{n,\mathfrak{E}} u_{|\partial\Omega} = 0 \right\} \end{aligned} \tag{2.1}$$

with $\partial_{n,\mathfrak{E}} u(x) = \frac{1}{p} \sum_{j=1}^{p} \langle n(x), e_j \rangle \partial_{e_j} u(x)$, $n(x)$ denoting the outgoing normal vector to the boundary at point x. If the domain Ω has smooth boundary, the normal derivative is well defined. In the case where it is Lipschitz, it can be defined by duality in the following way.

Define first the gradient and divergence associated to the family $\mathfrak{E}$, by $\mathrm{div}_{\mathfrak{E}} u = \frac{1}{p} \sum_{j=1}^{p} \partial_{e_j} u_j$ for any $u = (u_1, \ldots, u_p)$ and $\nabla_{\mathfrak{E}} u = (\partial_{e_1} u, \ldots, \partial_{e_p} u)$. Then, define a trace operator $\gamma_{\mathfrak{E}}$ by

$$\gamma_{\mathfrak{E}} : \left\{ u \in \left(L^2(\Omega) \right)^p, \mathrm{div}_{\mathfrak{E}}(u) \in L^2(\Omega) \right\} \to H^{-1/2}(\partial\Omega) \tag{2.2}$$

and and for $v \in H^1(\Omega)$,

$$\int_\Omega \mathrm{div}_{\mathfrak{E}}(u)(x) v(x)\,dx = -\frac{1}{p} \int_\Omega \langle u(x), \nabla_{\mathfrak{E}} v(x) \rangle_{\mathbb{C}^p}\,dx + \int_{\partial\Omega} \gamma_{\mathfrak{E}}(u) v_{|\partial\Omega}\,d\sigma(x). \tag{2.3}$$

In particular, for $u \in H^1(\Omega)$ satisfying $\Delta_{\mathfrak{E}} u = \frac{1}{6} \mathrm{div}_{\mathfrak{E}} \nabla_{\mathfrak{E}} u \in L^2(\Omega)$ define $\partial_{n,\mathfrak{E}} u_{|\partial\Omega} = \gamma_{\mathfrak{E}}(\nabla_{\mathfrak{E}} u) \in H^{-1/2}(\partial\Omega)$ and the set $D(-\Delta_{\mathfrak{E}})$ is well defined. The Dirichlet form associated with $-\Delta_{\mathfrak{E}}$ is

$$\mathcal{E}_{\mathfrak{E}}(u) = \frac{1}{6p} \sum_{j=1}^{p} \int_\Omega |\partial_{e_j} u(x)|^2\,dx. \tag{2.4}$$

Let $\mathfrak{E}_0$ be the canonical basis in $\mathbb{R}^d$. Then, $\Delta_{\mathfrak{E}_0} = \frac{1}{6d}\Delta$ where Δ is the usual Laplace operator and $\mathcal{E}_{\mathfrak{E}_0}(f) = \frac{1}{6d}\int_\Omega |\nabla f|^2 dx$ is the usual Dirichlet form. Since $\mathfrak{E} = \{e_1, \dots, e_p\}$ spans $\mathbb{R}^d$, a simple calculation shows that there exists a constants $C > 0$ such that

$$C^{-1}\mathcal{E}_{\mathfrak{E}_0}(f) \le \mathcal{E}_{\mathfrak{E}}(f) \le C\mathcal{E}_{\mathfrak{E}_0}(f). \tag{2.5}$$

Then, it is standard to show that $-\Delta_{\mathfrak{E}}$ is the self-adjoint realization of the Dirichlet form $\mathcal{E}_{\mathfrak{E}}$. A standard argument using Sobolev embedding shows that $-\Delta_{\mathfrak{E}}$ has compact resolvant. Denote its spectrum by $\nu_0 = 0 < \nu_1 < \nu_2 < \dots$ and by m_j the associated multiplicities. Observe that $m_0 = 1$. Section 4 shows that $h^{-2}(1 - M_h)$ converges to $-\Delta_{\mathfrak{E}}$ in the strong resolvent sense so that eigenvalues and eigenvectors converge; see [13].

The main theorem of this section follows.

Theorem 2.1. *Suppose that $\mathfrak{E}$ is weakly incoming to Ω, then the following hold true.*

i) There exists $h_0 > 0$, $\delta_0 \in]0, \frac{1}{2}[$ and a positive constant C such that for any $h \in]0, h_0]$, the spectrum of M_h is a subset of $[-1 + \delta_0, 1]$, 1 is a simple eigenvalue and $\mathrm{Spec}(M_h) \cap [1 - \delta_0, 1]$ is discrete.

ii) For any $h \in]0, h_0]$ and $0 \le \lambda \le \delta_0 h^{-2}$, the number of eigenvalues of M_h in $[1 - h^2\lambda, 1]$ (with multiplicity) is bounded by $C(1 + \lambda)^{d/2}$.

iii) 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 $h \in]0, h_1]$,

$$\mathrm{Spec}\left(\frac{1 - M_h}{h^2}\right) \cap]0, R] \subset \cup_{j \ge 1}[\nu_j - \varepsilon, \nu_j + \varepsilon], \tag{2.6}$$

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

A consequence of this theorem is that M_h has a spectral gap $g(h) = 1 - \sup(\mathrm{Spec}(M_h) \setminus \{1\}) > 0$ and that $\lim_{h \to 0^+} h^{-2}g(h) = \nu_1$. This will be used in the proof of total variation estimates.

The strategy used to prove the first part of Theorem 2.1 is very close to the one given in [4]. First, show that some iterate of the Markov kernel "controls" the random walk on a ball. Next, this ball walk on the polytope is compared to the same walk on a large torus containing Ω. Finally the information on the torus is transferred back to the original problem.

The proof of the last part of Theorem 2.1 is slightly different from the proof in [4]. Indeed, the starting point of the analysis in [4] is that for regular function φ with normal derivatives vanishing on the boundary, $h^{-2}(1 - T_h)\varphi$ is close to $-\Delta\varphi$ up to the boundary, where T_h is the Metropolis operator associated to the kernel $\mathrm{vol}(B(0,1))^{-1}h^{-d}1_{|x-y|<h}$. Here, this property fails to be true. Suppose for instance that $\Omega \subset \mathbb{R}^2$ and that its boundary is given near $(0,0)$ by $x_1 \ge 0$. Suppose that $e_1 = (a, b)$ and $e_2 = (b, -a)$ for some $a, b > 0$. Then

$$\begin{aligned} h^{-2}(1 - M_{1,h})f(x) &= \frac{1}{2h^3}\int_{|t|<h, x+te_1 \in \Omega} (f(x) - f(x + te_1))\, dt \\ &= -\frac{1}{2h^3}\partial_{e_1} f(x) \int_{|t|<h, x+te_1 \in \Omega} t\, dt + O(1) \\ &= \frac{1}{4h^3}\partial_{e_1} f(x) 1_{]0,ah]}(x_1)\left(h^2 - \frac{x_1^2}{a^2}\right) + O(1). \end{aligned} \tag{2.7}$$

A similar expression holds for $M_{2,h}$ and summing these equalities gives

$$h^{-2}(1-M_h)f(x)=\frac{1}{4h^3}\left(\partial_{e_1}f(0,x_2)1_{]0,ah]}(x_1)\left(h^2-\frac{x_1^2}{a^2}\right)\right.$$
$$\left.+\,\partial_{e_2}f(0,x_2)1_{]0,bh]}(x_1)\left(h^2-\frac{x_1^2}{b^2}\right)\right)+O(1) \quad (2.8)$$

If $a=b$, $\partial_{e_1}f+\partial_{e_2}f$ is proportional to the normal derivative of f and hence, the above quantity is bounded.

Suppose now that $a<b$. Then the above quantity is bounded on $x_1\in[ah,bh]$ provided $\partial_{e_2}f(0,x_2)=0$. Then the same argument on $[0,ah]$ shows that $\partial_{e_1}f(0,x_2)=0$ also.

In order to avoid these difficulties, we work directly on the quadratic form and show that the Dirichlet form associated to the Metropolis operator converges to the Dirichlet form of the Laplace operator with Neuman boundary conditions. The end of this section is devoted to the proof of Theorem 2.1.

Proposition 2.2. *There exists $N\in\mathbb{N}$ and constants $c_1,c_2>0$ such that for all $h\in]0,1]$*

$$M_h^N(x,dy)=\mu_h(x,dy)+c_1h^{-d}1_{|x-y|<c_2h}\,dy \quad (2.9)$$

where for all $x\in\Omega$, $\mu_h(x,dy)$ is a positive Borel measure.

Proof. The proof follows the lines of [4]. Denote $K_h=\frac{1}{p}\sum_{j=1}^p K_{j,h}$. Since for any $h_2>h_1>0$ and any non-negative function f, $h_2K_{h_2}f\geq h_1K_{h_1}f$, it is sufficient to prove the following: there exists $h_0>0$, $c_1,c_2>0$ and $N\in\mathbb{N}^*$ such that for all $h\in]0,h_0]$, one has, for all non-negative continuous functions f,

$$K_h^N(f)(x)\geq c_1h^{-d}\int_{y\in\Omega,|x-y|\leq c_2h}f(y)\,dy. \quad (2.10)$$

First note that it is sufficient to prove the weaker version: for all $x^0\in\overline{\Omega}$, there exist $N(x^0),\alpha=\alpha(x^0)>0$, $c_1=c_1(x_0)>0$, $c_2=c_2(x_0)>0$, $h_0=h_0(x_0)>0$ such that for all $h\in]0,h_0]$, all $x\in\Omega$ and all non-negative functions f

$$|x-x^0|\leq 2\alpha\Longrightarrow K_h^{N(x_0)}(f)(x)\geq c_1h^{-d}\int_{y\in\Omega,|x-y|\leq c_2h}f(y)\,dy. \quad (2.11)$$

Let us verify that (2.11) implies (2.10). Decreasing $\alpha(x_0)$ if necessary, it may be assumed that $2\alpha(x_0)<r(x_0)$, where $r(x_0)$ is given by Lemma 1.3. Since $\overline{\Omega}$ is compact, there exists a finite set F such that $\overline{\Omega}\subset\cup_{x_0\in F}\{|x-x_0|<\alpha(x_0)\}$. Let $N=\sup\{N(x_0),\,x_0\in F\}$, $c_i'=\min_{x_0\in F}c_i(x_0)$ and $h_0'=\min_{x_0\in F}h_0(x_0)$. One has to check that for any $x_0\in F$ and any x with $|x-x_0|\leq\alpha(x_0)$, the right inequality in (2.11) holds true with $N=N(x_0)+n$ in place of $N(x_0)$ for some constants c_1,c_2,h_0. Moreover, one may assume that $h_0\max|e_j|\leq\min_{x_0\in F}\alpha(x_0)/N$.

Let $\epsilon>0$, $q\in\{1,\dots,p\}$ and $\theta_q=\pm1$ be given by Corollary 1.4. Then for $|x-x_0|<(2-\frac{1}{N})\alpha(x_0)$, one has

$$\begin{aligned}K_h^{N(x_0)+1}f(x)&\geq\frac{1}{p}K_{h,\beta_q}K_h^{N(x_0)}f(x)\geq\frac{1}{p}\int_0^1K_h^{N(x_0)}f(x+ht\theta_qe_{\beta_q})\,dt\\&\geq c_1'\frac{h^{-d}}{p}\int_0^{\min(\epsilon,\frac{c_2}{2\max_j|e_j|})}\int_{y\in\Omega,\,|y-x-ht\theta_qe_{\beta_q}|<c_2h}f(y)\,dydt \quad (2.12)\\&\geq c_0'h^{-d}\int_{y\in\Omega,\,|y-x|<c_2h/2}f(y)\,dy\end{aligned}$$

since for any $t \in [0, \min(\epsilon, \frac{c_2}{2\max_j |e_j|})]$, $\{|y-x| < c_2h/2\} \subset \{|y - x - ht\theta_q e_{\beta_q}| < c_2 h\}$. Iterating this computation $n \leq N$ times gives (2.10).

It remains to prove (2.11). If $x_0 \in \Omega$, the proof is obvious. Indeed, since $\mathfrak{E}$ spans $\mathbb{R}^d$, it is easy to see that for any $\delta > 0$, there exists $c_3, c_4 > 0$ such that for any non-negative function f,

$$\mathrm{dist}(y, \partial\Omega) \geq \delta h \Longrightarrow K_h^d(f)(y) \geq c_3 h^{-d} \int_{z\in\Omega,\, |y-z|<c_4 h} f(z)\, dz \qquad \forall y \in \Omega. \tag{2.13}$$

Suppose that $x_0 \in \partial\Omega$ and denote $k = \mathfrak{c}(x_0)$. Let $(i_j)_{1\leq j\leq k}$, $(\beta_j)_{1\leq j\leq k}$, $(\theta_j)_{1\leq j\leq k}$ be as in Lemma 1.3. Let $1 = \gamma_1 > \gamma_2 > \cdots > \gamma_k > 0$ and $\delta_1, \ldots, \delta_k > 0$ be such that for all j, $\gamma_j - \delta_j > \gamma_{j+1}$. Let $G_j = [\gamma_j - \delta_j, \gamma_j]$ and $G = \Pi_{j=1}^k G_j$. In the following computation, c denotes a positive constant independant of f and h that may change from line to line. Since f is non-negative,

$$K_h^k(f)(x) \geq p^{-k} K_{\beta_1,h} \ldots K_{\beta_k,h} f(x) \geq c \int_{t\in A_h(x)} f(x + h\sum_{j=1}^k \theta_j t_j e_{\beta_j})\, dt \tag{2.14}$$

where $A_h(x) = \{t = (t_1, \ldots, t_k) \in G,\ \forall l = 1, \ldots, k,\ x + h\sum_{j=1}^l \theta_j t_j e_{\beta_j} \in \Omega\}$.

Since $\theta_1 e_{\beta_1}$ is strictly incoming to H_{i_1}, there exists some constant $c_5, c_6 > 0$ such that for any $t \in I$,

$$\begin{aligned} \mathrm{dist}\left(x + h\sum_{j=1}^k \theta_j t_j e_{\beta_j}, H_{i_1}\right) &\geq c_5 h t_1 - c_6 h (t_2 + \cdots + t_k) \\ &\geq c_5 h(\gamma_1 - \delta_1) - c_6 h(\gamma_2 + \cdots + \gamma_k) \\ &\geq c_5 h(\gamma_1 - \delta_1)/2 \end{aligned} \tag{2.15}$$

by taking $\gamma_2, \ldots, \gamma_k$ small with respect to γ_1. Similarly, by taking γ_j very small with respect to γ_{j+1} for $j = 2, \ldots, k$, there is $c_7 > 0$ such that for any $j = 1, \ldots, k$,

$$\forall (t_1, \ldots, t_j) \in G_1 \times \cdots \times G_j,\ \mathrm{dist}\left(x + h\sum_{i=1}^j \theta_i t_i e_{\beta_i}, \mathbb{R}^d \setminus \overline{\Omega}\right) \geq c_7 h. \tag{2.16}$$

Hence,

$$K_h^k f(x) \geq c \int_{t\in G} f\left(x + h\sum_{j=1}^k \theta_j t_j e_{\beta_j}\right)\, dt \tag{2.17}$$

and for any $N \geq 0$

$$K_h^{k+N} f(x) \geq c \int_{t\in G} K_h^N(f)\left(x + h\sum_{j=1}^k \theta_j t_j e_{\beta_j}\right)\, dt. \tag{2.18}$$

Combining (2.13), (2.16) and (2.18), there is $c_8 > 0$ small enough such that any $y \in \mathbb{R}^d$ such that $|x + h\sum_{j=1}^k t_j e_{\beta_j} - y| < c_8 h$ belongs to Ω and hence

$$K_h^{d+k} f(x) \geq c h^{-d} \int_{t\in G} \int_{|x+h\sum_{j=1}^k t_j e_{\beta_j} - y| < c_8 h} f(y)\, dy dt. \tag{2.19}$$

Since, $K_h^k f(y) \geq p^{-k} K_{h,\beta_k} \dots K_{h,\beta_1} f(y)$, then

$$K_h^{d+2k} f(x) \geq ch^{-d} \int_{(t,s,y)\in B_h(x)} f(y - h\sum_{j=1}^{k} s_j e_{\beta_j})\, dtdsdy \tag{2.20}$$

where

$$B_h(x) = \Bigg\{ (t,s,y) \in G \times G \times \mathbb{R}^d,\ |x + h\sum_{j=1}^{k} \theta_j t_j e_{\beta_j} - y| < c_8 h \text{ and}$$
$$\forall l = 1, \dots, k,\ y - h\sum_{j=l}^{k} \theta_j s_j e_{\beta_j} \in \Omega \Bigg\}. \tag{2.21}$$

Using the new variable $z = y - h\sum_{j=1}^{k} \theta_j s_j e_{\beta_j}$,

$$K_h^{d+2k} f(x) \geq ch^{-d} \int_{(t,s,z)\in D_h(x)} f(z)\, dtdsdz \tag{2.22}$$

with

$$D_h(x) = \Bigg\{ (t,s,z) \in G \times G \times \Omega,\ |x + h\sum_{j=1}^{k} (t_j - s_j)\theta_j e_{\beta_j} - z| < c_8 h \text{ and}$$
$$\forall l = 1, \dots, k,\ z + h\sum_{j=1}^{l-1} \theta_j s_j e_{\beta_j} \in \Omega \Bigg\}. \tag{2.23}$$

Since in the above integral, $|t_j - s_j| < \delta_j$, taking the δ_j's small enough gives

$$D_h(x) \supset \Bigg\{ (t,s,z) \in G \times G \times \Omega, |x - z| < c_8 h/2,\ \forall l = 1, \dots, k,\ z + h\sum_{j=1}^{l-1} \theta_j s_j e_{\beta_j} \in \Omega \Bigg\}. \tag{2.24}$$

Now using (2.16), it follows that

$$D_h(x) \supset \{(t,s,z) \in G \times G \times \Omega, |x - z| < c_8 h/2\}. \tag{2.25}$$

Combined with (2.22), this yields the announced result. □

Following the strategy of [4], introduce the Dirichlet form associated to the iterated kernel M_h^k:

$$\mathcal{E}_{h,k}(u) = \left\langle \left(1 - M_h^k\right) u, u \right\rangle_{L^2(\Omega)}. \tag{2.26}$$

Also, put Ω in a large box $B =]-A/2, A/2[^d$ and define an extension map $E : L^2(\Omega) \to L^2(B)$ which is continuous from $H^1(\Omega)$ into $H^1(B)$ and vanishes far from $\overline{\Omega}$. This is possible since $\partial\Omega$ has Lipschitz regularity. Finally, introduce the Dirichlet form on B:

$$\tilde{\mathcal{E}}_h(u) = h^{-d} \int_{B\times B, |x-y|<h} |u(x) - u(y)|^2\, dxdy. \tag{2.27}$$

Then Proposition 2.2 easily yields the following (see [4] for details).

Lemma 2.3. *There exists $C_0, h_0 > 0$ such that for any $h \in]0, h_0]$ and any $u \in L^2(\Omega)$,*

$$\tilde{\mathcal{E}}_h\left(E(u)\right) \leq C_0 \left(\mathcal{E}_{h,N}(u) + h^2 \|u\|^2_{L^2(\Omega)}\right). \tag{2.28}$$

Moreover, any function $u \in L^2(\Omega)$ such that

$$\|u\|^2_{L^2(\Omega)} + h^{-2} \left\langle (1 - M_h)u, u \right\rangle_{L^2(\Omega)} \leq 1$$

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

We are now in position to prove the first part of Theorem 2.1. First, assume that $M_h u = u$. Then, it follows from Proposition 2.2, that

$$c_1 h^{-d} \int_{\Omega\times\Omega, |x-y|<c_2 h} (u(x) - u(y))^2 \, dx dy \leq \int_{\Omega\times\Omega} (u(x) - u(y))^2 \, M_h^N(x, dy)\, dx. \tag{2.29}$$

On the other hand, the right hand side in the above inequality is equal to $\mathcal{E}_{h,N}(u)$ which is actually equal to zero. Hence, u is constant and 1 is a simple eigenvalue.

Using the Markov property of M_h^N, positivity of μ_h and the fact that $\partial\Omega$ has Lipschitz regularity, easily yields

$$\|\mu_h\|_{L^\infty \to L^\infty} = \mu_h(\Omega) \leq 1 - c_1 h^{-d} \min_{x\in\overline{\Omega}} \int_\Omega 1_{|x-y|<c_2 h}\, dy < 1 - \delta_0' \tag{2.30}$$

for some $\delta_0' > 0$ independent of h. Working as in the proof of Theorem 1 in [4] shows that there exists $\delta_0 \in]0, \frac{1}{2}[$ such that for any $u \in L^2(\Omega)$ and any $n \geq N$,

$$\langle M_h^n u, u \rangle_{L^2(\Omega)} \geq (-1 + \delta_0)\|u\|^2_{L^2(\Omega)}. \tag{2.31}$$

Hence, the same holds true for $n = 1$ with a possibly different δ_0.

To show that there is $\delta_0 > 0$ sufficiently small so that the spectrum of M_h is discrete in $[1-\delta_0, 1]$ it suffices to work as in the proof of Theorem 4.6 in [4], using again Proposition 2.2.

Similarly, the Weyl bound on the number of eigenvalues follows from Lemma 2.3 as in Lemma 4.8 in [4]. This proves Part *i*.

To prove the last part of the theorem, work on the Dirichlet form is needed. In the following, denote $\mathcal{E}_h = \mathcal{E}_{h,1}$. Introduce the bilinear form associated with $\mathcal{E}_h$:

$$\mathcal{B}_h(u, v) = \langle (1 - M_h)u, v \rangle_{L^2(\Omega)}, \ \forall u, v \in L^2(\Omega). \tag{2.32}$$

A standard computation shows that $\mathcal{B}_h(u, v) = \frac{1}{p}\sum_{j=1}^p \mathcal{B}_{j,h}(u, v)$ with

$$\mathcal{B}_{j,h}(u, v) = \frac{1}{4h} \int_{x\in\Omega, x+te_j\in\Omega, |t|<h} (u(x) - u(x + te_j)) \left(\overline{v}(x) - \overline{v}(x + te_j)\right) dx dt \tag{2.33}$$

Lemma 2.4. *Let $\theta \in C^\infty(\overline{\Omega})$ be fixed and let $(\varphi_h, r_h) \in H^1(\Omega) \times L^2(\Omega)$ be such that $\|r_h\|_{L^2(\Omega)} = O(h)$ and φ_h converges weakly in $H^1(\Omega)$ to some φ. Then*

$$\lim_{h\to 0^+} h^{-2}\mathcal{B}_h(r_h, \theta) = 0 \tag{2.34}$$

and

$$\lim_{h\to 0^+} h^{-2}\mathcal{B}_h(\varphi_h, \theta) = \frac{1}{6p} \int_\Omega \left\langle \nabla_{\mathfrak{E}}\varphi(x), \overline{\nabla_{\mathfrak{E}}\theta}(x) \right\rangle_{\mathbb{C}^p} dx. \tag{2.35}$$

Proof. To prove (2.34), observe that since θ is smooth,

$$\begin{aligned}(1-M_{j,h})\theta(x) &= \frac{h^{-1}}{2}\int_{|t|<h,x+te_j\in\Omega}(\theta(x)-\theta(x+te_j))\,dt\\ &= \frac{\partial_{e_j}\theta(x)}{2h}\int_{|t|<h,x+te_j\in\Omega}t\,dt+O(h^2).\end{aligned}\tag{2.36}$$

Denoting

$$\rho_h(x)=\frac{\partial_{e_j}\theta(x)}{2h}\int_{|t|<h,x+te_j\in\Omega}t\,dt$$

observe that $\mathrm{supp}(\rho_h)\subset\{x\in\Omega,\,d(x,\partial\Omega)<h\}$ and $\|\rho_h\|_{L^\infty}=O(h)$. Hence $\|\rho_h\|_{L^2}=O(h^{3/2})$ and since $\|r_h\|_{L^2}=O(h)$, it follows that

$$h^{-2}\mathcal{B}_{j,h}(r_h,\theta)=h^{-2}\langle r_h,(1-M_{j,h})\theta\rangle_{L^2}=\langle h^{-1}r_h,h^{-1}\rho_h\rangle_{L^2}+O(h)=O(h^{1/2})\tag{2.37}$$

which goes to zero as h goes to zero.

To prove (2.35) observe that

$$\begin{aligned}\theta(x+te_j)-\theta(x)&=t\psi(t,x)\\ \varphi_h(x+te_j)-\varphi_h(x)&=t\int_0^1\partial_{e_j}\varphi_h(x+tze_j)\,dz\end{aligned}\tag{2.38}$$

with $\psi(t,x)$ smooth and $\psi(0,x)=\partial_{e_j}\theta(x)$. Hence

$$\begin{aligned}h^{-2}\mathcal{B}_{j,h}(\varphi_h,\theta)&=\frac{1}{4h^3}\int_{x\in\Omega,x+te_j\in\Omega,|t|<h,z\in[0,1]}t^2\partial_{e_j}\varphi_h(x+tze_j)\overline{\psi}(t,x)\,dtdzdx\\ &=\frac14\int_{x\in\Omega,x+hue_j\in\Omega,|u|<1,z\in[0,1]}u^2\partial_{e_j}\varphi_h(x+huze_j)\overline{\psi}(hu,x)\,dudzdx\\ &=\frac14\int_{x-huze_j\in\Omega,x+hu(1-z)e_j\in\Omega,|u|<1,z\in[0,1]}u^2\partial_{e_j}\varphi_h(x)\overline{\psi}(hu,x-huze_j)\,dudzdx.\end{aligned}\tag{2.39}$$

Taylor expansion of ψ shows that $\psi(hu,x-huze_j)=\partial_{e_j}\theta(x)+O(h)$. Hence, for any $\delta>0$ and any $h\in]0,1]$,

$$\begin{aligned}h^{-2}\mathcal{B}_{j,h}(\varphi_h,\theta)&=\frac14\int_{x-huze_j\in\Omega,x+hu(1-z)e_j\in\Omega,|u|<1,z\in[0,1]}u^2\partial_{e_j}\varphi_h(x)\overline{\partial_{e_j}\theta}(x)\,dudzdx+O(h)\\ &=I_\delta(h)+J_\delta(h)+O(h)\end{aligned}\tag{2.40}$$

with $I_\delta(h)$ equal to the above integral over $d(x,\partial\Omega)\geq\delta$ and $J_\delta(h)$ the integral over $d(x,\partial\Omega)<\delta$. Then, by Cauchy–Schwartz, $|J_\delta(h)|\leq C(\theta)\delta^{1/2}\|\varphi_h\|_{H^1}$. On the other hand, for any $h\in]0,\delta[$,

$$\begin{aligned}I_\delta(h)&=\frac16\int_{x\in\Omega,d(x,\partial\Omega)>\delta}\partial_{e_j}\varphi_h(x)\overline{\partial_{e_j}\theta}(x)\,dx\\ &=\frac16\int_{x\in\Omega}\partial_{e_j}\varphi_h(x)\overline{\partial_{e_j}\theta}(x)\,dx+O\left(\delta^{1/2}\|\varphi_h\|_{H^1}\right).\end{aligned}\tag{2.41}$$

Given $\epsilon > 0$, it is easy to find $\delta > 0$ small enough such that for any $h \in]0, \delta[$, $|J_\delta(h)| < \epsilon$ and $|I_\delta(h) - \frac{1}{6}\int_{x\in\Omega} \partial_{e_j}\varphi_h(x)\partial_{e_j}\theta(x)dx| < \epsilon$. Now make $h \to 0^+$, δ being fixed, and use the fact that φ_h converges weakly in H^1 to get

$$\lim_{h\to 0^+} h^{-2}\mathcal{B}_{j,h}(\varphi_h, \theta) = \frac{1}{6}\int_\Omega \partial_{e_j}\varphi(x)\overline{\partial_{e_j}\theta}(x)\,dx \tag{2.42}$$

and the proof is complete. □

To complete the proof of Theorem 2.1, denote $|\Delta_h| = h^{-2}(1 - M_h)$. Let $R > 0$ be fixed and observe that if $\nu_h \in [0, R]$ and $f_h \in L^2(\Omega)$ satisfy $|\Delta_h| f_h = \nu_h f_h$ and $\|f_h\|_{L^2} = 1$, then, thanks to Lemma 2.3, f_h can be decomposed as $f_h = \varphi_h + r_h$ with $\|r_h\|_{L^2(\Omega)} = O(h)$ and φ_h bounded in H^1. Hence (extracting a subsequence if necessary) it may be assumed that φ_h weakly converges in H^1 to a limit φ and that ν_h converges to a limit ν. It now follows from Lemma 2.4 that for any $\theta \in C^\infty(\overline{\Omega})$,

$$\frac{1}{6p}\int_\Omega \langle \nabla_{\mathfrak{E}} f(x), \nabla_{\mathfrak{E}}\theta(x)\rangle_{\mathbb{C}^p}\; dx = \nu\langle\varphi, \theta\rangle_{L^2}. \tag{2.43}$$

Since θ is arbitrary, it follows that $(-\Delta_{\mathfrak{E}} - \nu)\varphi = 0$ and $\partial_{n,\mathfrak{E}}\varphi_{|\partial\Omega} = 0$. In fact, this also proves that for any $\epsilon > 0$ small, there exists $h_\epsilon > 0$ such that for $h \in]0, h_\epsilon]$, one has

$$\mathrm{Spec}(|\Delta_h|) \cap [0, R] \subset \cup_j [\nu_j - \epsilon, \nu_j + \epsilon] \tag{2.44}$$

and

$$\sharp\mathrm{Spec}(|\Delta_h|) \cap [\nu_j - \epsilon, \nu_j + \epsilon] \le m_j \tag{2.45}$$

In fact, there is equality in (2.45). The following proof is a simplification of the one in [4]. Proceed by induction on j: let $\epsilon > 0$, small, be given such that for $0 \le \nu_j \le M+1$, the intervals $I_j^\epsilon = [\nu_j - \epsilon, \nu_j + \epsilon]$ are disjoint. Let $(\mu_j^h)_{j\ge 0}$ be the increasing sequence of eigenvalues of $|\Delta_h|$, $\sigma_N = \sum_{j=1}^N m_j$ and $(e_k)_{k\ge 0}$ an othonormal basis of eigenfunctions of $-\Delta_{\mathfrak{E}}$ such that for all $k \in \{1+\sigma_N, \dots, \sigma_{N+1}\}$, one has $(-\Delta_{\mathfrak{E}} - \nu_{N+1})e_k = 0$. As 0 is a simple eigenvalue of both $-\Delta_{\mathfrak{E}}$ and $|\Delta_h|$, clearly $\nu_0 = \mu_0 = 0$ and $m_0 = 1 = \sharp\mathrm{Spec}(|\Delta_h|) \cap [\nu_0 - \epsilon, \nu_0 + \epsilon]$.

Suppose that for all $n \le N$, $m_n = \sharp\mathrm{Spec}(|\Delta_h|) \cap [\nu_n - \epsilon, \nu_n + \epsilon]$. Then by (2.44), for $h \le h_\varepsilon$,

$$\mu^h_{1+\sigma_N} \ge \nu_{N+1} - \epsilon. \tag{2.46}$$

By the min-max principle, if G is a finite dimensional subspace of H^1 with $\dim(G) = 1 + \sigma_{N+1}$,

$$\mu^h_{\sigma_{N+1}} \le \sup_{\psi\in G, \|\psi\|=1} \langle |\Delta_h|\psi, \psi\rangle_{L^2(\Omega)}. \tag{2.47}$$

Let G be the vector space spanned by the e_k, $0 \le k \le \sigma_{N+1}$. Then, $\dim(G) = 1 + \sigma_{N+1}$ and it follows from Lemma 2.4, for any $k, k' \le 1 + \sigma_{N+1}$,

$$\lim_{h\to 0^+} h^{-2}\mathcal{B}_h(e_k, e_{k'}) = \frac{1}{6p}\int_\Omega \langle \nabla_{\mathfrak{E}} e_k(x), \nabla_{\mathfrak{E}} e_{k'}(x)\rangle_{\mathbb{C}^p}\; dx. \tag{2.48}$$

Hence

$$\lim_{h\to 0^+} h^{-2}\mathcal{B}_h(\psi, \psi) = \frac{1}{6p}\int_\Omega |\nabla_{\mathfrak{E}}\psi(x|^2\,dx \le \nu_{N+1} \tag{2.49}$$

for any $\psi \in G$ with $\|\psi\|_{L^2} = 1$. Since G has finite dimension, a standard compactness argument shows that there exists $h_\epsilon > 0$ such that for any $h \in]0, h_\epsilon]$ and any $\psi \in G$ with $\|\psi\|_{L^2} \le 1$,

$$h^{-2}\mathcal{B}_h(\psi, \psi) \le \nu_{N+1} + \epsilon. \tag{2.50}$$

Therefore $\mu_{\sigma_{N+1}} \le \nu_{N+1} + \epsilon$. Combining this with (2.46) and (2.45) gives $m_{N+1} = \sharp\mathrm{Spec}(|\Delta_h|) \cap [\nu_{N+1} - \epsilon, \nu_{N+1} + \epsilon]$. The proof of Theorem 2.1 is complete.

3 Total Variation Estimates

This section gives estimates on the convergence speed of the iterated kernel $M_h^n(x,dy)$ towards its stationary measure $\frac{dy}{Vol(\Omega)}$. Recall that the total variation $\|\mu-\nu\|_{TV}$ between two probability measures μ and ν on Ω is defined by

$$\|\mu-\nu\|_{TV} = \sup|\mu(A)-\nu(A)| \tag{3.1}$$

where the sup is taken over all measurable sets A. Equivalently,

$$\|\mu-\nu\|_{TV} = \frac{1}{2}\sup_{f\in L^\infty,\|f\|_{L^\infty}=1}|\mu(f)-\nu(f)|. \tag{3.2}$$

Theorem 3.1. *Assume that $\mathfrak{E}$ is weakly incoming. Then there exists $C>0$ and $h_0>0$ such that for all $h\in]0,h_0]$ and all $n\in\mathbb{N}$, the following estimate holds true, with $g(h)$ the spectral gap studied in Section 2:*

$$\sup_{x\in\Omega}\left\|M_h^n(x,dy)-\frac{dy}{\mathrm{vol}(\Omega)}\right\|_{TV}\leq Ce^{-ng(h)}. \tag{3.3}$$

Proof. The proof is very close to the proof of Theorem 4.6 in [4] and is just sketched for the reader's convenience. Observe first that $n\geq h^{-2}$ can be assumed, since otherwise the estimate is trivial thanks to the lower bound on the spectral gap.

Let Π_0 be the othogonal projector in $L^2(\Omega)$ on the constant functions. Observe that

$$2\sup_{x\in\Omega}\left\|M_h^n(x,dy)-\frac{dy}{\mathrm{vol}(\Omega)}\right\|_{TV} = \|M_h^n-\Pi_0\|_{L^\infty\to L^\infty}. \tag{3.4}$$

Using the spectral decomposition of M_h, let $0<\lambda_{1,h}\leq\cdots\leq\lambda_{j,h}\leq\cdots\leq h^{-2}\delta_0$ be such that the eigenvalues of M_h in the interval $[1-\delta_0,1]$ are the $1-h^2\lambda_{j,h}$ with associated orthonormalized eigenfunctions $M_h(e_{j,h})=(1-h^2\lambda_{j,h})e_{j,h}$.

Then write $M_h-\Pi_0=M_{h,1}+M_{h,2}+M_{h,3}$, so that the operators $M_{h,1}$, $M_{h,2}$ have kernels

$$M_{h,1}(x,y) = \sum_{\lambda_{1,h}\leq\lambda_{j,h}\leq h^{-\alpha}}(1-h^2\lambda_{j,h})e_{j,h}(x)e_{j,h}(y) \tag{3.5}$$

$$M_{h,2}(x,y) = \sum_{h^{-\alpha}\leq\lambda_{j,h}\leq h^{-2}\delta_0}(1-h^2\lambda_{j,h})e_{j,h}(x)e_{j,h}(y) \tag{3.6}$$

where $\alpha\in]0,2]$ is a small constant that will be chosen later. Then

$$2\sup_{x\in\Omega}\left\|M_h^n(x,dy)-\frac{dy}{\mathrm{vol}(\Omega)}\right\|_{TV}\leq\sum_{j=1}^{3}\|M_{h,j}^n\|_{L^\infty\to L^\infty} \tag{3.7}$$

and terms on the right hand side must be estimated.

From (2.30), it is easy to prove that any eigenfunction $M_h(u)=\lambda u$ with $\lambda\in]1-\delta_0,1]$ satisfies

$$\|u\|_{L^\infty}\leq Ch^{-d/2}\|u\|_{L^2}. \tag{3.8}$$

As in [4], using in particular the bound on the number of eigenvalues, we show that for $n\in\mathbb{N}$,

$$\|M_{h,2}^n\|_{L^\infty\to L^\infty}+\|M_{h,3}^n\|_{L^\infty\to L^\infty}\leq C\big((1-h^{2-\alpha})^n+(1-\delta_0)^n\big)h^{-3d/2} \tag{3.9}$$

For $n \geq h^{-2}$, this implies that

$$\|M_{h,2}^n\|_{L^\infty \to L^\infty} + \|M_{h,3}^n\|_{L^\infty \to L^\infty} \leq C_\alpha e^{-nh^{2-\alpha}}. \tag{3.10}$$

It remains to estimate $M_{h,1}^n$. Let E_α denote the space spanned by the eigenvectors $e_{j,h}$ such that $\lambda_{j,h} \leq h^{-\alpha}$. Then, thanks to Part *ii* of Theorem 2.1, $\dim(E_\alpha) \leq h^{-d\alpha/2}$. As in [4], Lemma 2.3 shows that there exists $\alpha > 0$ and $p > 2$ such that for any $u \in E_\alpha$,

$$\|u\|_{L^p}^2 \leq Ch^{-2}\left(\mathcal{E}_{h,N}(u) + h^2\|u\|_{L^2}^2\right). \tag{3.11}$$

This gives the following Nash estimate, with $\frac{1}{D} = 2 - \frac{4}{p} > 0$:

$$\|u\|_{L^2}^{2+\frac{1}{D}} \leq Ch^{-2}\left(\mathcal{E}_{h,N}(u) + h^2\|u\|_{L^2}^2\right)\|u\|_{L^1}^{\frac{1}{D}} \qquad \forall u \in E_\alpha. \tag{3.12}$$

This inequality allows an estimate of $M_{h,1}$ from L^1 into L^2 and this leads to $\|M_{h,1}^{kN}\|_{L^\infty \to L^\infty} \leq Ce^{-kNg(h)}$ for $k \geq h^{-2}$. As M_h is bounded by 1 on L^∞ it follows that kN can be replaced by $n \geq h^{-2}$ in this estimate, and the proof of Theorem 3.1 is complete. □

4 Convergence of the Resolvants

Let us denote $|\Delta_h| = h^{-2}(1 - M_h)$. Recall $\Delta_{\mathfrak{E}}$ from (2.1). This section proves strong resolvent convergence of $|\Delta_h|$ to $\Delta_{\mathfrak{E}}$. For background and consequences, see [13].

Theorem 4.1. *Let $z \in \mathbb{C} \setminus [0, +\infty[$ and $g \in L^2(\Omega)$. Then*

$$\lim_{h \to 0^+} \left\|(|\Delta_h| - z)^{-1}g - (-\Delta_{\mathfrak{E}} - z)^{-1}g\right\|_{L^2(\Omega)} = 0. \tag{4.1}$$

Proof. Let $z \in \mathbb{C} \setminus [0, +\infty[$ and $g \in L^2(\Omega)$ be fixed. For any $h > 0$ let $f_h \in L^2(\Omega)$ be the solution of $(|\Delta_h| - z)f_h = g$. Hence

$$-z\langle f_h, f_h\rangle_{L^2} + \left\langle \frac{1 - M_h}{h^2} f_h, f_h \right\rangle_{L^2} = \langle g, f_h\rangle_{L^2}. \tag{4.2}$$

Since $z \notin [0, \infty[$ and $|\Delta_h|$ is a positive operator, it follows that $\|f_h\|_{L^2} \leq dist(z, [0, \infty[)^{-1}\|g\|_{L^2}$ is bounded uniformly with respect to h. It follows from the above equation that there exists $C_0 > 0$ such that

$$\|f_h\|_{L^2}^2 + h^{-2}\mathcal{E}_h(f_h) \leq C_0\|g\|_{L^2}^2. \tag{4.3}$$

It now follows from Lemma 2.4 that there exists $C > 0$ depending on z and $\|g\|_{L^2}$ such that for any $h \in]0, 1]$, we can write $f_h = \varphi_h + r_h$ with $\|\varphi_h\|_{H^1} \leq C$ and $\|r_h\|_{L^2} \leq Ch$. Let $f \in H^1(\Omega)$ and $(h_k)_{k \in \mathbb{N}}$ be a sequence of positive numbers such that $(\varphi_{h_k})_k$ converges weakly to f in H^1. Let $\theta \in C^\infty(\overline{\Omega})$ be fixed. Then

$$-z\langle f_{h_k}, \theta\rangle_{L^2} + h_k^{-2}\mathcal{B}_{h_k}(f_{h_k}, \theta) = \langle g, \theta\rangle_{L^2} \tag{4.4}$$

and taking the limit $k \to \infty$ it follows from Lemma 2.4 that

$$-z\langle f, \theta\rangle_{L^2} + \frac{1}{6p}\int_\Omega \nabla_{\mathfrak{E}} f(x)\overline{\nabla_{\mathfrak{E}}\theta(x)}\, dx = \int_\Omega g(x)\overline{\theta(x)}\, dx. \tag{4.5}$$

Since θ is arbitrary, this implies $(-\Delta_{\mathfrak{E}} - z)f = g$ and $\partial_{n,\mathfrak{E}} f_{|\partial\Omega} = 0$. Since, this is true for any subsequence (h_k), this shows that $\|f_h - f\|_{L^2} \to 0$ when $h \to 0$, which is exactly (4.1). □

5 Some Generalizations

Here we present a possible generalization of the previous results. It is still assumed that Ω is a convex polytope in $\mathbb{R}^d$. Suppose that $E \subset \mathbb{R}^d$ is endowed with a Borel probability measure μ. For any $e \in E$, define

$$K_{e,h}f(x) = \frac{1}{2}\int_{t\in[-1,1],x+hte\in\Omega} f(x+hte)\,dt \tag{5.1}$$

and

$$K_hf(x) = \int_{e\in E} K_{e,h}f(x)\,d\mu(e). \tag{5.2}$$

The associated Metropolis operator is defined by $M_hf(x) = m_h(x)f(x) + K_hf(x)$ with $m_h(x) = 1 - K_h(1)$.

Definition 5.1. Say that (E,μ) is weakly incoming to Ω if for any $x_0 \in \partial\Omega$ there exists $\epsilon > 0$, $\theta \in \{\pm 1\}$ and a measurable subset $F \subset E$ such that $\mu(F) > 0$ and

$$\mathfrak{c}(x_0 + \theta te) < \mathfrak{c}(x_0) \qquad \forall t \in]0,\epsilon],\ \forall e \in F. \tag{5.3}$$

Lemma 5.1. *There exists some measurable subsets $F_1,\dots,F_d \subset E$ such that $\mu(F_j) > 0$ for all j and any $(f_1,\dots,f_d) \in \Pi_{j=1}^d F_j$ spans $\mathbb{R}^d$. Moreover the sets F_j can be chosen with arbitrary small diameters.*

Proof. From the same argument as in remark following Definition 1.1, we can easily see that μ can not be supported in an hyperplane of $\mathbb{R}^d$. Let us prove by induction that for $k = 1,\dots,d$, there exists $F_1,\dots F_k \subset E$ such that $\mu(F_j) > 0$ for all j and for any $(f_1,\dots,f_k) \in \Pi_{j=1}^k F_j$, $rank(f_1,\dots,f_k) = k$.

If $k = 1$, it suffices to take $F_1 \subset F \setminus \{0\}$ with $\mu(F_1) > 0$, which is possible thanks to the fact that F is weakly incoming to Ω.

Assume that the property holds true at rank $k-1 < d$. There exists $F_1,\dots F_{k-1} \subset E$ such that $\mu(F_j) > 0$ for all j and any $(f_1,\dots,f_{k-1}) \in \Pi_{j=1}^{k-1} F_j$, $H = span(f_1,\dots,f_{k-1})$ has dimension $k-1$. Since $\mathrm{supp}(\mu)$ is not contained in H, there exists $F_k \subset F \setminus H$ with $\mu(F_k) > 0$. Then $F_1,\dots,F_k$ satisfy the property at rank k.

The fact that we can take $diam(F_j)$ arbitrary small can be shown as follows. Let $\epsilon > 0$ and assume by contradiction that there exists j_0 such that for any $f \in F_{j_0}$, $\mu(B(f,\epsilon) \cap F_{j_0}) = 0$. Then any compact subset of F_{j_0} would have measure zero, which is impossible since $\mu(F_{j_0}) > 0$. $\square$

Introduce the following differential operators associated to the set E:

$$\nabla_E :\ H^1(\Omega) \to L^\infty\left(E, L^2(\Omega)\right) \tag{5.4}$$

defined by $\nabla_E u(e,x) = \langle \nabla u(x), e\rangle_{\mathbb{C}^d}$ for any $(e,x) \in E \times \Omega$;

$$\mathrm{div}_E :\ L^\infty\left(E, H^1(\Omega)\right) \to L^2(\Omega) \tag{5.5}$$

defined by $\mathrm{div}_E f(x) = \int_E \langle \nabla_x f(e,x), e\rangle_{\mathbb{C}^d} d\mu(e)$ for any $x \in \Omega$; and

$$\Delta_E :\ H^2(\Omega) \to L^2(\Omega) \tag{5.6}$$

given by $\Delta_E = \frac{1}{6}\mathrm{div}_E \nabla_E$.

Define also the following trace operator:

$$\gamma_E^0 : \left\{ f \in L^\infty\left(E, H^1(\Omega)\right),\, \mathrm{div}_E f \in L^2(\Omega) \right\} \to H^{-\frac{1}{2}}(\partial\Omega) \tag{5.7}$$

by

$$\int_{\partial\Omega} \gamma_E^0 f(x) v_{|\partial\Omega}(x)\, d\sigma(x) = \int_\Omega \mathrm{div}_E f(x) v(x)\, dx + \int_E \int_\Omega f(e,x) \nabla_E v(e,x)\, dx d\mu(e) \tag{5.8}$$

for any $v \in H^1(\Omega)$. Observe that if $f \in L^\infty(E, C^1(\overline{\Omega}))$, then

$$\gamma_E^0 f(x) = \int_E \langle e, n(x) \rangle_{\mathbb{C}^d} f(x,e)\, d\mu(e) \tag{5.9}$$

where $n(x)$ denotes the unit outgoing normal vector to the boundary $\partial\Omega$ at point x.

For $u \in H^1(\Omega)$ such that $\Delta_E u \in L^2(\Omega)$, the function $f = \nabla_E u$ satisfies $\mathrm{div}_E f \in L^2(\Omega)$. Hence, the operator

$$\gamma_E^1 : \left\{ u \in H^1(\Omega),\, \Delta_E u \in L^2(\Omega) \right\} \to H^{-\frac{1}{2}}(\partial\Omega) \tag{5.10}$$

defined by $\gamma_E^1 u(x) = \gamma_E^0 \nabla_E u(x)$ is continuous.

Finally, introduce the following quadratic form on $H^1(\Omega)$:

$$\mathcal{E}_E(u) = \frac{1}{6} \int_E \int_\Omega |\nabla_E u(e,x)|^2\, dx d\mu(e) \qquad \forall u \in H^1(\Omega). \tag{5.11}$$

From Lemma 5.1 it follows that, since E is weakly incoming to Ω, there exists some subsets $F_1, \dots, F_d$ with arbitrary small diameters and $\mu(F_j) > 0$ such that any $(f_1, \dots, f_d) \in F_1 \times \cdots \times F_d$ spans $\mathbb{R}^d$. Taking the diameter of the F_j sufficiently small, it is easy to show that there exists $C > 0$ such that for any $u \in H^1(\Omega)$,

$$\frac{1}{C} \|\nabla u\|^2_{L^2(\Omega)} \le \mathcal{E}_E(u) \le C \|\nabla u\|^2_{L^2(\Omega)}. \tag{5.12}$$

Then, the operator $-\Delta_E = -\frac{1}{6}\mathrm{div}_E \nabla_E$ with domain $D(-\Delta_E) = \{u \in H^1(\Omega),\ \Delta_E u \in L^2(\Omega),\, \gamma_E^1 u = 0\}$ is the self-adjoint realization of the Dirichlet form $\mathcal{E}_E$. Moreover, it follows from (5.12) that $-\Delta_E$ has compact resolvant. Denote its spectrum by $\nu_0 = 0 < \nu_1 < \nu_2 < \dots$ and by m_j the multiplicity associated to ν_j. Observe that $m_0 = 1$.

Theorem 5.2. *Suppose that (E, μ) is weakly incoming to Ω, then the following hold true.*

i) There exists $h_0 > 0,\ \delta_0 \in]0, \frac{1}{2}[$ and a positive constant C such that for any $h \in]0, h_0]$, the spectrum of M_h is a subset of $[-1+\delta_0, 1]$, 1 is a simple eigenvalue and $\mathrm{Spec}(M_h) \cap [1-\delta_0, 1]$ is discrete.

ii) For any $h \in]0, h_0]$ and $0 \le \lambda \le \delta_0 h^{-2}$, the number of eigenvalues of M_h in $[1 - h^2\lambda, 1]$ (with multiplicity) is bounded by $C(1+\lambda)^{d/2}$.

iii) 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 $h \in]0, h_1]$,

$$\mathrm{Spec}\left(\frac{1 - M_h}{h^2}\right) \cap]0, R] \subset \cup_{j \ge 1} [\nu_j - \varepsilon, \nu_j + \varepsilon] \tag{5.13}$$

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

Here are two examples of (E, μ) which are weakly incoming to Ω. The first is the case where $E = \{e_1, \dots, e_p\}$ is discrete and μ is simply the measure $\frac{1}{p}\sum_{j=1}^{p} \delta_{e=e_j}$. Then it suffices to assume that E is weakly incoming to Ω in the sense of Definition 1.1. Moreover, in that case the conclusion of Theorem 5.2 are exactly those of Theorem 2.1.

A second example is the following. Let E be equal to the sphere S^{d-1} and $\mu = d\sigma_d$ be the surface measure. Assume that $\rho : S^{d-1} \to \mathbb{R}^+$ is a continuous function such that $\int_{S^{d-1}} \rho(\omega) d\sigma_d(\omega) = 1$ and let $\mu = \rho(\omega) d\sigma_d(\omega)$. Then (E, μ) will be weakly incoming to Ω iff there exists a family of vectors $e_1, \dots, e_p \in \operatorname{supp}(\rho)$ such that $(e_1, \dots, e_p)$ is weakly incoming in the sense of Definition 1.1. For instance, if ρ is strictly positive on S^{d-1} then these assumptions are automatically satisfied.

The proof of Theorem 5.2 is very close to that of Theorem 2.1 and only the main steps are given. The following proposition is a version of Lemma 1.3 adapted to the present setting.

Proposition 5.3. *Assume that (E, μ) is weakly incoming to Ω, let $x_0 \in \overline{\Omega}$ and denote $k = \mathfrak{c}(x_0)$. There exists $\epsilon > 0$ and some subsets $F_1, \dots, F_k \subset E$ such that $\mu(F_i) > 0$ for all $i = 1, \dots, k$ and*

- *there exists $r_0 > 0$ and $I \subset \{1, \dots, m\}$ with $\sharp I = k$ such that*

$$\Omega \cap B(x_0, r_0) = \left(\cap_{i \in I}^{k} H_i^+ \right) \cap B(x_0, r_0); \tag{5.14}$$

- *there exists $\theta_1, \dots, \theta_k \in \{\pm 1\}$ and a bijection $\{1, \dots, k\} \ni n \mapsto i_n \in I$ such that for any $n = 1, \dots, k$ and any $f_n \in F_n$,*

$$\theta_n f_n \textit{ is strictly incoming to } H_{i_n} \tag{5.15}$$

and

$$\theta_n f_n \textit{ is incoming to } H_{i_m} \qquad \forall m > n. \tag{5.16}$$

Moreover the sets $F_1, \dots, F_k$ can be chosen with arbitrary small diameter.

Proof. First, it is clear that (5.14) holds true. We prove (5.15) and (5.16) by induction on $k = \mathfrak{c}(x_0)$. For $k = 0$ there is nothing to prove.

Assume now that the property holds true for all x_0' such that $\mathfrak{c}(x_0') \leq k - 1$ and suppose that $\mathfrak{c}(x_0) = k$. Since (E, μ) is weakly incoming to Ω, there exists $F \subset E$, $\theta_1 \in \{\pm 1\}$ and $\epsilon > 0$ such that $\mathfrak{c}(x_0 + t\theta_1 f) < \mathfrak{c}(x_0)$ for all $t \in]0, \epsilon]$. Assume without loss of generality that $\theta_1 = 1$. Since $\mu(F) > 0$, there exists $f^0 \in F$ such that for all $\rho > 0$, $\mu(B(f^0, \rho) \cap F) > 0$ and

$$\mathfrak{c}(x_0 + tf) < \mathfrak{c}(x_0) \qquad \forall f \in B(f^0, \rho) \cap F, \ \forall t \in]0, \epsilon]. \tag{5.17}$$

In particular, there exists $q_1 \in \{1, \dots, k\}$ and $i_1, \dots, i_{q_1} \in I$ such that

$$f^0 \text{ is strictly incoming to } H_{i_q} \qquad \forall q = 1, \dots, q_1 \tag{5.18}$$

and

$$f^0 \text{ is parallel to } H_{i_q} \qquad \forall q \geq q_1 + 1. \tag{5.19}$$

Let $F_q = B(f_0, \rho) \cap F$ with $\rho > 0$ for $q = 1, \dots, q_1$. Then $\mu(F_q) > 0$ and it follows from (5.18) that for ρ small enough, any $f \in F_q$ is strictly incoming to H_{i_q}. Moreover, thanks to (5.17), any $f \in F_q$ is incoming to H_i for $i \in I \setminus \{i_1, \dots, i_{q_1}\}$. Then we can use the induction hypothesis with $x_0' = x_0 + \epsilon f_0$ close to x_0 such that $\mathfrak{c}(x_0') = k - q_1 < k$ to build $F_{q_1+1}, \dots, F_k$. The statement concerning the diameter of the F_j is a trivial consequence of the construction. ☐

Corollary 5.4. *Assume that* (E,μ) *is weakly incoming to* Ω *and let* $x_0 \in \overline{\Omega}$*. Then there exists* $r_0 > 0$, $\epsilon > 0$, $F \subset E$ *with* $\mu(F) > 0$ *and* $\theta \in \{\pm 1\}$ *such that*

$$\forall f \in F,\ \forall x \in B(x_0, r_0) \cap \Omega,\ \forall t \in [0, \epsilon],\ x + t\theta f \in \Omega \tag{5.20}$$

Using these results and working as in Section 2 easily proves the following.

Proposition 5.5. *There exists* $N \in \mathbb{N}$ *and* $c_1, c_2 > 0$ *such that*

$$M_h^N(x, dy) = \mu_h(x, dy) + c_1 h^{-d} 1_{|x-y|<c_2 h}\, dy \tag{5.21}$$

where for all $x \in \Omega$, $\mu_h(x, dy)$ *is a positive Borel measure.*

Proof. The starting point of the proof is to observe that for any $k \in \mathbb{N}$ and any non-negative function f,

$$K_h^k f(x) \geq \int_{e_1 \in F_1} \dots \int_{e_k \in F_k} K_{h,e_1} \dots K_{h,e_k} f(x)\, d\mu(e_k) \dots d\mu(e_1) \tag{5.22}$$

for any $F_1, \dots, F_k \subset E$. Then the proof is the same as the proof of Proposition 2.2. In fact, (2.13) remains valid thanks to Lemma 5.1. Then we can mimick the end of the proof, using the fact that in Proposition 5.3 the set F_j can be chosen with arbitrary small diameter. Details are left to the reader. □

Proposition 5.5 implies a lemma analogous to Lemma 2.3 for the operator M_h considered in this section. In particular, any function $u \in L^2(\Omega)$ satisfying

$$\|u\|_{L^2}^2 + \langle (1 - M_h)u, u \rangle_{L^2} \leq 1 \tag{5.23}$$

admits a decomposition $u = u_L + u_H$ with $\|u_L\|_{H^1} \leq 1$ and $\|u_H\|_{L^2} = O(h)$. Using Proposition 5.3 and the generalization of Lemma 2.3 gives Parts *i* and *ii* of Theorem 5.2.

Part *iii* is implied by the following lemma (where $\mathcal{B}_h$ still denotes the Dirichet form associated to M_h).

Lemma 5.6. *Let* $\theta \in C^\infty(\overline{\Omega})$ *be fixed and let* $(\varphi_h, r_h) \in H^1(\Omega) \times L^2(\Omega)$ *be such that* $\|r_h\|_{L^2(\Omega)} = O(h)$ *and* φ_h *converges weakly in* $H^1(\Omega)$ *to some* φ*. Then*

$$\lim_{h \to 0^+} h^{-2} \mathcal{B}_h(r_h, \theta) = 0 \tag{5.24}$$

and

$$\lim_{h \to 0^+} h^{-2} \mathcal{B}_h(\varphi_h, \theta) = \frac{1}{6} \int_E \int_\Omega \nabla_E \varphi(e, x) \overline{\nabla_E \theta}(e, x)\, dx d\mu(e). \tag{5.25}$$

Proof. The proof is the same as that of Lemma 2.4. □

Total variation estimates for rates of convergence now follow as in Section 3.

References

[1] Chatterjee, S., Diaconis, P. and Sly, A. (2010). Properties of uniform doubly stochastic matrices. *ArXiv e-prints*. `1010.6136`.

[2] Diaconis, P. and Anderson, H. C. (2007). Hit and run as a unifying device. *J. Soc. Francaise Statist.*, **148** 5–28.

[3] Diaconis, P. and Lebeau, G. (2009). Micro-local analysis for the Metropolis algorithm. *Math. Z.*, **262** 411–447. URL `http://dx.doi.org/10.1007/s00209-008-0383-9`.

[4] Diaconis, P., Lebeau, G. and Michel, L. (2011). Geometric analysis for the Metropolis algorithm on Lipshitz domains. To appear, *Invent. Math.*

[5] Diaconis, P. and Matchup-Wood, P. (2010). On random, doubly stochastic, tri-diagonal matrices. Tech. rep., Department of Mathematics, Stanford University, preprint.

[6] Diaconis, P. and Saloff-Coste, L. (1998). What do we know about the Metropolis algorithm? *J. Comput. System Sci.*, **57** 20–36. 27th Annual ACM Symposium on the Theory of Computing (STOC'95) (Las Vegas, NV).

[7] Lebeau, G. and Michel, L. (2010). Semiclassical analysis of a random walk on a manifold. *Ann. Probab.*, **38** 277–315.

[8] Liu, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. Springer Series in Statistics, Springer-Verlag, New York.

[9] Lovász, L. and Simonovits, M. (1993). Random walks in a convex body and an improved volume algorithm. *Random Structures Algorithms*, **4** 359–412. URL `http://dx.doi.org/10.1002/rsa.3240040402`.

[10] Lovász, L. and Vempala, S. (2003). Hit-and-run is fast and fun. Tech. rep., Microsoft Research, Microsoft Corporation. URL `citeseer.ist.psu.edu/lovasz03hitrun.html`.

[11] Lovász, L. and Vempala, S. (2007). The geometry of logconcave functions and sampling algorithms. *Random Structures Algorithms*, **30** 307–358. URL `http://dx.doi.org/10.1002/rsa.20135`.

[12] Meyn, S. P. and Tweedie, R. L. (1993). *Markov Chains and Stochastic Stability*. Communications and Control Engineering Series, Springer-Verlag London Ltd., London.

[13] Reed, M. and Simon, B. (1978). *Methods of Modern Mathematical Physics. IV. Analysis of Operators*. Academic Press [Harcourt Brace Jovanovich Publishers], New York.

[14] Rosenthal, J. S. (1995). Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.*, **90** 558–566. URL http://links.jstor.org/sici?sici=0162-1459(199506)90:430<558:MCACRF>2.0.CO;2-2&origin=MSN.

[15] Smith, R. L. (1984). Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Oper. Res.*, **32** 1296–1308. URL `http://dx.doi.org/10.1287/opre.32.6.1296`.

[16] Yuen, W. K. (2000). Applications of geometric bounds to the convergence rate of Markov chains on $\mathbb{R}^n$. *Stochastic Process. Appl.*, **87** 1–23. URL `http://dx.doi.org/10.1016/S0304-4149(99)00101-5`.