

# How to Clean a Dirty Floor: Probabilistic Potential Theory and the Dobrushin Uniqueness Theorem

T. de la Rue<sup>1</sup>, R. Fernández<sup>1</sup> and A.D. Sokal<sup>2,\*</sup>

<sup>1</sup>Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS, Université de Rouen,  
Avenue de l'Université, F-76801 Saint Etienne du Rouvray, France.

E-mail: thierry.de-la-rue@univ-rouen.fr, Roberto.Fernandez@univ-rouen.fr

<sup>2</sup>Department of Physics, New York University, 4 Washington Place, New York, NY 10003  
USA. E-mail: sokal@nyu.edu.

Received April 24, 2007

**Abstract.** Motivated by the Dobrushin uniqueness theorem in statistical mechanics, we consider the following situation: Let  $\alpha$  be a nonnegative matrix over a finite or countably infinite index set  $X$ , and define the “cleaning operators”  $\beta_h = I_{1-h} + I_h\alpha$  for  $h: X \rightarrow [0, 1]$  (here  $I_f$  denotes the diagonal matrix with entries  $f$ ). We ask: For which “cleaning sequences”  $h_1, h_2, \dots$  do we have  $c\beta_{h_1} \cdots \beta_{h_n} \rightarrow 0$  for a suitable class of “dirt vectors”  $c$ ? We show, under a modest condition on  $\alpha$ , that this occurs whenever  $\sum_i h_i = \infty$  everywhere on  $X$ . More generally, we analyze the cleaning of subsets  $\Lambda \subseteq X$  and the final distribution of dirt on the complement of  $\Lambda$ . We show that when  $\text{supp}(h_i) \subseteq \Lambda$  with  $\sum_i h_i = \infty$  everywhere on  $\Lambda$ , the operators  $\beta_{h_1} \cdots \beta_{h_n}$  converge as  $n \rightarrow \infty$  to the “balayage operator”  $\Pi_\Lambda = \sum_{k=0}^{\infty} (I_\Lambda \alpha)^k I_{\Lambda^c}$ . These results are obtained in two ways: by a fairly simple matrix formalism, and by a more powerful tree formalism that corresponds to working with formal power series in which the matrix elements of  $\alpha$  are treated as noncommuting indeterminates.

KEYWORDS: nonnegative matrix, spectral radius, probabilistic potential theory, discrete potential theory, balayage, Dobrushin uniqueness theorem

AMS SUBJECT CLASSIFICATION: Primary 60J99, Secondary 15A48, 31C20, 31C99, 60J10, 60J45, 82B20

---

\*Also at Department of Mathematics, University College London, London WC1E 6BT, England.

## Contents

<b>1. Introduction</b>	<b>2</b>
<b>2. Basic set-up</b>	<b>4</b>
2.1. Definition of operators . . . . .	4
2.2. The Fundamental Hypothesis . . . . .	6
2.3. Main results . . . . .	8
<b>3. Matrix approach to balayage</b>	<b>10</b>
3.1. Identities for cleaning operators . . . . .	10
3.2. Inequalities for cleaning operators . . . . .	12
3.3. Identities related to $\mathbf{\Pi}_\Lambda$ . . . . .	14
3.4. Comparison of cleaning operators with $\mathbf{\Pi}_\Lambda$ . . . . .	15
3.5. Convergence of cleaning operators to $\mathbf{\Pi}_\Lambda$ . . . . .	17
3.6. Some further identities and inequalities . . . . .	22
3.7. Converse results . . . . .	26
<b>4. Tree approach to balayage</b>	<b>27</b>
4.1. Markers, clouds and operators . . . . .	28
4.2. Some special clouds . . . . .	31
4.3. A partial ordering on clouds . . . . .	33
4.4. A fundamental inequality . . . . .	45
4.5. Clouds carried by $\Lambda$ . . . . .	46
4.6. $\Lambda$ -regular clouds . . . . .	49
4.7. Comparison of $T_\mu$ with $\mathbf{\Pi}_\Lambda$ . . . . .	50
4.8. Convergence of cleaning operators . . . . .	53
4.8.1. Pointwise (vector-norm) version . . . . .	53
4.8.2. Uniform (operator-norm) version . . . . .	59
4.8.3. Some final remarks . . . . .	62
4.9. Some further identities and inequalities . . . . .	62
<b>5. Alternate sufficient conditions for cleanability</b>	<b>65</b>
<b>6. Converse results</b>	<b>72</b>
6.1. General result . . . . .	72
6.2. Finite $\Lambda$ . . . . .	75

## 1. Introduction

Let  $X$  be a finite or countably infinite set, let  $\mathcal{T}$  be a collection of nonnegative real matrices indexed by  $X$ , and let  $\mathcal{C}$  be a class of nonnegative real vectors indexed by  $X$ . In this paper we want to ask variants of the following question: Under what conditions does there exist a sequence of elements  $(T_i)_{i=1}^\infty$  in  $\mathcal{T}$  such

that  $cT_1 \cdots T_n \rightarrow 0$  for all  $c \in \mathcal{C}$ ? (If  $X$  is infinite, we must of course specify the topology in which this convergence is to be understood.)

Here is a homely but suggestive interpretation: Think of the elements of  $X$  as the “sites” of a dirty floor, the nonnegative vectors  $c = (c_x)_{x \in X}$  as “distributions of dirt”, and the matrices  $T \in \mathcal{T}$  as the “cleaning operators” at our disposal. Application of a cleaning operator  $T$  transforms the dirt distribution from  $c$  to  $cT$ . (Note that we always write our dirt vectors on the left, in analogy with probability distributions in Markov-chain theory.) It is natural to ask: Under what conditions can the floor be completely cleaned? We will also ask: Under what conditions can a subset  $\Lambda \subseteq X$  be cleaned, and in this case, where in  $\Lambda^c := X \setminus \Lambda$  does the dirt go?

These questions arise in mathematical statistical mechanics in connection with the Dobrushin [7, 8, 10, 11, 13, 18, 21] and Dobrushin–Shlosman [1, 9, 22] uniqueness theorems. Indeed, the simplest proofs of these theorems employ a “cleaning” process of precisely the form just discussed.<sup>1</sup> This led us to investigate the cleaning process in its own right.

In this paper we shall not treat the case of an arbitrary family  $\mathcal{T}$  of cleaning operators<sup>2</sup>, but shall focus on the special case of “single-site” cleaning operators: for each  $x \in X$  we are given exactly one cleaning operator  $\beta_x$ , which leaves untouched the dirt on sites other than  $x$  and which distributes the dirt on  $x$  to sites  $y$  with a weight factor  $\alpha_{xy}$ . In other words,

$$(c\beta_x)_y := \begin{cases} c_y + c_x\alpha_{xy} & \text{if } y \neq x, \\ c_x\alpha_{xx} & \text{if } y = x. \end{cases} \quad (1.1)$$

This is the case that arises in the proof of the Dobrushin uniqueness theorem. The definition (1.1) can trivially be rewritten as

$$\beta_x = I_{\{x\}^c} + I_{\{x\}}\alpha, \quad (1.2)$$

---

<sup>1</sup>We learned the “cleaning” interpretation of the Dobrushin uniqueness theorem from Michael Aizenman in the mid-1980s.

<sup>2</sup>The general case leads, in fact, to interesting issues of computational complexity and decidability (see [5] for an excellent survey of closely related problems). Consider the following problem:

*Input:* A finite set  $T_1, \dots, T_m$  of  $n \times n$  matrices with nonnegative rational entries.

*Question:* Does there exist a sequence of indices  $i_1, \dots, i_k$  such that the product  $T_{i_1} \cdots T_{i_k}$  has spectral radius  $< 1$ ?

This problem turns out to be *NP-hard*, even when restricted to matrices with elements 0 and 1 [4, Remarks 2 and 3 after Theorem 2], i.e. an oracle for solving it would permit the polynomial-time solution of any problem in the class NP (e.g. the traveling salesman problem). Even more strikingly, it has very recently been proven [6] that this problem is *algorithmically undecidable* even when restricted to  $m = 3$ ,  $n = 46$  (indeed, even when  $T_1, T_2$  are stochastic matrices and  $T_3$  is a diagonal matrix with a single nonzero element). The proof uses a simple reduction from [3, Theorem 2.1].

where  $I_{\{x\}}$  and  $I_{\{x\}^c}$  are the projection operators on  $\{x\}$  and its complement, respectively. This way of writing the cleaning operators brings out the close relations between our subject and probabilistic potential theory [12, 14, 15, 19]. Indeed, probabilistic potential theory in discrete time and countable state space can be interpreted as the theory of the algebra of operators generated by a single nonnegative matrix  $\alpha = (\alpha_{xy})_{x,y \in X}$  together with all the multiplication operators  $I_f$  (where  $f$  is a real-valued function on  $X$ ) — or more specifically, as the theory of the multiplicative convex cone of operators generated by  $\alpha$  together with all the nonnegative multiplication operators  $I_f$ .

Our main result (Theorem 3.2) is that, under mild conditions on the matrix  $\alpha$  (see Section 2.2), *any* sequence of cleaning operations inside  $\Lambda$  that visits each site of  $\Lambda$  infinitely many times will lead, in the limit, to the *same* result: the dirt will be removed from  $\Lambda$  and transferred to  $\Lambda^c$  as specified by the “balayage operator”  $\Pi_\Lambda$ .

The plan of this paper is as follows: In Section 2 we set forth the basic definitions and state a few of our main results. In Section 3 we analyze the cleaning operators by deriving matrix identities and inequalities in the spirit of probabilistic potential theory [12, 14, 15, 19]; here  $\alpha = (\alpha_{xy})_{x,y \in X}$  is considered to be a fixed matrix of nonnegative real numbers. In Section 4 we introduce an alternate approach that we think clarifies the combinatorial structure of these identities and inequalities: it is based on the tree of finite sequences of elements of  $X$ . In essence, we are now treating the matrix elements  $\alpha_{xy}$  as noncommutative indeterminates; or in physical terms, we are keeping track of the entire trajectory of each particle of dirt, and not merely its endpoint. This approach allows a much finer analysis of the algebra of operators generated by  $\alpha$  and the multiplication operators  $I_f$ .<sup>3</sup> In Section 5 we present some alternative sufficient conditions that guarantee the cleanability of  $\Lambda$ . Finally, in Section 6 we present some converses to our results.

## 2. Basic set-up

Let  $X$  be a finite or countably infinite set (assumed nonempty), and let  $\alpha = (\alpha_{xy})_{x,y \in X}$  be a nonnegative matrix indexed by  $X$ .

### 2.1. Definition of operators

We shall employ the following classes of matrices. (We refer to them as “operators”, but for now we treat them simply as matrices. Later we shall make clear on what space of vectors they act.)

---

<sup>3</sup>Here our approach mirrors the spirit of modern enumerative combinatorics [20], where generating functions are considered in the first instance as *formal power series*, i.e. as an algebraic tool for efficiently manipulating collections of coefficients. Only at a second stage might one insert specific numerical values for the indeterminates and worry about convergence.

Multiplication operators. For each  $\Lambda \subseteq X$ , we denote by  $I_\Lambda$  the projection on  $\Lambda$ , i.e. the matrix

$$(I_\Lambda)_{xy} := \begin{cases} 1 & \text{if } x = y \in \Lambda, \\ 0 & \text{otherwise.} \end{cases} \quad (2.1)$$

More generally, if  $f$  is a real-valued function on  $X$ , we denote by  $I_f$  the operator of multiplication by  $f$ , i.e. the matrix

$$(I_f)_{xy} := f(x)\delta_{xy} = \begin{cases} f(x) & \text{if } x = y, \\ 0 & \text{otherwise.} \end{cases} \quad (2.2)$$

Clearly  $I_\Lambda = I_{\chi_\Lambda}$ , where  $\chi_\Lambda$  denotes the indicator function of  $\Lambda$ .

Cleaning operators. For each  $x \in X$ , we define the *cleaning operator*  $\beta_x$  by

$$\beta_x := I_{\{x\}^c} + I_{\{x\}}\alpha. \quad (2.3)$$

More generally, for each  $\Lambda \subseteq X$ , we define the cleaning operator  $\beta_\Lambda$  by

$$\beta_\Lambda := I_{\Lambda^c} + I_\Lambda\alpha, \quad (2.4)$$

where  $\Lambda^c := X \setminus \Lambda$ . More generally yet, for each function  $f: X \rightarrow [0, 1]$ , we define the cleaning operator  $\beta_f$  by

$$\beta_f := I_{1-f} + I_f\alpha, \quad (2.5)$$

so that  $\beta_\Lambda = \beta_{\chi_\Lambda}$ . We also introduce, for later use, the “dual” cleaning operators

$$\beta_f^* := I_{1-f} + \alpha I_f \quad (2.6)$$

(these have no obvious physical interpretation but will play an important role in our analysis). Clearly all the operators  $\beta_f$  and  $\beta_f^*$  are nonnegative (i.e. have nonnegative matrix elements). We have  $\beta_0 = \beta_0^* = I$  and  $\beta_1 = \beta_1^* = \alpha$ .

*Remark.* In the “single-site cleaning problem” as formulated in the Introduction, the only allowed cleaning operators are the  $\beta_x$  ( $x \in X$ ). When  $\sum_{x \in X} f(x) = 1$  (resp.  $\leq 1$ ), one can interpret  $\beta_f$  as a convex combination of the  $\beta_x$  (resp. of the  $\beta_x$  and the identity operator), hence as the expected output from a random choice of single-site cleaning operator. When  $\sum_{x \in X} f(x) > 1$ , the operators  $\beta_f$  have no such interpretation in terms of single-site cleaning; nevertheless, their introduction is natural from the point of view of probabilistic potential theory, as we shall see. Indeed, the quantity  $\sum_{x \in X} f(x)$  plays no role whatsoever in our analysis; the condition  $0 \leq f \leq 1$ , on the other hand, is crucial. Whenever we write  $\beta_f$  it will be assumed tacitly that  $0 \leq f \leq 1$ .

Balayage operators. For each  $\Lambda \subseteq X$  and each  $n \geq 0$ , we define  $\Pi_\Lambda^{(n)}$  to be the result of cleaning  $n$  times the set  $\Lambda$  and then keeping only the dirt outside  $\Lambda$ :

$$\Pi_\Lambda^{(n)} := \beta_\Lambda^n I_{\Lambda^c} = \sum_{k=0}^n (I_\Lambda \alpha)^k I_{\Lambda^c}. \quad (2.7)$$

Let us note the identities

$$\begin{aligned} \beta_\Lambda^n &= \Pi_\Lambda^{(n)} + (I_\Lambda \alpha I_\Lambda)^n I_\Lambda, \\ \Pi_\Lambda^{(n)} \Pi_\Lambda^{(m)} &= \Pi_\Lambda^{(n)}. \end{aligned} \quad (2.8)$$

We then define the *balayage operator*  $\Pi_\Lambda$  by a limiting process:

$$\Pi_\Lambda := \lim_{n \rightarrow \infty} \uparrow \Pi_\Lambda^{(n)} = \sum_{k=0}^{\infty} (I_\Lambda \alpha)^k I_{\Lambda^c} \quad (2.9)$$

Please note that, with no hypotheses other than the nonnegativity of  $\alpha$ , the matrix elements of  $\Pi_\Lambda$  are well-defined (and nonnegative) but might be  $+\infty$ . However, we shall soon introduce a condition (the Fundamental Hypothesis) under which the matrix elements of  $\Pi_\Lambda$  are finite and indeed  $\Pi_\Lambda$  is a contraction operator on a suitable space of dirt vectors (Lemma 2.1). It furthermore follows from (2.8) that  $\Pi_\Lambda$  is a projection operator, i.e.  $\Pi_\Lambda^2 = \Pi_\Lambda$  (see also Lemma 3.4).

## 2.2. The Fundamental Hypothesis

From now through the end of Section 4 (with the exception of Sections 3.6 and 3.7), we shall make the following

**Fundamental Hypothesis.** There exists a vector  $w = (w_x)_{x \in X}$  with  $w_x > 0$  for all  $x$ , such that  $\alpha w \leq w$ .

In potential theory, a vector  $w$  satisfying  $\alpha w \leq w$  is called *subinvariant*, *superharmonic* or *excessive* with respect to  $\alpha$  [12, 14, 15, 19]. Thus, the Fundamental Hypothesis asserts the existence of a strictly positive subinvariant vector.

*For clarity, all results in this paper that assume the Fundamental Hypothesis will be marked “(FH)”.*

*Remarks.* 1. The Fundamental Hypothesis implies that the operator  $P := I_w^{-1} \alpha I_w$  is submarkovian, i.e. satisfies  $P \mathbf{1} \leq \mathbf{1}$ . Our results could therefore be given a more probabilistic flavor, reminiscent of Markov-chain theory, by rewriting them in terms of  $P$  (this is easy because  $I_w$  commutes with all the other multiplication operators  $I_f$ ). We shall not need this interpretation, however, so we leave this translation to the reader.

2. If  $X$  is finite and  $\alpha$  is irreducible, the Fundamental Hypothesis means simply that the Perron–Frobenius eigenvalue (= spectral radius) of  $\alpha$  is  $\leq 1$ . In this case it is natural to take  $w$  to be the Perron–Frobenius eigenvector (this is the unique choice if the spectral radius equals 1, but is nonunique otherwise). But if  $\alpha$  fails to be irreducible, the Fundamental Hypothesis is *stronger* than this assertion about the spectral radius: consider, for instance,  $\alpha = \begin{pmatrix} 1 & a \\ 0 & b \end{pmatrix}$  with  $a > 0$  and  $0 \leq b \leq 1$ , which has spectral radius 1 but does not satisfy the Fundamental Hypothesis. It can be shown [16] that, when  $X$  is finite, the Fundamental Hypothesis holds if and only if the spectral radius of  $\alpha$  is  $\leq 1$  and each class  $J$  for which the square submatrix  $\alpha_{JJ}$  has spectral radius 1 is a final class (i.e.  $\alpha_{jk} = 0$  whenever  $j \in J$  and  $k \notin J$ ).<sup>4</sup>

3. Example 6.2 below shows that if  $X$  is infinite, then the Fundamental Hypothesis can fail even though each matrix  $I_\Lambda \alpha I_\Lambda$  for  $\Lambda$  finite has spectral radius 0 (i.e. is nilpotent).

We fix once and for all a vector  $w > 0$  satisfying  $\alpha w \leq w$ . For any vector  $c = (c_x)_{x \in X}$ , we define

$$\|c\|_w := \sum_{x \in X} |c_x| w_x \quad (2.10)$$

and we denote by  $l^1(w)$  the space of vectors  $c$  satisfying  $\|c\|_w < \infty$ .

For any matrix  $A = (A_{xy})_{x,y \in X}$ , we define the operator norm

$$\|A\|_{w \rightarrow w} := \sup_{\substack{c \text{ of finite support} \\ c \neq 0}} \frac{\|cA\|_w}{\|c\|_w}. \quad (2.11)$$

If  $\|A\|_{w \rightarrow w} < \infty$ , then  $A$  defines a bounded operator on  $l^1(w)$  of norm  $\|A\|_{w \rightarrow w}$ . It is not hard to see that (2.11) is equivalent to

$$\|A\|_{w \rightarrow w} = \sup_{x \in X} w_x^{-1} \sum_{y \in X} |A_{xy}| w_y. \quad (2.12)$$

In particular, when  $A$  is a nonnegative matrix, (2.11)–(2.12) reduces simply to

$$\|A\|_{w \rightarrow w} = \sup_{x \in X} \frac{(Aw)_x}{w_x}. \quad (2.13)$$

Thus, for nonnegative matrices,  $\|A\|_{w \rightarrow w} \leq 1$  if and only if  $Aw \leq w$ .

The norms of the multiplication operators are trivially given by

$$\|I_f\|_{w \rightarrow w} = \|f\|_\infty := \sup_{x \in X} |f(x)|. \quad (2.14)$$

---

<sup>4</sup>We recall that the *classes* of a nonnegative matrix  $\alpha$  are defined as the strongly connected components of the directed graph with vertex set  $X$  and edge set  $\{xy : \alpha_{xy} > 0\}$ : see e.g. [2, 17].

The Fundamental Hypothesis immediately gives  $\|\alpha\|_{w \rightarrow w} \leq 1$ . More generally, for any  $f: X \rightarrow [0, 1]$ , the cleaning operator  $\beta_f$  satisfies

$$\beta_f w = I_{1-f} w + I_f \alpha w \leq I_{1-f} w + I_f w = w \quad (2.15)$$

and hence  $\|\beta_f\|_{w \rightarrow w} \leq 1$ . A similar argument shows, yet more generally, that for any functions  $(f_i)_{i=1}^n$  and operators  $(A_i)_{i=1}^n$ , we have

$$\left\| \sum_{i=1}^n I_{f_i} A_i \right\|_{w \rightarrow w} \leq \left\| \sum_{i=1}^n \|A_i\|_{w \rightarrow w} |f_i| \right\|_{\infty} \leq \sum_{i=1}^n \|A_i\|_{w \rightarrow w} \|f_i\|_{\infty}. \quad (2.16)$$

Finally, for the balayage operator  $\Pi_{\Lambda}$  we have the easy result:

**Lemma 2.1.** (FH) *For each  $\Lambda \subseteq X$ , the matrix  $\Pi_{\Lambda}$  has finite matrix elements, and indeed satisfies*

$$0 \leq \Pi_{\Lambda} w \leq w, \quad (2.17)$$

so that  $\|\Pi_{\Lambda}\|_{w \rightarrow w} \leq 1$ . More specifically,  $\|\Pi_{\Lambda}\|_{w \rightarrow w}$  equals 1 if  $\Lambda \neq X$ , and equals 0 if  $\Lambda = X$ .

*Proof.* By (2.15) we have  $\Pi_{\Lambda}^{(n)} w = \beta_{\Lambda}^n I_{\Lambda^c} w \leq \beta_{\Lambda}^n w \leq w$ . Since  $0 \leq \Pi_{\Lambda}^{(n)} \uparrow \Pi_{\Lambda}$  elementwise, we have  $\Pi_{\Lambda}^{(n)} w \uparrow \Pi_{\Lambda} w$  elementwise by the monotone convergence theorem, and hence  $\Pi_{\Lambda} w \leq w$ . Since  $\Pi_{\Lambda}$  is nonnegative, we conclude from (2.11) or (2.12) that  $\|\Pi_{\Lambda}\|_{w \rightarrow w} \leq 1$ .

On the other hand, we have  $c \Pi_{\Lambda} = c$  for any vector  $c$  supported on  $\Lambda^c$ ; when  $\Lambda \neq X$  (so that  $c$  can be chosen nonzero) this implies that  $\|\Pi_{\Lambda}\|_{w \rightarrow w} \geq 1$ . When  $\Lambda = X$ , by contrast, we have  $\Pi_{\Lambda} = 0$ .  $\square$

### 2.3. Main results

Let us now state briefly a few of our main results, just to give their flavor. Our principal result on cleaning is the following:

**Theorem 2.1 (= Theorem 3.2).** (FH) *Consider a region  $\Lambda \subseteq X$  and functions  $0 \leq h_i \leq \chi_{\Lambda}$  ( $i \geq 1$ ) such that*

$$\sum_{i=1}^{\infty} h_i(x) = \infty \quad \forall x \in \Lambda. \quad (2.18)$$

*If the nonnegative vector  $c \in l^1(w)$  is such that  $\|c(I_{\Lambda} \alpha I_{\Lambda})^{\ell}\|_w \rightarrow 0$  as  $\ell \rightarrow \infty$ , then*

$$\|c(\beta_{h_1} \cdots \beta_{h_N} - \Pi_{\Lambda})\|_w \xrightarrow{N \rightarrow \infty} 0. \quad (2.19)$$



Intuitively, this says that any sequence of cleaning operations inside  $\Lambda$  that covers  $\Lambda$  infinitely many times will lead, in the limit, to removing the dirt from  $\Lambda$  and transferring it to  $\Lambda^c$  as specified by the balayage operator  $\Pi_\Lambda$ .

We also have a variant of this result (Theorem 3.1) in which the vector-norm convergence (2.19) is strengthened to operator-norm convergence, at the price of a stronger hypothesis on the functions  $h_i$  (namely, requiring that they cover the set  $\Lambda$  uniformly). Finally, we will prove some far-reaching extensions of this result, in terms of the tree formalism, in Section 4.8.

We also have some sufficient conditions for cleanability when the Fundamental Hypothesis is not assumed. For instance:

**Theorem 2.2 (= Theorem 5.2).** *Let  $\Lambda \subseteq X$ , and let  $c \geq 0$  and  $w \geq 0$  be vectors satisfying*

$$(a) \quad c(I_\Lambda \alpha I_\Lambda)^k w < \infty \text{ for all } k \geq 0$$

and

$$(b) \quad \lim_{k \rightarrow \infty} c(I_\Lambda \alpha I_\Lambda)^k w = 0.$$

*Then it is possible to find a sequence of sites  $x_1, x_2, \dots \in \Lambda$  and a sequence of numbers  $\varepsilon_1, \varepsilon_2, \dots \in (0, 1]$  such that*

$$\lim_{n \rightarrow \infty} c \beta_{\varepsilon_1 \delta_{x_1}} \cdots \beta_{\varepsilon_n \delta_{x_n}} I_\Lambda w = 0. \quad (2.20)$$

Although this result refers only to matrices, its proof uses the tree formalism (see Section 5).

Finally, we have a converse result when the Fundamental Hypothesis is not assumed:

**Theorem 2.3 (= Theorem 3.3).** *Let  $X$  be a finite or countably infinite set, let  $\Lambda \subseteq X$ , and let  $c \geq 0$  and  $w \geq 0$  be vectors that are strictly positive on  $\Lambda$ . Consider the following conditions on a matrix  $\alpha$ :*

$$(a) \quad \sum_{k=0}^{\infty} c(I_\Lambda \alpha I_\Lambda)^k w < \infty.$$

$$(b) \quad \text{For all } h: X \rightarrow [0, 1] \text{ with } \text{supp } h = \Lambda \text{ such that } h \geq \varepsilon \chi_\Lambda \text{ for some } \varepsilon > 0, \\ \text{we have } \sum_{k=0}^{\infty} c(I_\Lambda \beta_h I_\Lambda)^k w < \infty.$$

$$(b') \quad \text{There exists } h: X \rightarrow [0, 1] \text{ with } \text{supp } h = \Lambda \text{ such that } \sum_{k=0}^{\infty} c(I_\Lambda \beta_h \times \\ I_\Lambda)^k w < \infty.$$

$$(c) \quad \text{For every finite sequence } f_1, \dots, f_m \text{ of functions } X \rightarrow [0, 1] \text{ with } \text{supp}(f_i) \subseteq \\ \Lambda \text{ such that } \sum_i f_i \geq \varepsilon \chi_\Lambda \text{ for some } \varepsilon > 0, \text{ we have } \sum_{k=0}^{\infty} c(I_\Lambda \beta_{f_1} \cdots \times \\ \beta_{f_m} I_\Lambda)^k w < \infty.$$

(c') *There exists a finite sequence  $f_1, \dots, f_m$  of functions  $X \rightarrow [0, 1]$  with  $\text{supp}(f_i) \subseteq \Lambda$  such that  $\sum_{k=0}^{\infty} c(I_\Lambda \beta_{f_1} \cdots \beta_{f_m} I_\Lambda)^k w < \infty$ .*

*Then (a)–(c) are all equivalent and imply (c'); and for matrices  $\alpha$  satisfying the additional hypothesis*

$$\text{There exists a constant } C < \infty \text{ such that } I_\Lambda \alpha I_\Lambda w \leq Cw \quad (2.21)$$

*all five conditions are equivalent.*

Intuitively, this says (at least if  $\Lambda$  is a finite set) that if the spectral radius of  $I_\Lambda \alpha I_\Lambda$  is  $\geq 1$ , then there is no way to clean the set  $\Lambda$  completely. In fact, we shall prove a stronger version of this result in Section 6, using the tree formalism (see Theorem 6.1 and Corollary 6.1).

### 3. Matrix approach to balayage

In this section we study the algebra of operators generated by  $\alpha$  and the multiplication operators  $I_f$ , by deriving matrix identities and inequalities in the spirit of probabilistic potential theory [12, 14, 15, 19]. Our main goal is to study the convergence of a product of cleaning operators  $\beta_{h_1} \cdots \beta_{h_n}$  where all the  $h_i$  have support contained in  $\Lambda$ . We shall show that, under very general conditions, any such product must converge to  $\Pi_\Lambda$  (see Theorems 3.1 and 3.2). We shall also prove a converse result when the Fundamental Hypothesis is not assumed (Theorem 3.3).

#### 3.1. Identities for cleaning operators

**Proposition 3.1 (Fundamental identities).** *For all  $h, h_1, h_2: X \rightarrow [0, 1]$ , the following identities hold:*

(i) *Intertwining I:*

$$(I - \alpha)\beta_h = \beta_h^*(I - \alpha). \quad (3.1)$$

(ii) *Intertwining II:*

$$(\beta_{h_1} - I)I_{h_2} = I_{h_1}(\beta_{h_2}^* - I), \quad (3.2)$$

$$\beta_h I_h = I_h \beta_h^*. \quad (3.3)$$

(iii) *Comparison:*

$$\beta_{h_1} - \beta_{h_2} = I_{h_2 - h_1}(I - \alpha), \quad (3.4)$$

$$I - \beta_h = I_h(I - \alpha). \quad (3.5)$$

(iv) *Collapse:*

$$\beta_{h_1}\beta_{h_2} = \beta_{1-(1-h_1)(1-h_2)} - I_{h_1}\alpha I_{h_2}(I - \alpha). \quad (3.6)$$

*Proof.* (i)

$$\alpha\beta_h - \beta_h^*\alpha = \alpha I_{1-h} - I_{1-h}\alpha = I_h\alpha - \alpha I_h = \beta_h - \beta_h^*. \quad (3.7)$$

(ii) Both sides of (3.2) are equal to  $I_{h_1}(\alpha - I)I_{h_2}$ . Equation (3.3) follows by setting  $h_1 = h_2 = h$ .

(iii)

$$\beta_{h_1} - \beta_{h_2} = I_{1-h_1} - I_{1-h_2} + (I_{h_1} - I_{h_2})\alpha = I_{h_2-h_1}(I - \alpha). \quad (3.8)$$

Equation (3.5) follows by setting  $h_1 = 0$ ,  $h_2 = h$ .

(iv)

$$\begin{aligned} \beta_{h_1}\beta_{h_2} + I_{h_1}\alpha I_{h_2}(I - \alpha) &= I_{(1-h_1)(1-h_2)} + I_{h_1}\alpha I_{1-h_2} + I_{1-h_1}I_{h_2}\alpha + I_{h_1}\alpha I_{h_2} \\ &= I_{(1-h_1)(1-h_2)} + I_{h_1}\alpha + I_{1-h_1}I_{h_2}\alpha \\ &= \beta_{1-(1-h_1)(1-h_2)}. \end{aligned} \quad (3.9)$$

□

**Corollary 3.1 (Telescoping comparison).** *For all  $g_1, \dots, g_n, h_1, \dots, h_n: X \rightarrow [0, 1]$ , we have*

$$\begin{aligned} \beta_{g_1} \cdots \beta_{g_n} - \beta_{h_1} \cdots \beta_{h_n} &= \sum_{i=1}^n \beta_{g_1} \cdots \beta_{g_{i-1}} I_{h_i-g_i}(I - \alpha) \beta_{h_{i+1}} \cdots \beta_{h_n} \\ &= \sum_{i=1}^n \beta_{g_1} \cdots \beta_{g_{i-1}} I_{h_i-g_i} \beta_{h_{i+1}}^* \cdots \beta_{h_n}^* (I - \alpha). \end{aligned} \quad (3.10)$$

*Proof.* This is an immediate consequence of the telescopic decomposition

$$\beta_{g_1} \cdots \beta_{g_n} - \beta_{h_1} \cdots \beta_{h_n} = \sum_{i=1}^n \beta_{g_1} \cdots \beta_{g_{i-1}} (\beta_{g_i} - \beta_{h_i}) \beta_{h_{i+1}} \cdots \beta_{h_n}, \quad (3.11)$$

the comparison identity (3.4), and the intertwining relation (3.1). □

**Lemma 3.1 (Cleaners with restricted support).** *Suppose that  $\text{supp}(g_i) \subseteq \Lambda$  for  $i = 1, \dots, n$ , and let  $h_i: X \rightarrow \mathbb{R}$  be any functions satisfying  $h_i \upharpoonright \Lambda \equiv 1$  for  $i = 1, \dots, n$ . Then*

$$\beta_{g_1} \cdots \beta_{g_n} I_\Lambda = I_{h_1} \beta_{g_1} \cdots I_{h_n} \beta_{g_n} I_\Lambda \quad (3.12)$$

and

$$I_\Lambda \beta_{g_1}^* \cdots \beta_{g_n}^* = I_\Lambda \beta_{g_1}^* I_{h_1} \cdots \beta_{g_n}^* I_{h_n}. \quad (3.13)$$

Furthermore,

$$I_{\Lambda^c} \beta_{g_1} \cdots \beta_{g_n} = I_{\Lambda^c} = \beta_{g_1}^* \cdots \beta_{g_n}^* I_{\Lambda^c}. \quad (3.14)$$

*Proof.* Note first that if  $\text{supp}(g) \subseteq \Lambda$ , we have  $I_{\Lambda^c}\beta_g I_\Lambda = 0$ ; therefore, for any  $h$  satisfying  $h \upharpoonright \Lambda \equiv 1$ , we have

$$\beta_g I_\Lambda = I_\Lambda \beta_g I_\Lambda = I_h \beta_g I_\Lambda. \quad (3.15)$$

So, starting with  $\beta_{g_1} \cdots \beta_{g_n} I_\Lambda$ , we successively use the first equality in (3.15), working from right to left in the product, to transform it into  $I_\Lambda \beta_{g_1} \cdots I_\Lambda \beta_{g_n} I_\Lambda$ . We then successively use the second equality in (3.15), working from left to right in the product, to transform it into  $I_{h_1} \beta_{g_1} \cdots I_{h_n} \beta_{g_n} I_\Lambda$ . This proves (3.12). An analogous argument gives (3.13). Identities (3.14) are immediately verified by induction.  $\square$

### 3.2. Inequalities for cleaning operators

We now turn our attention to proving inequalities that say, roughly speaking, that one operator “cleans better” than another.

**Lemma 3.2.** (FH) *Let  $\Lambda$  be any subset of  $X$ , and let  $f$  be any function satisfying  $\chi_\Lambda \leq f \leq 1$ . Then*

$$I_\Lambda(I - \alpha)I_f w \geq 0. \quad (3.16)$$

*If, in addition,  $0 \leq h_i \leq \chi_\Lambda$  for  $i = 1, \dots, n$ , then*

$$I_\Lambda(I - \alpha)\beta_{h_1} \cdots \beta_{h_n} I_f w \geq 0. \quad (3.17)$$

*Proof.* We have

$$\begin{aligned} I_\Lambda(I - \alpha)I_f w &= I_\Lambda f w - I_\Lambda \alpha f w \\ &\geq I_\Lambda f w - I_\Lambda \alpha w && \text{[since } f \leq 1\text{]} \\ &\geq I_\Lambda f w - I_\Lambda w && \text{[since } \alpha w \leq w\text{]} \\ &= I_\Lambda(f - 1)w = 0 && \text{[since } f \equiv 1 \text{ on } \Lambda\text{]}. \end{aligned} \quad (3.18)$$

This proves (3.16). Then

$$\begin{aligned} I_\Lambda(I - \alpha)\beta_{h_1} \cdots \beta_{h_n} I_f w &= I_\Lambda \beta_{h_1}^* \cdots \beta_{h_n}^* (I - \alpha)I_f w \\ &= I_\Lambda \beta_{h_1}^* \cdots \beta_{h_n}^* I_\Lambda (I - \alpha)I_f w \\ &\geq 0 \end{aligned} \quad (3.19)$$

where the first equality uses the intertwining relation (3.1), the second uses the identity (3.13), and the final inequality uses (3.16) and the nonnegativity of  $\beta_{h_1}^*, \dots, \beta_{h_n}^*$ .  $\square$

**Proposition 3.2 (Multi-monotonicity).** (FH) *Suppose that  $0 \leq g_i \leq h_i \leq \chi_\Lambda \leq f \leq 1$  for  $i = 1, \dots, n$ . Then*

$$\beta_{h_1} \cdots \beta_{h_n} I_f w \leq \beta_{g_1} \cdots \beta_{g_n} I_f w. \quad (3.20)$$

*Proof.* We will prove (3.20) by proving separately that  $I_{\Lambda^c}(\text{LHS}) \leq I_{\Lambda^c}(\text{RHS})$  and that  $I_{\Lambda}(\text{LHS}) \leq I_{\Lambda}(\text{RHS})$ . The former is in fact equality, since  $I_{\Lambda^c}\beta_{h_1} \cdots \times \beta_{h_n} = I_{\Lambda^c}\beta_{g_1} \cdots \beta_{g_n} = I_{\Lambda^c}$ . The latter follows from the telescoping comparison identity (3.10) and the inequality (3.17), along with the nonnegativity of the operators  $\beta_{g_1}, \dots, \beta_{g_{i-1}}$  that lie on the left in (3.10).  $\square$

**Proposition 3.3 (Collapse inequality).** (FH) *Suppose that  $0 \leq g_i, h_j \leq \chi_{\Lambda} \leq f \leq 1$  for  $i = 1, \dots, n$  and  $j = 1, \dots, m$ . Then*

$$\beta_{g_1} \cdots \beta_{g_n} \beta_{h_1} \cdots \beta_{h_m} I_f w \leq \beta_{1 - \prod_{i=1}^n (1 - g_i)} \beta_{h_1} \cdots \beta_{h_m} I_f w. \quad (3.21)$$

*Proof.* For  $n = 1$  this is trivial. For  $n = 2$  it follows from the collapse identity (3.6) together with the inequality (3.17). The cases  $n \geq 3$  are obtained by an easy induction from the case  $n = 2$ .  $\square$

Combining Propositions 3.2 and 3.3, we obtain the following comparison result:

**Corollary 3.2 (Multi-monotonicity + collapse).** (FH) *Suppose that  $0 \leq h_i, g_j \leq \chi_{\Lambda} \leq f \leq 1$  for  $i = 1, \dots, N$  and  $j = 1, \dots, k$ . Suppose further that there exist integers  $0 \leq n_0 < n_1 < \dots < n_k \leq N$  such that*

$$1 - \prod_{i=n_{j-1}+1}^{n_j} (1 - h_i) \geq g_j \quad (3.22)$$

for all  $j$ . Then

$$\beta_{h_1} \cdots \beta_{h_N} I_f w \leq \beta_{g_1} \cdots \beta_{g_k} I_f w. \quad (3.23)$$

Now we develop some analogous inequalities going in the *reverse* direction provided that we look only at the dirt *outside*  $\Lambda$ ; moreover, these inequalities hold pointwise.

**Lemma 3.3.** *Let  $\Lambda$  be any subset of  $X$ . Then*

$$I_{\Lambda}(I - \alpha)I_{\Lambda^c} \leq 0. \quad (3.24)$$

If, in addition,  $0 \leq h_i \leq \chi_{\Lambda}$  for  $i = 1, \dots, n$ , then

$$I_{\Lambda}(I - \alpha)\beta_{h_1} \cdots \beta_{h_n} I_{\Lambda^c} \leq 0. \quad (3.25)$$

*Proof.* The inequality (3.24) is trivial since  $\alpha \geq 0$ . To prove (3.25), we make a computation analogous to that in the proof of Lemma 3.2:

$$\begin{aligned} I_{\Lambda}(I - \alpha)\beta_{h_1} \cdots \beta_{h_n} I_{\Lambda^c} &= I_{\Lambda}\beta_{h_1}^* \cdots \beta_{h_n}^*(I - \alpha)I_{\Lambda^c} \\ &= I_{\Lambda}\beta_{h_1}^* \cdots \beta_{h_n}^* I_{\Lambda}(I - \alpha)I_{\Lambda^c} \leq 0 \end{aligned} \quad (3.26)$$

where the final step uses (3.24).  $\square$

**Proposition 3.4 (Reverse multi-monotonicity).** *Suppose that  $0 \leq h_i \leq g_i \leq \chi_\Lambda$  for  $i = 1, \dots, n$ . Then*

$$\beta_{g_1} \cdots \beta_{g_n} I_{\Lambda^c} \geq \beta_{h_1} \cdots \beta_{h_n} I_{\Lambda^c}. \quad (3.27)$$

*Proof.* We will prove (3.27) by proving separately that  $I_{\Lambda^c}(\text{LHS}) \leq I_{\Lambda^c}(\text{RHS})$  and that  $I_\Lambda(\text{LHS}) \leq I_\Lambda(\text{RHS})$ . The former is in fact equality because of (3.14). The latter follows from the telescoping comparison identity (3.10) and the inequality (3.25), along with the nonnegativity of the operators  $\beta_{h_1}, \dots, \beta_{h_{i-1}}$  that lie on the left in (3.10).  $\square$

**Proposition 3.5 (Reverse collapse inequality).** *Suppose that  $0 \leq g_i, h_j \leq \chi_\Lambda$  for  $i = 1, \dots, n$  and  $j = 1, \dots, m$ . Then*

$$\beta_{g_1} \cdots \beta_{g_n} \beta_{h_1} \cdots \beta_{h_m} I_{\Lambda^c} \geq \beta_{1 - \prod_{i=1}^n (1 - g_i)} \beta_{h_1} \cdots \beta_{h_m} I_{\Lambda^c}. \quad (3.28)$$

*Proof.* For  $n = 1$  this is trivial. For  $n = 2$  it follows from the collapse identity (3.6) together with the inequality (3.24). The cases  $n \geq 3$  are obtained by an easy induction from the case  $n = 2$ .  $\square$

Combining Propositions 3.4 and 3.5, we obtain:

**Corollary 3.3 (Reverse multi-monotonicity + collapse).** *Suppose that  $0 \leq h_i, g_j \leq \chi_\Lambda$  for  $i = 1, \dots, N$  and  $j = 1, \dots, k$ . Suppose further that there exist integers  $0 \leq n_0 < n_1 < \dots < n_k \leq N$  such that*

$$1 - \prod_{i=n_{j-1}+1}^{n_j} (1 - h_i) \geq g_j \quad (3.29)$$

for all  $j$ . Then

$$\beta_{h_1} \cdots \beta_{h_N} I_{\Lambda^c} \geq \beta_{g_1} \cdots \beta_{g_k} I_{\Lambda^c}. \quad (3.30)$$

### 3.3. Identities related to $\Pi_\Lambda$

**Lemma 3.4 (Properties of  $\Pi_\Lambda$ ).** *Let  $\Lambda \subseteq X$ . Then the following identities hold:*

(i) *Basic properties of  $\Pi_\Lambda$ :*

$$\Pi_\Lambda^2 = \Pi_\Lambda, \quad (3.31)$$

$$I_\Lambda(I - \Pi_\Lambda) = I - \Pi_\Lambda. \quad (3.32)$$

(ii) *Absorption of cleaning operators: If  $\text{supp}(h) \subseteq \Lambda$ , then*

$$I_h \Pi_\Lambda = I_h \alpha \Pi_\Lambda, \quad (3.33)$$

$$\beta_h \Pi_\Lambda = \Pi_\Lambda, \quad (3.34)$$

$$\Pi_\Lambda \beta_h = \Pi_\Lambda. \quad (3.35)$$

*Proof.* It follows immediately from the definition (2.9) of  $\Pi_\Lambda$  that  $\Pi_\Lambda I_{\Lambda^c} = \Pi_\Lambda$  and  $I_{\Lambda^c} \Pi_\Lambda = I_{\Lambda^c}$ , hence  $\Pi_\Lambda \Pi_\Lambda = \Pi_\Lambda I_{\Lambda^c} \Pi_\Lambda = \Pi_\Lambda I_{\Lambda^c} = \Pi_\Lambda$ . This proves (3.31). Equation (3.32) is trivially equivalent to  $I_{\Lambda^c} \Pi_\Lambda = I_{\Lambda^c}$ . It also follows immediately from the definition of  $\Pi_\Lambda$  that

$$I_\Lambda \Pi_\Lambda = \Pi_\Lambda - I_{\Lambda^c} = I_\Lambda \alpha \Pi_\Lambda, \quad (3.36)$$

and premultiplying this by  $I_h$  yields (3.33). The identity (3.34) follows immediately from (3.33) and the definition of  $\beta_h$ . The identity (3.35) follows from  $\Pi_\Lambda \beta_h = \Pi_\Lambda I_{\Lambda^c} (I_{1-h} + I_h \alpha) = \Pi_\Lambda I_{\Lambda^c} = \Pi_\Lambda$ .  $\square$

**Proposition 3.6 (Convergence-to-balayage identity).** *Suppose that  $0 \leq h_i \leq \chi_\Lambda$  for  $i = 1, \dots, n$ . Then*

$$\begin{aligned} \beta_{h_1} \cdots \beta_{h_n} - \Pi_\Lambda &= \beta_{h_1} \cdots \beta_{h_n} I_\Lambda (I - \Pi_\Lambda) \\ &= (I_\Lambda \beta_{h_1} I_\Lambda) \cdots (I_\Lambda \beta_{h_n} I_\Lambda) I_\Lambda (I - \Pi_\Lambda). \end{aligned} \quad (3.37)$$

*Proof.* We shall prove the first line of (3.37) by induction on  $n$ . It is true for  $n = 0$ , by (3.32). So assume it is true for  $n - 1$ , i.e. that

$$\beta_{h_2} \cdots \beta_{h_n} - \Pi_\Lambda = \beta_{h_2} \cdots \beta_{h_n} I_\Lambda (I - \Pi_\Lambda). \quad (3.38)$$

Left-multiplying by  $\beta_{h_1}$  and using (3.34), the desired identity follows. The alternative form (second line of (3.37)) is then an immediate consequence of Lemma 3.1.  $\square$

*Remark.* In the absence of the Fundamental Hypothesis, some of the matrix elements of  $\Pi_\Lambda$  could be  $+\infty$ , but the identities (3.31)–(3.35) and (3.37) continue to hold (with some matrix elements possibly  $+\infty$  or  $-\infty$ ).

### 3.4. Comparison of cleaning operators with $\Pi_\Lambda$

**Definition 3.1.** Let  $\Lambda \subseteq X$ . We say that an operator  $A \geq 0$  is *absorbed by  $\Pi_\Lambda$*  in case  $A \Pi_\Lambda = \Pi_\Lambda$ .

By Lemma 3.4, operators  $\beta_{h_1} \cdots \beta_{h_n}$  with  $\text{supp}(h_i) \subseteq \Lambda$  are absorbed by  $\Pi_\Lambda$ , as are all convex combinations thereof.

Operators absorbed by  $\Pi_\Lambda$  obey some elementary but remarkable identities and inequalities:

**Lemma 3.5 (Comparison with  $\Pi_\Lambda$ ).** (FH) *Let  $\Lambda \subseteq X$ , and let  $A \geq 0$  be an operator absorbed by  $\Pi_\Lambda$ . Then:*

- (i)  $0 \leq \Pi_\Lambda w \leq Aw$ . [In particular, if  $\Lambda \neq X$ , we have  $\|A\|_{w \rightarrow w} \geq 1$ .]

(ii) *The operator  $A - \Pi_\Lambda$  can be decomposed in the form*

$$A - \Pi_\Lambda = (A - \Pi_\Lambda)I_\Lambda + (A - \Pi_\Lambda)I_{\Lambda^c} \quad (3.39)$$

$$= AI_\Lambda - (\Pi_\Lambda - AI_{\Lambda^c}) \quad (3.40)$$

$$= AI_\Lambda(I - \Pi_\Lambda) \quad (3.41)$$

where

$$AI_\Lambda \geq 0, \quad (3.42)$$

$$\Pi_\Lambda - AI_{\Lambda^c} = AI_\Lambda \Pi_\Lambda \geq 0, \quad (3.43)$$

$$(\Pi_\Lambda - AI_{\Lambda^c})w \leq AI_\Lambda w. \quad (3.44)$$

(iii) *The following norm relations hold:*

$$\|A - \Pi_\Lambda\|_{w \rightarrow w} = \|AI_\Lambda(I - \Pi_\Lambda)\|_{w \rightarrow w} \leq 2\|AI_\Lambda\|_{w \rightarrow w} \quad (3.45)$$

and, for every vector  $c \geq 0$ ,

$$\|c(A - \Pi_\Lambda)\|_w = \|cAI_\Lambda\|_w + \|c(\Pi_\Lambda - AI_{\Lambda^c})\|_w \leq 2\|cAI_\Lambda\|_w. \quad (3.46)$$

*Proof.* By Lemma 2.1 we have  $0 \leq \Pi_\Lambda w \leq w$ . Applying  $A$  on the left and using  $A\Pi_\Lambda = \Pi_\Lambda$ , we obtain (i). The remark in brackets is obtained by left-multiplying with any vector  $c \geq 0$  ( $c \neq 0$ ) supported on  $\Lambda^c$ .

Equalities (3.39), (3.40) are trivial, and (3.41) follows from  $A\Pi_\Lambda = \Pi_\Lambda$  using (3.32):  $A - \Pi_\Lambda = A(I - \Pi_\Lambda) = AI_\Lambda(I - \Pi_\Lambda)$ . (3.42) is trivial. The hypothesis  $A\Pi_\Lambda = \Pi_\Lambda$  yields  $\Pi_\Lambda - AI_{\Lambda^c} = A(\Pi_\Lambda - I_{\Lambda^c}) = AI_\Lambda \Pi_\Lambda \geq 0$ , which is (3.43). Finally, (3.44) is a rewriting of (i).

The identity in (3.45) is an application of (3.41), and the inequality follows from the fact that  $\|I - \Pi_\Lambda\|_{w \rightarrow w} \leq 2$  (cf. Lemma 2.1).

Since  $cAI_\Lambda$  is supported on  $\Lambda$  while  $c(\Pi_\Lambda - AI_{\Lambda^c})$  is supported on  $\Lambda^c$ , the equality in (3.46) follows from (3.39)–(3.41). The inequality follows from (3.44) together with the nonnegativity of the operators  $\Pi_\Lambda - AI_{\Lambda^c}$  and  $AI_\Lambda$ .  $\square$

We can now compare the “efficiency of cleaning” of two operators  $A$  and  $B$ .

**Corollary 3.4. (Comparison of cleaners).** (FH) *Let  $\Lambda \subseteq X$ , and let  $A, B \geq 0$  be operators absorbed by  $\Pi_\Lambda$ . Suppose further that*

$$AI_\Lambda w \leq BI_\Lambda w, \quad (3.47)$$

$$AI_{\Lambda^c} w \geq BI_{\Lambda^c} w. \quad (3.48)$$

Then

$$\|A - \Pi_\Lambda\|_{w \rightarrow w} \leq \|B - \Pi_\Lambda\|_{w \rightarrow w} \quad (3.49)$$

and, for every vector  $c \geq 0$ ,

$$\|c(A - \Pi_\Lambda)\|_w \leq \|c(B - \Pi_\Lambda)\|_w. \quad (3.50)$$



*Proof.* The vector-norm inequality (3.50) is an immediate consequence of the equality in (3.46) together with the hypotheses (3.47), (3.48).

For the operator-norm inequality (3.49), note that the decomposition (3.40) yields

$$(A - \Pi_\Lambda)_{xy} = \begin{cases} A_{xy} & \text{for } y \in \Lambda, \\ -(\Pi_\Lambda - A)_{xy} & \text{for } y \in \Lambda^c. \end{cases} \quad (3.51)$$

Inequalities (3.42) and (3.43) therefore imply the componentwise identity

$$|A - \Pi_\Lambda| = AI_\Lambda + (\Pi_\Lambda - AI_{\Lambda^c}). \quad (3.52)$$

Inequality (3.49) is an immediate consequence of this identity and the hypotheses (3.47), (3.48) together with (2.12).  $\square$

The applications of interest to us follow from the multi-monotonicity bounds (3.20) and (3.27) and the collapse inequalities (3.21) and (3.28):

**Corollary 3.5.** (FH) *Let  $\Lambda \subseteq X$ .*

(i) *Suppose that  $0 \leq g_i \leq h_i \leq \chi_\Lambda$  for  $i = 1, \dots, n$ . Then*

$$\|\beta_{h_1} \cdots \beta_{h_n} - \Pi_\Lambda\|_{w \rightarrow w} \leq \|\beta_{g_1} \cdots \beta_{g_n} - \Pi_\Lambda\|_{w \rightarrow w} \quad (3.53)$$

and, for each vector  $c \geq 0$ ,

$$\|c(\beta_{h_1} \cdots \beta_{h_n} - \Pi_\Lambda)\|_w \leq \|c(\beta_{g_1} \cdots \beta_{g_n} - \Pi_\Lambda)\|_w. \quad (3.54)$$

(ii) *Suppose that  $0 \leq h_i \leq \chi_\Lambda$  for  $i = 1, \dots, N$ . Then, for every choice of integers  $0 \leq n_1 \leq \dots \leq n_k \leq N$ , we have*

$$\|\beta_{h_1} \cdots \beta_{h_N} - \Pi_\Lambda\|_{w \rightarrow w} \leq \|\beta_{1 - \prod_{i=1}^{n_1} (1-h_i)} \cdots \beta_{1 - \prod_{i=n_{k-1}+1}^{n_k} (1-h_i)} - \Pi_\Lambda\|_{w \rightarrow w} \quad (3.55)$$

and, for each vector  $c \geq 0$ ,

$$\|c(\beta_{h_1} \cdots \beta_{h_N} - \Pi_\Lambda)\|_w \leq \|c(\beta_{1 - \prod_{i=1}^{n_1} (1-h_i)} \cdots \beta_{1 - \prod_{i=n_{k-1}+1}^{n_k} (1-h_i)} - \Pi_\Lambda)\|_w. \quad (3.56)$$

### 3.5. Convergence of cleaning operators to $\Pi_\Lambda$

We are now ready to study the convergence of cleaning operators  $\beta_{h_1} \cdots \beta_{h_n}$  to  $\Pi_\Lambda$ . We shall prove the main result in two versions: a uniform (operator-norm) version, and a dust-dependent (vector-norm) version. A central role in these analyses will be played, respectively, by the quantities

$$\rho_\Lambda(\ell) := \|(I_\Lambda \alpha I_\Lambda)^\ell\|_{w \rightarrow w} \quad (3.57)$$

and, for each vector  $c \geq 0$ ,

$$\rho_\Lambda(\ell; c) := \|c(I_\Lambda \alpha I_\Lambda)^\ell\|_w. \quad (3.58)$$

Clearly we have  $0 \leq \rho_\Lambda(\ell) \leq 1$  and  $0 \leq \rho_\Lambda(\ell; c) \leq \|c\|_w$ . Note also that both  $\rho_\Lambda(\ell)$  and [since  $c \geq 0$ ]  $\rho_\Lambda(\ell; c)$  are increasing functions of  $\Lambda$ .

For brevity let us denote

$$\beta_{\varepsilon\Lambda} := \beta_{\varepsilon\chi_\Lambda} \quad (3.59)$$

for  $0 < \varepsilon < 1$  and a set  $\Lambda \subseteq X$ .

**Lemma 3.6.** *Fix a region  $\Lambda \subseteq X$  and a number  $\varepsilon > 0$ .*

(i) *If  $\rho_\Lambda(\ell) \rightarrow 0$  as  $\ell \rightarrow \infty$ , then*

$$\|(I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^N\|_{w \rightarrow w} \xrightarrow[N \rightarrow \infty]{} 0 \quad (3.60)$$

*uniformly for all regions  $\Lambda' \subseteq \Lambda$ .*

(ii) *If the nonnegative vector  $c \in l^1(w)$  is such that  $\rho_\Lambda(\ell; c) \rightarrow 0$  as  $\ell \rightarrow \infty$ , then*

$$\|c(I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^N\|_w \xrightarrow[N \rightarrow \infty]{} 0 \quad (3.61)$$

*uniformly for all regions  $\Lambda' \subseteq \Lambda$ .*

*Proof.* The obvious identity

$$(I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^N = [(1 - \varepsilon)I_{\Lambda'} + \varepsilon I_{\Lambda'} \alpha I_{\Lambda'}]^N = \sum_{\ell=0}^N \binom{N}{\ell} (1 - \varepsilon)^{N-\ell} \varepsilon^\ell (I_{\Lambda'} \alpha I_{\Lambda'})^\ell I_{\Lambda'} \quad (3.62)$$

(valid for  $N \geq 1$ ) allows us to write

$$\left. \begin{aligned} &\|(I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^N\|_{w \rightarrow w} \\ &\|c(I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^N\|_w \end{aligned} \right\} \leq E(F_{\Lambda'}(X_N)) \quad (3.63)$$

where  $X_N$  is a Binomial( $N, \varepsilon$ ) random variable and

$$F_{\Lambda'}(\ell) = \begin{cases} \rho_{\Lambda'}(\ell) & \text{in case (i),} \\ \rho_{\Lambda'}(\ell; c) & \text{in case (ii).} \end{cases} \quad (3.64)$$

In both cases,

$$F_{\Lambda'}(\ell) \leq F_\Lambda(\ell) \xrightarrow[\ell \rightarrow \infty]{} 0. \quad (3.65)$$

Denoting

$$M_K := \sup_{\ell \geq K} F_\Lambda(\ell), \quad (3.66)$$

we can decompose

$$E(F_{\Lambda'}(X_N)) \leq M_0 P(X_N \leq K) + M_K, \quad (3.67)$$

where  $M_0 < \infty$  [in case (ii) because  $c \in l^1(w)$ ]. By hypothesis, given  $\delta > 0$  we can choose  $K$  so that  $M_K \leq \delta/2$ . For such  $K$ ,

$$P(X_N \leq K) \leq (1 - \varepsilon)^{N-K} N^K \sum_{\ell=0}^K (1 - \varepsilon)^{K-\ell} \varepsilon^\ell \xrightarrow{N \rightarrow \infty} 0. \quad (3.68)$$

Thus for  $N$  large enough the first term on the right-hand-side of (3.67) is also smaller than  $\delta/2$ .  $\square$

We remark that in the situation (i), the inequality

$$\|(I_{\Lambda'} \beta_{\varepsilon \Lambda'} I_{\Lambda'})^{N_1 + N_2}\|_{w \rightarrow w} \leq \|(I_{\Lambda'} \beta_{\varepsilon \Lambda'} I_{\Lambda'})^{N_1}\|_{w \rightarrow w} \|(I_{\Lambda'} \beta_{\varepsilon \Lambda'} I_{\Lambda'})^{N_2}\|_{w \rightarrow w} \quad (3.69)$$

implies that the convergence in (3.60) is actually exponentially fast in  $N$ .

**Theorem 3.1 (Uniform cleaning).** (FH) Consider a region  $\Lambda \subseteq X$  and functions  $0 \leq h_i \leq \chi_\Lambda$  ( $i \geq 1$ ) such that

$$\sum_{i=1}^{\infty} h_i(x) = \infty \quad \text{uniformly for } x \in \Lambda. \quad (3.70)$$

(i) If  $\rho_\Lambda(\ell) \equiv \|(I_\Lambda \alpha I_\Lambda)^\ell\|_{w \rightarrow w} \rightarrow 0$  as  $\ell \rightarrow \infty$ , then

$$\|\beta_{h_1} \cdots \beta_{h_N} - \Pi_\Lambda\|_{w \rightarrow w} \xrightarrow{N \rightarrow \infty} 0. \quad (3.71)$$

(ii) If the nonnegative vector  $c \in l^1(w)$  is such that  $\rho_\Lambda(\ell; c) \equiv \|c(I_\Lambda \alpha I_\Lambda)^\ell\|_w \rightarrow 0$  as  $\ell \rightarrow \infty$ , then

$$\|c(\beta_{h_1} \cdots \beta_{h_N} - \Pi_\Lambda)\|_w \xrightarrow{N \rightarrow \infty} 0. \quad (3.72)$$

We remark that if  $\Lambda$  is a *finite* set, then the hypothesis (3.70) is equivalent to the apparently weaker hypothesis that  $\sum_{i=1}^{\infty} h_i(x) = \infty$  for all  $x \in \Lambda$  [cf. (3.83) below]. But if  $\Lambda$  is an *infinite* set, then (3.70) is stronger.

*Proof of Theorem 3.1.* By Proposition 3.6 and the fact that  $\|I - \Pi_\Lambda\|_{w \rightarrow w} \leq 2$  [or alternatively by Lemma 3.5(iii)], it is enough to show that

$$\|\beta_{h_1} \cdots \beta_{h_N} I_\Lambda\|_{w \rightarrow w} \xrightarrow{N \rightarrow \infty} 0 \quad \text{in case (i),} \quad (3.73)$$

$$\|c \beta_{h_1} \cdots \beta_{h_N} I_\Lambda\|_w \xrightarrow{N \rightarrow \infty} 0 \quad \text{in case (ii).} \quad (3.74)$$

Now, since the  $h_i$  are bounded, (3.70) is equivalent to the existence of  $\delta > 0$  and a sequence  $0 = n_0 < n_1 < n_2 < \dots$  satisfying

$$\sum_{i=n_{j-1}+1}^{n_j} h_i \geq \delta \chi_\Lambda \quad (3.75)$$

for all  $j$ . This, in turn, is equivalent to the existence of  $\varepsilon > 0$  such that

$$1 - \prod_{i=n_{j-1}+1}^{n_j} (1 - h_i) \geq \varepsilon \chi_\Lambda \quad (3.76)$$

for every  $j$  (indeed, we can set  $\varepsilon = 1 - \exp(-\delta)$ ). Therefore, by Corollary 3.2, if  $N \geq n_k$  we have

$$\|\beta_{h_1} \cdots \beta_{h_N} I_\Lambda\|_{w \rightarrow w} \leq \|\beta_{\varepsilon \Lambda}^k I_\Lambda\|_{w \rightarrow w}, \quad (3.77)$$

$$\|c\beta_{h_1} \cdots \beta_{h_N} I_\Lambda\|_w \leq \|c\beta_{\varepsilon \Lambda}^k I_\Lambda\|_w. \quad (3.78)$$

The theorem then follows from Lemmas 3.1 and 3.6.  $\square$

*Remark.* The standard “cleaning” proof of the Dobrushin uniqueness theorem [10, 11, 13, 18, 21] proves a very special case of Theorem 3.1(ii): namely, one assumes the very strong hypothesis  $\alpha w \leq (1 - \varepsilon)w$  for some  $\varepsilon > 0$ , and one takes  $h_i = \delta_{x_i}$ , where  $x_1, x_2, \dots$  is a sequence that visits each site of  $\Lambda$  infinitely many times. The correlations between  $\Lambda$  and  $\Lambda^c$  can then be bounded in terms of  $\Pi_\Lambda$  (see [10, 11, 18] for variants of this idea).

Let us also remark that, in the application to the Dobrushin uniqueness theorem, it appears to be necessary to take  $w = \mathbf{1}$ . This choice plays no role in the “cleaning” proof itself, but plays a role in the final step of the argument, where the total oscillation of a function of many variables is bounded by the sum of its single-variable oscillations — *not* a weighted sum.

The non-uniform (dust-dependent) version of the previous result relies on the following decomposition:

**Lemma 3.7.** *Fix  $\varepsilon > 0$  and a set  $\Lambda \subseteq X$ . Then, for all integers  $N \geq K \geq 0$ ,*

$$\beta_{\varepsilon \Lambda}^N = \beta_{\varepsilon \Lambda}^K I_{\Lambda^c} + (I_\Lambda \beta_{\varepsilon \Lambda} I_\Lambda)^K I_\Lambda \beta_{\varepsilon \Lambda}^{N-K}. \quad (3.79)$$

*Proof.* By (3.14) we have

$$I_{\Lambda^c} \beta_{\varepsilon \Lambda}^N = I_{\Lambda^c} \quad (3.80)$$

and hence

$$\beta_{\varepsilon \Lambda}^N = I_{\Lambda^c} + I_\Lambda \beta_{\varepsilon \Lambda}^N. \quad (3.81)$$

It follows that

$$\begin{aligned}
 \beta_{\varepsilon\Lambda}^N &= \beta_{\varepsilon\Lambda}^K \beta_{\varepsilon\Lambda}^{N-K} \\
 &= \beta_{\varepsilon\Lambda}^K (I_{\Lambda^c} + I_{\Lambda} \beta_{\varepsilon\Lambda}^{N-K}) && \text{[by (3.81)]} \\
 &= \beta_{\varepsilon\Lambda}^K I_{\Lambda^c} + (I_{\Lambda} \beta_{\varepsilon\Lambda} I_{\Lambda})^K I_{\Lambda} \beta_{\varepsilon\Lambda}^{N-K} && \text{[by (3.12)].}
 \end{aligned} \tag{3.82}$$

□

**Theorem 3.2 (Dust-dependent cleaning).** (FH) Consider a region  $\Lambda \subseteq X$  and functions  $0 \leq h_i \leq \chi_{\Lambda}$  ( $i \geq 1$ ) such that

$$\sum_{i=1}^{\infty} h_i(x) = \infty \quad \text{for all } x \in \Lambda. \tag{3.83}$$

If the nonnegative vector  $c \in l^1(w)$  is such that  $\rho_{\Lambda}(\ell; c) \equiv \|c(I_{\Lambda} \alpha I_{\Lambda})^{\ell}\|_w \rightarrow 0$  as  $\ell \rightarrow \infty$ , then

$$\|c(\beta_{h_1} \cdots \beta_{h_N} - \Pi_{\Lambda})\|_w \xrightarrow{N \rightarrow \infty} 0. \tag{3.84}$$

*Proof.* As in the proof of Theorem 3.1, it is enough to show that

$$\|c\beta_{h_1} \cdots \beta_{h_N} I_{\Lambda}\|_w \xrightarrow{N \rightarrow \infty} 0. \tag{3.85}$$

Fix  $\delta > 0$  and  $0 < \varepsilon < 1$ . By Lemma 3.6 we can choose  $K$  so that

$$\|c(I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^K\|_w \leq \delta/2 \tag{3.86}$$

uniformly for all regions  $\Lambda' \subseteq \Lambda$ . On the other hand, since  $c \in l^1(w)$ , we can choose a finite set  $\Lambda' \subseteq \Lambda$  so that

$$\|c(I_{\Lambda} \alpha I_{\Lambda})^{\ell} I_{\Lambda \setminus \Lambda'}\|_w \leq \delta/[2(K+1)] \tag{3.87}$$

for  $\ell = 0, 1, \dots, K$ . Then hypothesis (3.83) guarantees that there exists a sequence of integers  $0 = n_0 < n_1 < n_2 < \dots$  such that

$$1 - \prod_{i=n_{j-1}+1}^{n_j} (1 - h_i) \geq \varepsilon \chi_{\Lambda'} \tag{3.88}$$

for all  $j$ . Therefore, by Corollary 3.2, if  $N \geq n_k$  we have

$$\|c\beta_{h_1} \cdots \beta_{h_N} I_{\Lambda}\|_w \leq \|c\beta_{\varepsilon\Lambda'}^k I_{\Lambda}\|_w. \tag{3.89}$$

Now, by the decomposition (3.79), whenever  $k \geq K$  we have

$$\beta_{\varepsilon\Lambda'}^k I_{\Lambda} = \beta_{\varepsilon\Lambda'}^K I_{\Lambda \setminus \Lambda'} + (I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^K I_{\Lambda'} \beta_{\varepsilon\Lambda'}^{k-K} I_{\Lambda}. \tag{3.90}$$

Applying this to the vector  $w$ , we deduce the vector inequalities

$$\begin{aligned}
\beta_{\varepsilon\Lambda'}^k I_\Lambda w &\leq \beta_{\varepsilon\Lambda'}^K I_{\Lambda\setminus\Lambda'} w + (I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^K I_{\Lambda'} w && \text{[by (2.15)]} \\
&\leq \beta_{\Lambda'}^K I_{\Lambda\setminus\Lambda'} w + (I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^K I_{\Lambda'} w && \text{[by (3.27)]} \\
&= \sum_{\ell=0}^K (I_{\Lambda'} \alpha)^\ell I_{\Lambda\setminus\Lambda'} w + (I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^K I_{\Lambda'} w \\
&\leq \sum_{\ell=0}^K (I_\Lambda \alpha I_\Lambda)^\ell I_{\Lambda\setminus\Lambda'} w + (I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^K I_{\Lambda'} w. && (3.91)
\end{aligned}$$

Thus,

$$\|c\beta_{\varepsilon\Lambda'}^k I_\Lambda\|_w \leq \sum_{\ell=0}^K \|c(I_\Lambda \alpha I_\Lambda)^\ell I_{\Lambda\setminus\Lambda'}\|_w + \|c(I_{\Lambda'} \beta_{\varepsilon\Lambda'} I_{\Lambda'})^K\|_w \leq \delta. \quad (3.92)$$

□

Let us observe that, in the absence of some uniformity hypothesis on the  $h_i$  [like (3.70)], the convergence in (3.84) can be *arbitrarily slow*, even if  $\rho_\Lambda(\ell; c) \rightarrow 0$  arbitrarily rapidly:

**Example 3.1.** Let  $\Lambda$  be countably infinite (say,  $\Lambda = \{1, 2, 3, \dots\}$ ); let  $w = \mathbf{1}$  and choose any  $c > 0$  with  $c \in l^1$ . Let us consider the *best possible case* for cleaning, namely  $\alpha = 0$ , so that  $\rho_\Lambda(\ell; c) = 0$  for all  $\ell \geq 1$ . Then  $\|c(\beta_h^N - \Pi_\Lambda)\|_w = \sum_{i=1}^\infty c_i [1 - h(i)]^N =: \varepsilon_N$ . It is a fairly simple analysis exercise to show that the  $(h_i)_{i \geq 1}$  can be chosen so that  $\varepsilon_N$  decays more slowly with  $N$  than any specified convergent-to-zero sequence  $(\delta_N)_{N \geq 0}$ .<sup>5</sup>

### 3.6. Some further identities and inequalities

Let us now prove a beautiful identity for the sum of a geometric series  $\sum_{N=0}^\infty (I_\Lambda \beta_h I_\Lambda)^N$ . This identity will play a central role in the next subsection in the proof of the converse theorem on cleaning (Theorem 3.3).

<sup>5</sup>Let  $(c_i)_{i \geq 1}$  be any strictly positive sequence, and let  $(\delta_N)_{N \geq 0}$  be any sequence of non-negative numbers converging to zero. We claim that one can choose a sequence  $(h_i)_{i \geq 1}$  of numbers in  $(0, 1]$  such that  $\varepsilon_N := \sum_{i=1}^\infty c_i (1 - h_i)^N \geq \delta_N$  for all but finitely many  $N$ .

*Proof.* Choose  $N_1$  such that  $\delta_N \leq c_1/2$  for all  $N \geq N_1$ ; and for  $i \geq 2$ , inductively choose  $N_i > N_{i-1}$  such that  $\delta_N \leq c_i/2$  for all  $N \geq N_i$ . Then, for each  $i \geq 1$ , choose  $h_i$  small enough so that  $(1 - h_i)^{N_{i+1}} \geq 1/2$ . It follows that, for  $N_i \leq N \leq N_{i+1}$  we have

$$\delta_N \leq c_i/2 \leq c_i(1 - h_i)^{N_{i+1}} \leq c_i(1 - h_i)^N \leq \varepsilon_N.$$

Since every  $N \geq N_1$  belongs to some such interval, we are done. □

**Lemma 3.8.** *Let  $\Lambda \subseteq X$ , and let  $h: X \rightarrow [0, 1]$  be strictly positive on  $\Lambda$  and zero outside  $\Lambda$ . (In other words,  $\text{supp } h = \Lambda$ .) Then*

$$\sum_{N=0}^{\infty} (I_{\Lambda} \beta_h I_{\Lambda})^N I_h = \sum_{k=0}^{\infty} (I_{\Lambda} \alpha I_{\Lambda})^k I_{\Lambda}. \quad (3.93)$$

*Proof.* Everything occurs within  $\Lambda$ , so for notational simplicity let us suppose that  $\Lambda = X$ . Write  $\beta_h = I_{1-h} + I_h \alpha$ , expand out the  $N$ th power, and sum over  $N$ . We get a sum over all finite sequences (including the empty sequence) of factors  $I_{1-h}$  and  $I_h \alpha$ . Now let us treat the matrix elements of  $\alpha$  as noncommuting indeterminates and extract the coefficient of a monomial  $\alpha_{x_0 x_1} \alpha_{x_1 x_2} \cdots \alpha_{x_{k-1} x_k}$  with  $k \geq 0$  (it is easy to see that these are the only monomials that arise). To the left of  $\alpha_{x_0 x_1}$  we have an arbitrary number (including zero) of factors  $1 - h(x_0)$  followed by one factor  $h(x_0)$ : this gives

$$\sum_{n=0}^{\infty} [1 - h(x_0)]^n h(x_0) = 1$$

since  $0 < h(x_0) \leq 1$ . Likewise to the immediate left of each  $\alpha_{x_{i-1} x_i}$ . Finally, to the right of  $\alpha_{x_{k-1} x_k}$  we have an arbitrary number (including zero) of factors  $1 - h(x_k)$ : this gives  $\sum_{n=0}^{\infty} [1 - h(x_k)]^n = 1/h(x_k)$ , and this factor is cancelled by the  $I_h$  on the left-hand side of (3.93). So each monomial  $\alpha_{x_0 x_1} \alpha_{x_1 x_2} \cdots \alpha_{x_{k-1} x_k}$  gets a coefficient 1, which corresponds exactly to the right-hand side of (3.93).  $\square$

*Important Remark.* By treating the matrix elements of  $\alpha$  as noncommuting indeterminates, we are in essence using the tree formalism that will be described in detail in Section 4.

If we use a product  $\beta_{f_1} \cdots \beta_{f_m}$  in place of the single cleaning operator  $\beta_h$ , then we can obtain an *inequality* in place of the identity (3.93):

**Lemma 3.9.** *Let  $\Lambda \subseteq X$ , and let  $f_1, \dots, f_m: X \rightarrow [0, 1]$  be supported on  $\Lambda$ . Define  $h := 1 - \prod_{i=1}^m (1 - f_i)$ . Then*

$$\sum_{k=0}^{\infty} (I_{\Lambda} \beta_{f_1} \cdots \beta_{f_m} I_{\Lambda})^k I_h \leq \sum_{k=0}^{\infty} (I_{\Lambda} \alpha I_{\Lambda})^k I_{\Lambda} \quad (3.94)$$

*provided that the right-hand side is elementwise finite.*

*Proof.* As in the previous lemma, everything here occurs within  $\Lambda$ , so we can assume for notational simplicity that  $\Lambda = X$ . We apply Corollary 3.1 with  $g_1 = \dots = g_m = 0$  and  $h_i = f_i$ , to obtain

$$I - \beta_{f_1} \cdots \beta_{f_m} = \left( \sum_{i=1}^m I_{f_i} \beta_{f_{i+1}}^* \cdots \beta_{f_m}^* \right) (I - \alpha) := P(I - \alpha) \quad (3.95)$$

where  $P$  is a sum of products of the operators  $\alpha, I_{f_j}$  and  $I_{1-f_j}$  ( $1 \leq j \leq m$ ). Furthermore, the term in  $P$  containing no factors of  $\alpha$  is the operator of multiplication by

$$\sum_{i=1}^m f_i \prod_{j=i+1}^m (1 - f_j) = 1 - \prod_{i=1}^m (1 - f_i) = h. \quad (3.96)$$

Since the other terms are nonnegative, we have  $P \geq I_h$ .

Let us now abbreviate  $B = \beta_{f_1} \cdots \beta_{f_m}$ . Under the assumption that  $\sum_{k=0}^{\infty} \alpha^k$  is elementwise finite, we have

$$(I - \alpha) \left( \sum_{k=0}^{\infty} \alpha^k \right) = I. \quad (3.97)$$

Therefore, right-multiplying (3.95) by  $\sum_{k=0}^{\infty} \alpha^k$  yields

$$I_h \leq P = (I - B) \sum_{k=0}^{\infty} \alpha^k. \quad (3.98)$$

We now left-multiply this inequality by  $B^k$  and sum from  $k = 0$  to  $N$ : since  $B^k(I - B) = B^k - B^{k+1}$ , the sum telescopes and we have

$$\sum_{k=0}^N B^k I_h \leq (I - B^{N+1}) \sum_{k=0}^{\infty} \alpha^k \leq \sum_{k=0}^{\infty} \alpha^k. \quad (3.99)$$

Taking  $N \rightarrow \infty$  gives the result.  $\square$

Lemma 3.9 is a special case of a result to be proven in Section 4.9 using the tree formalism [see Lemmas 4.13(b) and 4.14(a)]. Indeed, the mysterious operator  $P$  in (3.95) will correspond to the cloud  $\mu$  in Lemma 4.14(a). Furthermore, in the tree context the summability condition on  $\sum_{k=0}^{\infty} \alpha^k$  can be removed.

Finally, we have a reverse inequality:

**Lemma 3.10.** *Let  $\Lambda \subseteq X$ , and let  $f_1, \dots, f_m: X \rightarrow [0, 1]$  be supported on  $\Lambda$ . Then*

$$\sum_{k=0}^{\infty} (I_{\Lambda} \alpha I_{\Lambda})^k \leq \left( \sum_{n=0}^{\infty} (I_{\Lambda} \beta_{f_1} \cdots \beta_{f_m} I_{\Lambda})^n \right) \left( \sum_{k=0}^{m-1} (I_{\Lambda} \alpha I_{\Lambda})^k \right) \quad (3.100)$$

*provided that  $\sum_{n=0}^{\infty} (I_{\Lambda} \beta_{f_1} \cdots \beta_{f_m} I_{\Lambda})^n$  is elementwise finite.*

The main tool in the proof is the following bound, which uses only the nonnegativity of the matrix  $\alpha$ :



**Lemma 3.11.** *Let  $f_1, \dots, f_m: X \rightarrow [0, 1]$ . Then*

$$0 \leq \sum_{i=1}^m I_{f_i} \beta_{f_{i+1}}^* \cdots \beta_{f_m}^* \leq \sum_{k=0}^{m-1} \alpha^k. \quad (3.101)$$

*Proof.* Write  $\beta_f^* = I_{1-f} + \alpha I_f$  and expand out the left-hand side of (3.101). Let us once again treat the matrix elements of  $\alpha$  as noncommuting indeterminates and extract the coefficient of a monomial  $\alpha_{x_0 x_1} \alpha_{x_1 x_2} \cdots \alpha_{x_{k-1} x_k}$  with  $0 \leq k \leq m-1$  (it is easy to see that these are the only monomials that arise). We need to show that each such coefficient is  $\leq 1$  (the nonnegativity is obvious). We have already computed in (3.96) the term with no powers of  $\alpha$ : it is  $1 - \prod_{i=1}^m [1 - f_i(x_0)] \leq 1$ . Now suppose that there are  $k$   $\alpha$ 's, occurring at positions  $j_1, \dots, j_k$  with  $i+1 \leq j_1 < j_2 < \dots < j_k \leq m$ . The coefficient of such a term will be

$$\begin{aligned} & [f_i(1 - f_{i+1}) \cdots (1 - f_{j_1-1})](x_0) \times [f_{j_1}(1 - f_{j_1+1}) \cdots (1 - f_{j_2-1})](x_1) \\ & \times \cdots \times [f_{j_k}(1 - f_{j_k+1}) \cdots (1 - f_m)](x_k) \end{aligned} \quad (3.102)$$

and we then need to sum over all choices of indices  $1 \leq i < j_1 < j_2 < \dots < j_k \leq m$ . First fix  $j_1, \dots, j_k$  and sum over  $i$ : one gets

$$\sum_{i=1}^{j_1-1} [f_i(1 - f_{i+1}) \cdots (1 - f_{j_1-1})](x_0) = 1 - \prod_{i=1}^{j_1-1} [1 - f_i(x_0)] \leq 1 \quad (3.103)$$

[just as in (3.96)]. We then sum over  $j_1$  and so forth, each time bounding the sum by 1.  $\square$

*Proof of Lemma 3.10.* Once again, we can assume that  $\Lambda = X$ . Let us abbreviate  $B = \beta_{f_1} \cdots \beta_{f_m}$ . From (3.95) we have

$$I - B = P(I - \alpha) \quad (3.104)$$

where Lemma 3.11 gives

$$0 \leq P \leq \sum_{k=0}^{m-1} \alpha^k. \quad (3.105)$$

Under the assumption that  $\sum_{n=0}^{\infty} B^n$  is elementwise finite, we have

$$\left( \sum_{n=0}^{\infty} B^n \right) (I - B) = I, \quad (3.106)$$

so we can left-multiply (3.104) by  $\sum_{n=0}^{\infty} B^n$  to obtain

$$I = \left( \sum_{n=0}^{\infty} B^n \right) P(I - \alpha). \quad (3.107)$$

Now right-multiply this inequality by  $\alpha^k$  and sum from  $k = 0$  to  $N$ : the sum on the right telescopes and we obtain

$$\sum_{k=0}^N \alpha^k = \left( \sum_{n=0}^{\infty} B^n \right) P(I - \alpha^{N+1}) \leq \left( \sum_{n=0}^{\infty} B^n \right) P \leq \left( \sum_{n=0}^{\infty} B^n \right) \left( \sum_{k=0}^{m-1} \alpha^k \right). \quad (3.108)$$

Taking  $N \rightarrow \infty$  gives the result.  $\square$

Lemma 3.10 is a special case of a result to be proven in Section 4.9 using the tree formalism [see Lemma 4.14(c,d)], where moreover the summability condition on  $\sum_{n=0}^{\infty} (I_{\Lambda} \beta_{f_1} \cdots \beta_{f_m} I_{\Lambda})^n$  can be removed. The pair of inequalities (3.94), (3.100) will play a crucial role in the proof of the converse theorem on cleaning (Theorem 3.3), just as their tree generalizations will do in the proof of the strong form of this result (Theorem 6.1).

### 3.7. Converse results

*In this subsection we do **not** assume the Fundamental Hypothesis.* Rather, our goal is to study what happens in case the Fundamental Hypothesis fails. Here we are entitled to use the algebraic *identities* that were proven in the preceding subsections, since such identities are valid irrespective of the Fundamental Hypothesis. Furthermore, we are entitled to use those *inequalities that do not refer to the vector  $w$* , since they are based only on the nonnegativity of the matrix elements of  $\alpha$ . But we must be very careful to avoid using any result that depends on the Fundamental Hypothesis.

Our main result is the following:

**Theorem 3.3.** *Let  $X$  be a finite or countably infinite set, let  $\Lambda \subseteq X$ , and let  $c \geq 0$  and  $w \geq 0$  be vectors that are strictly positive on  $\Lambda$ . Consider the following conditions on a matrix  $\alpha$ :*

- (a)  $\sum_{k=0}^{\infty} c(I_{\Lambda} \alpha I_{\Lambda})^k w < \infty$ .
- (b) For all  $h: X \rightarrow [0, 1]$  with  $\text{supp } h = \Lambda$  such that  $h \geq \varepsilon \chi_{\Lambda}$  for some  $\varepsilon > 0$ , we have  $\sum_{k=0}^{\infty} c(I_{\Lambda} \beta_h I_{\Lambda})^k w < \infty$ .
- (b') There exists  $h: X \rightarrow [0, 1]$  with  $\text{supp } h = \Lambda$  such that  $\sum_{k=0}^{\infty} c(I_{\Lambda} \beta_h \times I_{\Lambda})^k w < \infty$ .
- (c) For every finite sequence  $f_1, \dots, f_m$  of functions  $X \rightarrow [0, 1]$  with  $\text{supp}(f_i) \subseteq \Lambda$  such that  $\sum_i f_i \geq \varepsilon \chi_{\Lambda}$  for some  $\varepsilon > 0$ , we have  $\sum_{k=0}^{\infty} c(I_{\Lambda} \beta_{f_1} \cdots \beta_{f_m} \times I_{\Lambda})^k w < \infty$ .
- (c') There exists a finite sequence  $f_1, \dots, f_m$  of functions  $X \rightarrow [0, 1]$  with  $\text{supp}(f_i) \subseteq \Lambda$  such that  $\sum_{k=0}^{\infty} c(I_{\Lambda} \beta_{f_1} \cdots \beta_{f_m} I_{\Lambda})^k w < \infty$ .

Then (a)–(c) are all equivalent and imply (c'); and for matrices  $\alpha$  satisfying the additional hypothesis

$$\text{There exists a constant } C < \infty \text{ such that } I_\Lambda \alpha I_\Lambda w \leq Cw \quad (3.109)$$

all five conditions are equivalent.

*Proof.* (b)  $\implies$  (b') is trivial, while (b')  $\implies$  (a)  $\implies$  (b) are immediate consequences of Lemma 3.8.

(a)  $\implies$  (c) is an immediate consequence of Lemma 3.9, and (c)  $\implies$  (b) is trivial, as is (c)  $\implies$  (c').

Finally, Lemma 3.10 entails (c')  $\implies$  (a) under the hypothesis (3.109).  $\square$

Please note that the hypothesis (3.109) is automatic whenever  $\Lambda$  is a *finite* set. On the other hand, the following two examples show that, when  $\Lambda$  is infinite, hypothesis (3.109) cannot be dispensed with in proving that (c')  $\implies$  (a):

**Example 3.2.** Let  $X = \{0, 1, 2, 3, \dots\}$ ; set  $\alpha_{0j} = 1$  for all  $j \geq 1$ , and set all other matrix elements of  $\alpha$  to 0. Set  $w = \mathbf{1}$  and let  $c$  be any strictly positive vector in  $l^1$ . Now let  $\Lambda$  be any infinite subset of  $X$  containing 0. We have

$$c(I_\Lambda \alpha I_\Lambda)^k w = \begin{cases} \|c\|_1 \in (0, \infty) & \text{for } k = 0, \\ +\infty & \text{for } k = 1, \\ 0 & \text{for } k \geq 2, \end{cases} \quad (3.110)$$

so that  $\sum_{k=0}^{\infty} c(I_\Lambda \alpha I_\Lambda)^k w = +\infty$ . On the other hand, if we take  $f_1 = \dots = f_m = \chi_\Lambda$  for any  $m \geq 2$ , we have  $\sum_{k=0}^{\infty} c(I_\Lambda \beta_{f_1} \cdots \beta_{f_m} I_\Lambda)^k w = \sum_{k=0}^{\infty} c(I_\Lambda \alpha \times I_\Lambda)^{mk} w = \|c\|_1 < \infty$ . Indeed, we have  $\beta_\Lambda^\ell I_\Lambda = 0$  for all  $\ell \geq 2$ .

**Example 3.3.** In the preceding example, one of the components of the vector  $I_\Lambda \alpha I_\Lambda w$  was  $+\infty$ . Here is a variant in which the vector  $I_\Lambda \alpha I_\Lambda w$  is pointwise finite but is not bounded by any multiple of  $w$ . Take  $\Lambda = X = \{x_1, x_2, x_3, \dots\} \cup \{y_1, y_2, y_3, \dots\}$ ; set  $\alpha_{x_i y_j} = 1$  if  $j \leq i$ , and set all other matrix elements of  $\alpha$  to 0. Set  $w = \mathbf{1}$  and  $c_{x_i} = c_{y_i} = 1/i^2$ . Then  $(\alpha w)_{x_i} = i$ ,  $(\alpha w)_{y_i} = 0$  and  $\alpha^k w = 0$  for  $k \geq 2$ ; so (3.110) again holds and the same choice of  $f_1, \dots, f_m$  provides a counterexample.

We will return to these questions in Section 6, where we will prove a significant extension of Theorem 3.3 (see Theorem 6.1).

#### 4. Tree approach to balayage

In this section we introduce an alternate approach to studying the algebra of operators generated by  $\alpha$  and the multiplication operators  $I_f$ , which brings

out more clearly its underlying combinatorial structure and which permits a far-reaching generalization of the results obtained in the preceding section. This approach is based on considering the tree of finite sequences of elements of  $X$ <sup>6</sup>. Another way of phrasing matters is to say that we are working in the algebra of formal power series in noncommuting indeterminates  $\{\alpha_{xy}\}_{x,y \in X}$  subject to the relations  $\alpha_{xy}\alpha_{uv} = 0$  whenever  $y \neq u$ .

The plan of this section is as follows: In Section 4.1 we introduce the formalism of “markers” and “clouds”, and we analyze its relation with the operator formalism of Section 3. In particular, we introduce the key operation of convolution of clouds, which corresponds to multiplication of operators. In Section 4.2 we define clouds to represent each of the special operators  $\alpha$ ,  $I_f$ ,  $\beta_f$  and  $\Pi_\Lambda$ . In Section 4.3 we introduce a very important partial ordering  $\leq$  on clouds, which formalizes (roughly speaking) the comparison of operators by their “efficiency of cleaning”. We study the circumstances under which the partial ordering  $\leq$  is preserved by convolution from the left or the right, and we introduce several important subclasses of clouds ( $\mathcal{B} \subsetneq \mathcal{P} \subsetneq \mathcal{R} \subsetneq \mathcal{S}$ ). In Section 4.4 we prove a fundamental comparison inequality, which substantiates our assertion that the partial ordering  $\leq$  is related to the efficiency of cleaning. In Section 4.5 we introduce the notion of a cloud being “carried” by a subset  $\Lambda$ , and in Section 4.6 we introduce the stricter notion of being “ $\Lambda$ -regular”. In Section 4.7 we show that the cloud  $\pi_\Lambda$  (which is associated to the balayage operator  $\Pi_\Lambda$ ) plays a special role among  $\Lambda$ -regular clouds, by virtue of its minimality with respect to  $\leq$ . In Section 4.8 we put all these tools together, and study the convergence of cleaning operators  $\beta_{h_1} \cdots \beta_{h_n}$  to  $\Pi_\Lambda$ . In the cloud context we can shed additional light on this convergence, by distinguishing convergence of clouds from convergence of the corresponding operators. In Section 4.9 we prove some further identities and inequalities that will play a crucial role in the converse results of Section 6.

#### 4.1. Markers, clouds and operators

A nonempty finite sequence  $\eta = (x_0, x_1, \dots, x_k)$  of elements of  $X$  will be called a *marker*. We denote by  $X^{[\infty]} = \bigcup_{k=0}^{\infty} X^{k+1}$  the set of all markers. Given a marker  $\eta = (x_0, x_1, \dots, x_k)$ , we define

$$\text{level}(\eta) := k, \tag{4.1}$$

$$\text{first}(\eta) := x_0, \tag{4.2}$$

$$\text{last}(\eta) := x_k, \tag{4.3}$$

$$\eta_i^j := (x_i, x_{i+1}, \dots, x_j). \tag{4.4}$$

An *ancestor* (or *prefix*) of  $\eta$  is any one of the markers  $\eta_0^j$  ( $0 \leq j \leq k$ ); we write  $\eta' \preceq \eta$  to denote that  $\eta'$  is an ancestor of  $\eta$ . We write  $\eta' < \eta$  to denote that  $\eta' \preceq \eta$

<sup>6</sup>We use the word “tree” even though it turns out to be convenient, in our approach, to suppress the root of the tree (that is, the empty sequence).

and  $\eta' \neq \eta$ . A *child* of  $\eta$  is any marker of the form  $(x_0, x_1, \dots, x_k, x_{k+1})$  for some  $x_{k+1} \in X$ . A *suffix* of  $\eta$  is any one of the markers  $\eta_j^k$  ( $0 \leq j \leq k$ ). A *subsequence* of  $\eta = (x_0, x_1, \dots, x_k)$  is any marker of the form  $\eta' = (x_{j_0}, x_{j_1}, \dots, x_{j_l})$  for some choice of indices  $0 \leq j_0 < j_1 < \dots < j_l \leq k$ .

A *cloud*  $\nu = (\nu_\eta)$  is a real-valued function on the set  $X^{[\infty]}$  of markers. [We shall sometimes write  $\nu(\eta)$  as a synonym for  $\nu_\eta$ .] We say that  $\nu$  has *finite support* if  $\nu_\eta = 0$  for all but finitely many markers  $\eta$ . We say that  $\nu$  is *supported on levels*  $\leq N$  if  $\nu_\eta = 0$  whenever  $\text{level}(\eta) > N$ . We define

$$\|\nu\| := \sup_{\eta} \sum_{\sigma \preceq \eta} |\nu_\sigma|, \quad (4.5)$$

and we say that  $\nu$  has *finite norm* if  $\|\nu\| < \infty$ . The set of clouds of finite norm forms a Banach space with the norm  $\|\cdot\|$ .

Clouds allow us to give an abstract combinatorial representation of the algebra of operators generated by  $\alpha$  and the multiplication operators, independently of any particular choice of the matrix  $\alpha$ . To see this, let us associate to the marker  $\eta = (x_0, x_1, \dots, x_k)$  the operator

$$T_\eta := I_{\{x_0\}} \alpha I_{\{x_1\}} \alpha \cdots \alpha I_{\{x_{k-1}\}} \alpha I_{\{x_k\}}. \quad (4.6)$$

[The level of a marker thus corresponds to the number of factors  $\alpha$  in the corresponding operator. In physical terms, a marker  $\eta = (x_0, x_1, \dots, x_k)$  represents a piece of dirt that has traveled from  $x_0$  to  $x_k$  via the path  $x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_k$ .] More generally, to a cloud  $\nu$  we associate the operator  $T_\nu$  defined by

$$T_\nu := \sum_{\eta} \nu_\eta T_\eta. \quad (4.7)$$

(Initially this formula makes sense only for clouds of finite support, but we will soon extend the definition to clouds of finite norm.) Conversely, given any finite sum of operators of the form  $I_{f_0} \alpha I_{f_1} \cdots I_{f_{k-1}} \alpha I_{f_k}$ , we can expand it out as a (possibly infinite) sum of terms  $I_{\{x_0\}} \alpha I_{\{x_1\}} \cdots I_{\{x_{k-1}\}} \alpha I_{\{x_k\}}$  and thus represent it in the form  $T_\nu$  for some cloud  $\nu$  supported on a finite number of levels; furthermore,  $\nu$  has finite norm if the functions  $f_i$  are bounded. Of course, to make these considerations precise we will have to deal with possibly infinite sums and specify the exact classes of clouds and operators under consideration.

**Proposition 4.1 (Extension of  $T_\nu$ ).** (FH) *Let  $\nu$  be a cloud of finite support. Then*

$$\|T_\nu\|_{w \rightarrow w} \leq \|\nu\|. \quad (4.8)$$

*Therefore, the definition of the operator  $T_\nu$  can be extended by linearity and continuity to all clouds of finite norm, and the map  $\nu \mapsto T_\nu$  is a contraction.*

*Proof.* Because  $\alpha$  has nonnegative entries, it suffices to consider the case in which  $\nu_\eta \geq 0$  for all  $\eta$ . We assume that the cloud  $\nu$  is supported on levels  $\leq N$ , and we shall prove the proposition by induction on  $N$ . If  $N = 0$ , (4.8) is straightforward. Suppose therefore that  $N > 0$ , and consider a vector  $c = (c_x)_{x \in X}$ . (Again, we can suppose  $c_x \geq 0$  for all  $x \in X$ .) We have

$$\|cT_\nu\|_w = \sum_{k=0}^N \sum_{\eta=(x_0, \dots, x_k)} \nu_\eta c_{x_0} \alpha_{x_0 x_1} \alpha_{x_1 x_2} \cdots \alpha_{x_{k-1} x_k} w_{x_k}. \quad (4.9)$$

In this sum, the contribution of the markers  $\eta$  such that  $\text{level}(\eta) \in \{N-1, N\}$  is

$$\sum_{\eta=(x_0, \dots, x_{N-1})} c_{x_0} \alpha_{x_0 x_1} \cdots \alpha_{x_{N-2} x_{N-1}} \left( \nu_\eta w_{x_{N-1}} + \sum_{x_N \in X} \nu_{\eta \circ x_N} \alpha_{x_{N-1} x_N} w_{x_N} \right) \quad (4.10)$$

where  $\eta \circ x_N$  denotes the marker obtained by appending the element  $x_N$  to  $\eta$ . Using the Fundamental Hypothesis in the form

$$\sum_{x_N \in X} \alpha_{x_{N-1} x_N} w_{x_N} \leq w_{x_{N-1}}, \quad (4.11)$$

the contribution (4.10) can be bounded by

$$\sum_{\eta=(x_0, \dots, x_{N-1})} c_{x_0} \alpha_{x_0 x_1} \cdots \alpha_{x_{N-2} x_{N-1}} w_{x_{N-1}} \left( \nu_\eta + \sup_{x_N \in X} \nu_{\eta \circ x_N} \right). \quad (4.12)$$

Thus, we obtain

$$\|cT_\nu\|_w \leq \|cT_{\nu'}\|_w \quad (4.13)$$

where  $\nu'$  is the cloud defined by

$$\nu'_\eta = \begin{cases} 0 & \text{if level}(\eta) = N, \\ \nu_\eta + \sup_{x_N \in X} \nu_{\eta \circ x_N} & \text{if level}(\eta) = N-1, \\ \nu_\eta & \text{otherwise.} \end{cases} \quad (4.14)$$

The cloud  $\nu'$  is supported on levels  $\leq N-1$  and satisfies  $\|\nu'\| = \|\nu\|$ ; this completes the proof.  $\square$

Since our goal is to represent combinatorially the algebra of operators generated by  $\alpha$  and the  $I_f$ , we need to introduce an operation on clouds that corresponds to the multiplication of operators. We do this as follows: Given two markers  $\eta$  and  $\eta'$ , we say that  $\eta$  *leads to*  $\eta'$  (and write  $\eta \rightarrow \eta'$ ) in case  $\text{last}(\eta) = \text{first}(\eta')$ . If  $\eta \rightarrow \eta'$ , we define  $\eta * \eta'$  to be the concatenation of  $\eta$  with  $\eta'$  with the proviso that the element  $\text{last}(\eta) = \text{first}(\eta')$  is not repeated. That is, if  $\eta = (x_0, x_1, \dots, x_k)$  and  $\eta' = (y_0, y_1, \dots, y_{k'})$  with  $x_k = y_0$ , then

$\eta * \eta' := (x_0, x_1, \dots, x_k, y_1, \dots, y_{k'})$ . Note that  $\text{level}(\eta * \eta') = \text{level}(\eta) + \text{level}(\eta')$ . Finally, given clouds  $\nu_1$  and  $\nu_2$ , we define the *convolution*  $\nu_1 * \nu_2$  by

$$(\nu_1 * \nu_2)(\eta) = \sum_{\substack{\eta_1 \rightarrow \eta_2 \\ \eta_1 * \eta_2 = \eta}} \nu_1(\eta_1) \nu_2(\eta_2). \quad (4.15)$$

(Note that this convolution is associative but non-commutative.) Since the sum (4.15) is finite for each  $\eta$ , convolution is well-defined for arbitrary pairs of clouds  $\nu_1, \nu_2$ . Moreover, if  $\nu_1$  and  $\nu_2$  are of finite support (resp. supported on finitely many levels), then so is  $\nu_1 * \nu_2$ . Furthermore, we have:

**Proposition 4.2 (Norm boundedness of convolution).** *Let  $\nu_1$  and  $\nu_2$  be clouds of finite norm. Then*

$$\|\nu_1 * \nu_2\| \leq \|\nu_1\| \|\nu_2\|. \quad (4.16)$$

*Proof.* For any marker  $\eta$  of level  $k$ , we have

$$\begin{aligned} \sum_{\sigma \preceq \eta} |(\nu_1 * \nu_2)(\sigma)| &= \sum_{j=0}^k |(\nu_1 * \nu_2)(\eta_0^j)| = \sum_{j=0}^k \left| \sum_{i=0}^j \nu_1(\eta_0^i) \nu_2(\eta_i^j) \right| \\ &\leq \sum_{i=0}^k \sum_{j=i}^k |\nu_1(\eta_0^i)| |\nu_2(\eta_i^j)| = \sum_{i=0}^k |\nu_1(\eta_0^i)| \sum_{\sigma \preceq \eta_i^k} |\nu_2(\sigma)| \\ &\leq \sum_{i=0}^k |\nu_1(\eta_0^i)| \|\nu_2\| \leq \|\nu_1\| \|\nu_2\|. \end{aligned} \quad (4.17)$$

□

It is now easily verified that convolution of clouds corresponds to multiplication of operators, i.e.

$$T_{\nu_1 * \nu_2} = T_{\nu_1} T_{\nu_2} \quad (4.18)$$

whenever  $\nu_1, \nu_2$  are clouds of finite norm. Indeed, this formula is easily seen to be true for clouds of finite support; it then extends by continuity to clouds of finite norm as a consequence of Propositions 4.1 and 4.2. Of course, the identity (4.18) is no accident: we defined convolution so that (4.18) would hold! Note in particular the importance of the “leads to” relation and the non-repetition of last ( $\eta$ ) = first ( $\eta'$ ): these implement the identity  $I_{\{x\}} I_{\{y\}} = \delta_{xy} I_{\{x\}}$ .

## 4.2. Some special clouds

Let us now define clouds to represent each of the special operators introduced in Section 2.1.

Level indicators. For each  $k \geq 0$ , we define a cloud  $\rho^k$  by

$$(\rho^k)_\eta = \begin{cases} 1 & \text{if level}(\eta) = k, \\ 0 & \text{otherwise.} \end{cases} \quad (4.19)$$

It is immediate that  $\|\rho^k\| = 1$  and that  $T_{\rho^k} = \alpha^k$ ; in particular,  $T_{\rho^0} = I$ . Note also that  $\rho^0$  is the two-sided identity for convolution:  $\rho^0 * \nu = \nu * \rho^0 = \nu$  for every cloud  $\nu$ . Furthermore,  $\rho^k * \rho^\ell = \rho^{k+\ell}$ .

For any cloud  $\nu$ , we shall denote its convolution powers by  $\nu^{*n}$  (or simply  $\nu^n$ ), with the convention that  $\nu^{*0} = \rho^0$ . In particular, we have  $(\rho^k)^{* \ell} = \rho^{k\ell}$  for all  $k, \ell \geq 0$ .

Clouds associated to multiplication operators. To each function  $f: X \rightarrow \mathbb{R}$ , we associate a cloud that we shall call (by slight abuse of notation)  $I_f$ :

$$(I_f)_{(x_0, \dots, x_k)} = \begin{cases} f(x_0) & \text{if } k = 0, \\ 0 & \text{if } k \geq 1. \end{cases} \quad (4.20)$$

We also write  $I_\Lambda$  as a shorthand for  $I_{\chi_\Lambda}$ . It is easy to verify that  $\|I_f\| = \|f\|_\infty$  and  $T_{I_f} = I_f$ .

Clouds associated to cleaning operators. To each function  $f: X \rightarrow [0, 1]$ , we associate a cloud that we shall call (by slight abuse of notation)  $\beta_f$ :

$$(\beta_f)_{(x_0, \dots, x_k)} = \begin{cases} 1 - f(x_0) & \text{if } k = 0, \\ f(x_0) & \text{if } k = 1, \\ 0 & \text{if } k \geq 2. \end{cases} \quad (4.21)$$

We also write  $\beta_\Lambda$  as a shorthand for  $\beta_{\chi_\Lambda}$ . It is easy to verify that  $\|\beta_f\| = 1$ ,  $\beta_f = I_{1-f} + I_f * \rho^1$  and  $T_{\beta_f} = \beta_f$ .

Clouds associated to balayage operators. For each  $\Lambda \subseteq X$ , we denote by  $\partial\Lambda$  the set of markers of the form  $\eta = (x_0, \dots, x_k)$  [ $k \geq 0$ ] for which  $x_0, \dots, x_{k-1} \in \Lambda$  and  $x_k \in \Lambda^c$ . We then define a cloud  $\pi_\Lambda$  by

$$(\pi_\Lambda)_\eta = \begin{cases} 1 & \text{if } \eta \in \partial\Lambda, \\ 0 & \text{otherwise.} \end{cases} \quad (4.22)$$

It is easy to verify that  $\|\pi_\Lambda\| = 1$  whenever  $\Lambda \neq X$ ,  $\pi_X = 0$ , and  $T_{\pi_\Lambda} = \Pi_\Lambda$ . Note also that  $\pi_\Lambda * \pi_\Lambda = \pi_\Lambda$ .

Indicator clouds. We denote by  $\mathbf{1}$  the cloud that takes the value 1 on all markers. More generally, for  $\Lambda \subseteq X$ , we denote by  $\mathbf{1}_\Lambda$  the cloud that takes the values

$$(\mathbf{1}_\Lambda)_\eta = \begin{cases} 1 & \text{if } \eta \text{ has all its entries in } \Lambda, \\ 0 & \text{otherwise.} \end{cases} \quad (4.23)$$



Formally we have

$$T_{\mathbf{1}} = \sum_{k=0}^{\infty} \alpha^k = (I - \alpha)^{-1}$$

and

$$T_{\mathbf{1}_\Lambda} = I_\Lambda \sum_{k=0}^{\infty} (I_\Lambda \alpha I_\Lambda)^k I_\Lambda = I_\Lambda (I - I_\Lambda \alpha I_\Lambda)^{-1} I_\Lambda;$$

but since  $\|\mathbf{1}\| = \|\mathbf{1}_\Lambda\| = \infty$  (for  $\Lambda \neq \emptyset$ ), there is no guarantee that these series converge. The advantage of the tree formalism is that it makes sense to speak of the clouds  $\mathbf{1}$  and  $\mathbf{1}_\Lambda$  without worrying about convergence questions.

### 4.3. A partial ordering on clouds

The space of clouds is obviously endowed with the “pointwise” partial ordering defined by

$$\mu \leq \nu \quad \text{if and only if} \quad \mu_\eta \leq \nu_\eta \quad \text{for all } \eta. \quad (4.24)$$

(In particular,  $\nu \geq 0$  means that a cloud is nonnegative.) We would now like to introduce a weaker partial ordering that will play an important role in the sequel; roughly speaking, it formalizes the comparison of operators by their “efficiency of cleaning”.

To each cloud  $\nu$ , we associate another cloud  $\tilde{\nu}$ , called its *cumulative distribution*, defined by

$$\tilde{\nu}_\eta = \sum_{\sigma \preceq \eta} \nu_\sigma \quad (4.25)$$

or in other words by

$$\tilde{\nu} = \nu * \mathbf{1}. \quad (4.26)$$

(Note also that  $\|\nu\| = \|\tilde{\nu}\|_\infty$ .) We then introduce a partial ordering  $\trianglelefteq$  on the space of clouds, defined by:

**Definition 4.1 (Definition of  $\trianglelefteq$ ).**  $\mu \trianglelefteq \nu$  if and only if  $\tilde{\mu} \leq \tilde{\nu}$ , i.e.

$$\mu \trianglelefteq \nu \quad \text{if and only if} \quad \tilde{\mu}_\eta \leq \tilde{\nu}_\eta \quad \text{for all } \eta. \quad (4.27)$$

Obviously  $\mu \leq \nu$  implies  $\mu \trianglelefteq \nu$  but not conversely.

Let us begin by proving two useful formulae for the cumulative distribution of a convolution. If  $\eta = (x_0, \dots, x_k)$  is any marker and  $\sigma = (x_0, \dots, x_j)$  is any ancestor of  $\eta$ , we denote by  $\eta \setminus \sigma$  the marker  $(x_j, \dots, x_k)$ ; it is the unique marker  $\xi$  satisfying  $\sigma * \xi = \eta$ . Also, if  $\eta$  is any marker of level  $k \geq 1$ , we define  $\eta^- = \eta_0^{k-1}$  and  $\eta_- = \eta_1^k$ . If  $\eta$  is a marker of level 0, we define  $\eta^- = \eta_- = \emptyset$  (the empty sequence).

**Lemma 4.1 (Cumulative distribution of a convolution).** *Let  $\mu, \nu$  be clouds. Then*

$$\widetilde{(\mu * \nu)} = \mu * \tilde{\nu} \quad (4.28)$$

or equivalently

$$\widetilde{(\mu * \nu)}(\eta) = \sum_{\sigma \preceq \eta} \mu(\sigma) \tilde{\nu}(\eta \parallel \sigma) \quad (4.29)$$

$$= \sum_{\sigma \preceq \eta} \tilde{\mu}(\sigma) [\tilde{\nu}(\eta \parallel \sigma) - \tilde{\nu}((\eta \parallel \sigma)_-)] \quad (4.30)$$

where in (4.30) we make the convention that  $\tilde{\nu}(\emptyset) = 0$ .

*Proof.* Identity (4.28) is an immediate consequence of the associativity of convolution:

$$\widetilde{(\mu * \nu)} = (\mu * \nu) * \mathbf{1} = \mu * (\nu * \mathbf{1}) = \mu * \tilde{\nu}. \quad (4.31)$$

(4.29) is just a rewriting of this.

We now insert in (4.29) the identity  $\mu(\sigma) = \tilde{\mu}(\sigma) - \tilde{\mu}(\sigma^-)$  with the convention  $\tilde{\mu}(\emptyset) = 0$ . This yields

$$\widetilde{(\mu * \nu)}(\eta) = \sum_{\sigma \preceq \eta} \tilde{\mu}(\sigma) \tilde{\nu}(\eta \parallel \sigma) - \sum_{\substack{\sigma \preceq \eta \\ \text{level}(\sigma) \geq 1}} \tilde{\mu}(\sigma^-) \tilde{\nu}(\eta \parallel \sigma). \quad (4.32)$$

In the second sum on the right, we make the change of variables  $\xi = \sigma^-$  and observe that  $\eta \parallel \sigma = (\eta \parallel \xi)_-$ ; this sum therefore becomes

$$\sum_{\xi \preceq \eta^-} \tilde{\mu}(\xi) \tilde{\nu}((\eta \parallel \xi)_-). \quad (4.33)$$

The term  $\xi = \eta$  can now be adjoined to this sum, thanks to the convention  $\tilde{\nu}(\emptyset) = 0$ . This proves (4.30).  $\square$

In the special case  $\nu = \beta_h$ , we have the following formulae:

**Lemma 4.2 (Cumulative distribution of  $\beta_h$ ).** *Let  $h: X \rightarrow [0, 1]$ . Then*

$$\widetilde{\beta_h} = \mathbf{1} - I_h. \quad (4.34)$$

*Proof.* From the definition (4.21) we easily obtain

$$(\widetilde{\beta_h})_{(x_0, \dots, x_k)} = \begin{cases} 1 - h(x_0) & \text{if } k = 0, \\ 1 & \text{if } k \geq 1, \end{cases} \quad (4.35)$$

which is equivalent to (4.34).  $\square$

**Corollary 4.1. (Cumulative distribution of a convolution with  $\beta_h$ ).**

Let  $\mu$  be a cloud, and let  $h: X \rightarrow [0, 1]$ . Then

$$\widetilde{\mu * \beta_h} = \tilde{\mu} - \mu * I_h \quad (4.36)$$

or equivalently

$$(\widetilde{\mu * \beta_h})(\eta) = \tilde{\mu}(\eta) - h(\text{last}(\eta))\mu(\eta) \quad (4.37)$$

$$= \tilde{\mu}(\eta^-) + [1 - h(\text{last}(\eta))]\mu(\eta) \quad (4.38)$$

$$= h(\text{last}(\eta))\tilde{\mu}(\eta^-) + [1 - h(\text{last}(\eta))]\tilde{\mu}(\eta). \quad (4.39)$$

*Proof.* The formula (4.36) is an immediate consequence of (4.28) and (4.34). The alternate forms (4.37)–(4.39) are trivial rewritings.  $\square$

**Corollary 4.2.** If  $\mu \geq 0$  and  $0 \leq g \leq h \leq 1$ , then  $\mu * \beta_h \trianglelefteq \mu * \beta_g$ . In particular,  $\beta_h \trianglelefteq \beta_g$ .

*Proof.* This is an immediate consequence of (4.38).  $\square$

**Corollary 4.3 (Convolution from the right by  $\beta_h$ ).** Let  $\mu_1, \mu_2$  be clouds and let  $0 \leq h \leq 1$ . If  $\mu_1 \trianglelefteq \mu_2$ , then  $\mu_1 * \beta_h \trianglelefteq \mu_2 * \beta_h$ .

*Proof.* This is an immediate consequence of (4.39).  $\square$

For general clouds  $\nu$ , one can obtain an *inequality* analogous to (4.38):

**Corollary 4.4. (Inequality for cumulative distribution of a convolution).** Let  $\mu, \nu \geq 0$  be clouds with  $\|\nu\| < \infty$ . Then, for every marker  $\eta$ ,

$$(\widetilde{\mu * \nu})(\eta) \leq \tilde{\mu}(\eta^-)\|\nu\| + \mu(\eta)\nu(\text{last}(\eta)). \quad (4.40)$$

*Proof.* Rewrite (4.29) as

$$(\widetilde{\mu * \nu})(\eta) = \sum_{\sigma \prec \eta} \mu(\sigma)\tilde{\nu}(\eta \setminus \sigma) + \mu(\eta)\nu(\text{last}(\eta)). \quad (4.41)$$

Using  $\tilde{\nu}(\eta \setminus \sigma) \leq \|\nu\|$  in the first term, we obtain (4.40).  $\square$

See Lemma 4.13 below for an interesting application of Corollary 4.4.

It is also useful to know under what conditions a cloud  $\mu$  is majorized by a cloud  $\beta_h$ . The easy answer is the following:

**Lemma 4.3 (Majorization by  $\beta_h$ ).** Let  $\mu \geq 0$  be a cloud and let  $h: X \rightarrow [0, 1]$ . Then the following are equivalent:

- (a)  $\mu \preceq \beta_h$ .
- (b)  $\|\mu\| \leq 1$  and  $\mu((x)) \leq 1 - h(x)$  for every marker  $(x)$  of level 0.

*Proof.*  $\mu \preceq \beta_h$  means that  $\tilde{\mu}(\eta) \leq \tilde{\beta}_h(\eta)$  for every marker  $\eta$ . By (4.35), this means precisely that  $\mu((x)) \leq 1 - h(x)$  for every  $x \in X$  and that  $\tilde{\mu}(\eta) \leq 1$  for all  $\eta$ .  $\square$

**Corollary 4.5 (Comparison of  $\beta_g$  with  $\beta_h$ ).** *Let  $g, h: X \rightarrow [0, 1]$ . Then  $\beta_g \preceq \beta_h$  if and only if  $g \geq h$ .*

For the cumulative distribution of  $\beta_g * \beta_h$  we obtain the following important identity:

**Lemma 4.4 (Collapse identity).** *Let  $g, h: X \rightarrow [0, 1]$ . Then*

$$\widetilde{\beta_g * \beta_h} = \tilde{\beta}_{1-(1-g)(1-h)} - I_g * \rho^1 * I_h. \quad (4.42)$$

*Proof.* We use the Kronecker delta notation  $\delta_{ij} = 1$  if  $i = j$  and 0 otherwise. Let  $\eta = (x_0, \dots, x_k)$ . Then

$$\begin{aligned} (\widetilde{\beta_g * \beta_h})(\eta) &= \tilde{\beta}_g(\eta^-) + [1 - h(x_k)]\beta_g(\eta) \\ &= 1 - g(x_0)\delta_{k1} + [1 - h(x_k)][[1 - g(x_0)]\delta_{k0} + g(x_0)\delta_{k1}] \\ &= 1 - [1 - g(x_0)][1 - h(x_0)]\delta_{k0} - g(x_0)h(x_1)\delta_{k1} \\ &= \tilde{\beta}_{1-(1-g)(1-h)}(\eta) - (I_g * \rho^1 * I_h)(\eta) \end{aligned} \quad (4.43)$$

where the first equality uses (4.38), the second uses (4.35) and the definition of  $\beta_g$ , the third is an easy rewriting, and the fourth again uses (4.35).  $\square$

**Corollary 4.6 (Collapse inequality).** *Let  $g_i: X \rightarrow [0, 1]$  for  $i = 1, \dots, n$ . Then*

$$\beta_{g_1} * \dots * \beta_{g_n} \preceq \beta_{1 - \prod_{i=1}^n (1-g_i)}. \quad (4.44)$$

*Proof.* The claim is trivial for  $n = 1$ . For  $n = 2$  it follows immediately from the identity (4.42) and the fact that  $I_{g_1} * \rho^1 * I_{g_2} \geq 0$ . The cases  $n \geq 3$  can be proven by an elementary induction using the case  $n = 2$  along with Corollary 4.3.  $\square$

We next need to know under what circumstances the partial ordering  $\preceq$  is preserved by convolution. Convolution from the left is easy:

**Proposition 4.3 (Convolution from the left).** *Let  $\mu, \nu_1, \nu_2$  be clouds satisfying  $\mu \geq 0$  and  $\nu_1 \preceq \nu_2$ . Then  $\mu * \nu_1 \preceq \mu * \nu_2$ .*

*Proof.* This is an immediate consequence of the formula (4.29).  $\square$

Physically, Proposition 4.3 says that if  $\nu_1$  is a better cleaner than  $\nu_2$ , then this same relation holds if both cleaners are preceded by an arbitrary nonnegative operation  $\mu$ .

The behavior of the partial ordering  $\leq$  under convolution from the right is considerably more subtle; this reflects the fact that the partial ordering between cleaners is not preserved by arbitrary *subsequent* nonnegative operations. Rather, we need to limit the class of subsequent operations that are allowed:

**Proposition 4.4 (Convolution from the right).** *Let  $\nu \geq 0$  be a cloud. Then the following are equivalent:*

- (a) *For all pairs of clouds  $\mu_1, \mu_2$  satisfying  $\mu_1 \leq \mu_2$ , we have  $\mu_1 * \nu \leq \mu_2 * \nu$ .*
- (b)  *$\rho^k * \nu \leq \nu$  for all  $k \geq 0$ .*
- (b')  *$\rho^1 * \nu \leq \nu$ .*
- (c)  *$\tilde{\nu}_{\eta'} \leq \tilde{\nu}_\eta$  whenever  $\eta'$  is a suffix of  $\eta$ .*
- (c')  *$\tilde{\nu}_{\eta'} \leq \tilde{\nu}_\eta$  whenever  $\eta$  is a marker of level  $k \geq 1$  and  $\eta' = \eta_1^k$ .*

*Proof.* The implications (b)  $\implies$  (b') and (c)  $\implies$  (c') are trivial, the implication (c')  $\implies$  (c) is easy.

From  $\rho^1 * \nu \leq \nu$  it follows, using Proposition 4.3 and the associativity of convolution, that

$$\rho^{k+1} * \nu = \rho^k * (\rho^1 * \nu) \leq \rho^k * \nu. \quad (4.45)$$

Induction on  $k$  then gives (b')  $\implies$  (b).

Now let  $\eta$  be a marker of level  $\ell$ . By (4.29) we have

$$\widetilde{(\rho^k * \nu)}(\eta) = \begin{cases} \tilde{\nu}(\eta_k^\ell) & \text{if } \ell \geq k, \\ 0 & \text{if } \ell < k. \end{cases} \quad (4.46)$$

From this, we easily get (b)  $\Leftrightarrow$  (c) and (b')  $\Leftrightarrow$  (c'). Hence (b), (b'), (c) and (c') are all equivalent.

The implication (a)  $\implies$  (b) is also trivial, because  $\rho^0 * \nu = \nu$  and  $\rho^k \leq \rho^0$  for all  $k \geq 0$ .

Finally, (c')  $\implies$  (a) is an immediate consequence of (4.30), since (c') ensures that the square brackets in (4.30) are always nonnegative.  $\square$

**Definition 4.2.** We denote by  $\mathcal{R}$  the class of clouds  $\nu \geq 0$  satisfying any one (hence all) of the equivalent conditions of Proposition 4.4.

*Remark.* Given a cloud  $\nu$ , one can define a “dual” cloud  $\nu^*$  by the intertwining relation  $(\rho^0 - \rho^1) * \nu = \nu^* * (\rho^0 - \rho^1)$  [compare (3.1)]. Convoluting on the right with  $\mathbf{1}$ , we obtain  $\nu^*(\eta) = \tilde{\nu}(\eta) - \tilde{\nu}(\eta_-)$ . Thus, the “dual” of a nonnegative cloud  $\nu$  is nonnegative if and only if  $\nu \in \mathcal{R}$ . (Note, however, that  $(\nu^*)^* \neq \nu$ , so this is not a true duality.)

**Corollary 4.7 (Multi-monotonicity).** For  $1 \leq i \leq n$ , let  $\mu_i, \nu_i \geq 0$  be clouds satisfying  $\mu_i \trianglelefteq \nu_i$  and  $\nu_i \in \mathcal{R}$ . Then  $\mu_1 * \dots * \mu_n \trianglelefteq \nu_1 * \dots * \nu_n$ .

*Proof.* We use the telescoping identity

$$\mu_1 * \dots * \mu_n - \nu_1 * \dots * \nu_n = \sum_{i=1}^n \mu_1 * \dots * \mu_{i-1} * (\mu_i - \nu_i) * \nu_{i+1} * \dots * \nu_n. \quad (4.47)$$

By hypothesis, we have  $\mu_i - \nu_i \trianglelefteq 0$  for all  $i$ . By Proposition 4.3, we can convolve on the left with  $\mu_1 * \dots * \mu_{i-1}$ ; and by Proposition 4.4(a) and Definition 4.2, we can convolve on the right successively by  $\nu_{i+1}, \dots, \nu_n$ .  $\square$

*Remark.* In Corollary 4.7, only  $\nu_2, \dots, \nu_n$  really need to belong to  $\mathcal{R}$ , as we never convolve on the right with  $\nu_1$ .

**Corollary 4.8 (Multi-monotonicity for cleaners).** Let  $0 \leq g_i \leq h_i \leq 1$  for  $i = 1, \dots, n$ . Then

$$\beta_{h_1} * \dots * \beta_{h_n} \trianglelefteq \beta_{g_1} * \dots * \beta_{g_n}. \quad (4.48)$$

*Proof.* An immediate consequence of Corollaries 4.5 and 4.7.  $\square$

We now resume our study of the class  $\mathcal{R}$ :

**Proposition 4.5.** The class  $\mathcal{R}$  forms a multiplicative convex cone. That is,

- (a) If  $\mu, \nu \in \mathcal{R}$  and  $a, b \geq 0$ , then  $a\mu + b\nu \in \mathcal{R}$ .
- (b) If  $\mu, \nu \in \mathcal{R}$ , then  $\mu * \nu \in \mathcal{R}$ .

*Proof.* This is immediate from property (a) of Proposition 4.4.  $\square$

We know (Corollary 4.3) that all the clouds  $\beta_f$  ( $0 \leq f \leq 1$ ) belong to  $\mathcal{R}$ , as do all sums of convolutions thereof. But the latter turn out to constitute a strictly smaller class, as they satisfy a condition like that of Proposition 4.4(c) not only for suffixes but also for arbitrary subsequences:

**Proposition 4.6.** Let  $\mu$  be a finite sum of clouds of the form  $a\beta_{f_1} * \dots * \beta_{f_n}$  with  $n \geq 0$ ,  $0 \leq f_1, \dots, f_n \leq 1$  and  $a \geq 0$ . Then  $\tilde{\mu}_{\eta'} \leq \tilde{\mu}_{\eta}$  whenever  $\eta'$  is a subsequence of  $\eta$ .

*Proof.* It suffices to prove the result for  $\mu = \beta_{f_1} * \dots * \beta_{f_n}$ . The case  $n = 0$  (i.e.  $\mu = \rho^0$ ) is trivial, so assume  $n \geq 1$ . First we consider the special case in which each  $f_i = \chi_{\Lambda_i}$ . In this case, we have  $\tilde{\mu}_{\eta} = 0$  or 1 according to the following rule:

$$\tilde{\mu}_{(x_0, \dots, x_k)} = 0 \iff \text{there exist indices } 1 \leq i_0 < \dots < i_k \leq n \text{ such that } x_j \in \Lambda_{i_j} \text{ for } j = 0, \dots, k. \quad (4.49)$$

This is easily seen by induction on  $n$ : recall that  $\beta_\Lambda$  is supported on levels 0 and 1, with  $\beta_\Lambda(x_0) = \chi_{\Lambda^c}(x_0)$  and  $\beta_\Lambda(x_0, x_1) = \chi_\Lambda(x_0)$ , so that (4.49) is clear for  $n = 1$ . For  $n \geq 2$ , let  $\mu = \beta_{\Lambda_1} * \dots * \beta_{\Lambda_n}$  and  $\nu = \beta_{\Lambda_1} * \dots * \beta_{\Lambda_{n-1}}$ . By (4.39), we have for  $\eta = (x_1, \dots, x_k)$

$$\tilde{\mu}_\eta = \chi_{\Lambda_n}(x_k) \tilde{\nu}_{\eta_0^{k-1}} + \chi_{\Lambda_n^c}(x_k) \tilde{\nu}_\eta, \quad (4.50)$$

from which we can see that if  $\nu$  satisfies (4.49), then so does  $\mu$ .

Now, (4.49) obviously implies that if  $\tilde{\mu}_\eta = 0$  and  $\eta'$  is a subsequence of  $\eta$ , then  $\tilde{\mu}_{\eta'} = 0$ . This proves the Proposition in the special case  $f_i = \chi_{\Lambda_i}$ .

To handle the general case, we note that each function  $f: X \rightarrow [0, 1]$  can be written in the form

$$f = \sum_{k \geq 0} a_k \chi_{\Lambda_k} \quad (4.51)$$

where  $a_k \geq 0$ ,  $\sum_k a_k = 1$ , and  $(\Lambda_k)$  is a sequence of (possibly empty) subsets of  $X$ . [One way to get such a decomposition is to use the binary expansion of  $f(x)$ ,

$$f(x) = \sum_{k=1}^{\infty} d_k(x) 2^{-k} \quad (4.52)$$

with  $d_k(x) \in \{0, 1\}$ , and then to set  $a_k := 2^{-k}$  and  $\Lambda_k := \{x \in X : d_k(x) = 1\}$ .] From (4.51) together with the fact that  $\beta_f$  is affine in  $f$ , it follows that each cloud of the form  $\beta_{f_1} * \dots * \beta_{f_n}$  is a convex combination (with a finite or countably infinite number of terms) of clouds of the form  $\beta_{\chi_{\Lambda_1}} * \dots * \beta_{\chi_{\Lambda_n}}$ . Since the set of clouds satisfying the conclusion of the Proposition is obviously stable under convex combination, the proof is complete.  $\square$

**Definition 4.3.** We denote by  $\mathcal{B}$  the class consisting of finite sums of clouds of the form  $a\beta_{f_1} * \dots * \beta_{f_n}$  with  $n \geq 0$ ,  $0 \leq f_1, \dots, f_n \leq 1$  and  $a \geq 0$ .

We denote by  $\mathcal{P}$  the class of clouds  $\mu \geq 0$  for which  $\tilde{\mu}_{\eta'} \leq \tilde{\mu}_\eta$  whenever  $\eta'$  is a subsequence of  $\eta$ .

We have just shown that  $\mathcal{B} \subseteq \mathcal{P} \subseteq \mathcal{R}$ . The following examples show that both these inclusions are strict:

**Example 4.1.** Let  $|X| \geq 2$  and  $\Lambda \subseteq X$  with  $\Lambda \neq \emptyset, X$ . Then  $\pi_\Lambda \in \mathcal{P}$  but  $\pi_\Lambda \notin \mathcal{B}$ . One might worry that this example is somehow “pathological” because  $\pi_\Lambda$  is not supported on finitely many levels. But the next example avoids this objection...

**Example 4.2.** Let  $X = \{x, y\}$ , and let  $\mu$  be the cloud that takes the value 1 on the markers  $xy, xxx, xxy, yx, yyx$  and  $yyy$  and takes the value 0 elsewhere (Figure 1). This is a cloud of norm 1, supported on levels  $\leq 2$ . To verify that

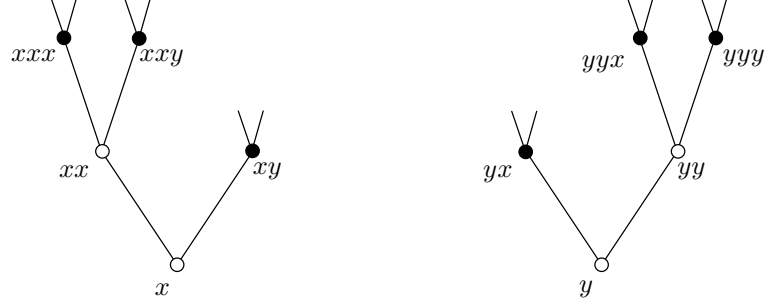


Figure 1. The cloud  $\mu$  in Example 4.2 takes the value 1 (resp. 0) on the markers indicated by a full (resp. empty) circle. This cloud belongs to  $\mathcal{P}$  but not to  $\mathcal{B}$ .

$\mu \in \mathcal{P}$ , it suffices to check that for each  $\eta$  satisfying  $\tilde{\mu}_\eta = 0$ , one has  $\tilde{\mu}_{\eta'} = 0$  for each subsequence  $\eta'$  of  $\eta$ . We leave this verification to the reader.

On the other hand, we claim that  $\mu \notin \mathcal{B}$ . Indeed, suppose that we could write  $\mu$  in the form

$$\mu = \sum_{i=1}^n a_i \beta_{f_1^{(i)}} * \dots * \beta_{f_{n_i}^{(i)}} \quad (4.53)$$

with all  $a_i > 0$ . Then all markers on which  $\mu$  takes the value 0 must also be given mass 0 by the cloud  $\nu_i := \beta_{f_1^{(i)}} * \dots * \beta_{f_{n_i}^{(i)}}$ . Furthermore, it is easily seen that  $\nu_i$  satisfies  $\tilde{\nu}_i(\eta) = 1$  for each marker  $\eta$  of level  $\geq n_i$  (see also Corollary 4.9 below). Since from Figure 1 we see that each marker  $\eta$  has at most one ancestor lying in the support of  $\mu$ , we conclude that we must have  $\nu_i = \mu$  for all  $i$ , so that the right-hand side of (4.53) can be reduced to a single term. We would then have  $\mu = \beta_{f_1} * \dots * \beta_{f_n}$ , where we can assume that none of the functions  $f_i$  are identically 0. Since  $\mu$  charges markers of level 2 but no higher, we must have  $n = 2$ ; and since  $\mu$  takes only the values 0 and 1, we must have  $f_1 = \chi_{\Lambda_1}$  and  $f_2 = \chi_{\Lambda_2}$  for some subsets  $\Lambda_1, \Lambda_2 \subseteq \{x, y\}$ . Since  $\mu(xxx) = 1$ , we must have  $x \in \Lambda_1$  and  $x \in \Lambda_2$ ; likewise, since  $\mu(yyy) = 1$ , we must have  $y \in \Lambda_1$  and  $y \in \Lambda_2$ ; but then  $\mu$  should take the value 0 on the marker  $xy$ , which it does not. This proves that  $\mu \notin \mathcal{B}$ .

**Example 4.3.** Let  $X = \{x, y\}$ , and let  $\mu$  be the cloud that takes the value 1 on the markers  $xx, xyxx, xyxy, xyy, yxx, yxy, yxx$  and  $yyy$  and takes the value 0 elsewhere (Figure 2). This is a cloud of norm 1, supported on levels  $\leq 3$ . To verify condition (c) of Proposition 4.4, it suffices to check that for each  $\eta$  satisfying  $\tilde{\mu}_\eta = 0$ , one has  $\tilde{\mu}_{\eta'} = 0$  for each suffix  $\eta'$  of  $\eta$ . We leave this verification to the reader. Hence  $\mu \in \mathcal{R}$ .



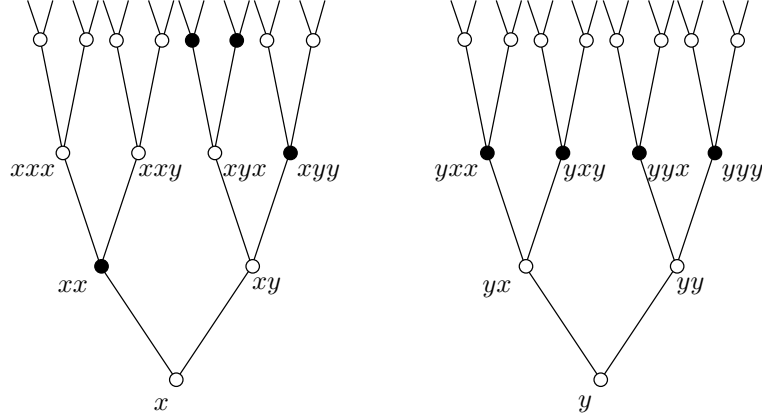


Figure 2. The cloud  $\mu$  in Example 4.3 takes the value 1 (resp. 0) on the markers indicated by a full (resp. empty) circle. This cloud belongs to  $\mathcal{R}$  but not to  $\mathcal{P}$ .

On the other hand,  $\mu$  does not satisfy the analogous condition for arbitrary subsequences, because  $\tilde{\mu}_{xyx} = 0$  while  $\tilde{\mu}_{xx} = 1$ . Hence  $\mu \notin \mathcal{P}$ .

**Proposition 4.7.** *The class  $\mathcal{P}$  forms a multiplicative convex cone.*

*Proof.* The only nontrivial fact to prove is the stability under convolution. So let  $\mu$  and  $\nu$  be clouds in  $\mathcal{P}$ , let  $\eta = (y_0, \dots, y_k)$  be a marker, and let  $\eta' = (y_{i_0}, \dots, y_{i_r})$  ( $0 \leq i_0 < \dots < i_r \leq k$ ) be a subsequence of  $\eta$ . For  $j = 0, \dots, r$ , we set

$$\sigma_j := \eta_0^{y_{i_j}} \preceq \eta. \quad (4.54)$$

Then, for each  $\sigma \preceq \eta$ , we set

$$\sigma \cap \eta' := (y_{i_0}, \dots, y_{i_\ell}), \quad (4.55)$$

where  $\ell$  is the largest index such that  $\sigma_\ell \preceq \sigma$ .

Observing that each  $\sigma' \preceq \eta'$  can be written as  $\sigma_j \cap \eta'$  for a unique  $j$ , we obtain from (4.30)

$$\widetilde{(\mu * \nu)}(\eta') = \sum_{j=0}^r \tilde{\mu}(\sigma_j \cap \eta') [\tilde{\nu}(\eta' \setminus (\sigma_j \cap \eta')) - \tilde{\nu}((\eta' \setminus (\sigma_j \cap \eta'))_-)]. \quad (4.56)$$

Since  $\nu \in \mathcal{P} \subseteq \mathcal{R}$ , the square brackets in the preceding equation are always nonnegative. Moreover, since  $\sigma_j \cap \eta'$  is a subsequence of  $\sigma_j$ , the fact that  $\mu$  belongs to  $\mathcal{P}$  ensures that

$$\tilde{\mu}(\sigma_j \cap \eta') \leq \tilde{\mu}(\sigma_j). \quad (4.57)$$

Hence we get

$$\widetilde{(\mu * \nu)}(\eta') \leq \sum_{j=0}^r \tilde{\mu}(\sigma_j) [\tilde{\nu}(\eta' \parallel (\sigma_j \cap \eta')) - \tilde{\nu}((\eta' \parallel (\sigma_j \cap \eta'))_-)]. \quad (4.58)$$

On the other hand, the right-hand side can also be written as  $\widetilde{(\mu' * \nu)}(\eta')$ , where  $\mu'$  is any cloud giving mass  $\tilde{\mu}(\sigma_j) - \tilde{\mu}(\sigma_{j-1})$  to the marker  $\sigma_j \cap \eta'$  for each  $j = 0, \dots, r$ .

Using now (4.29) and again the hypothesis  $\nu \in \mathcal{P}$ , we get

$$\begin{aligned} \widetilde{(\mu * \nu)}(\eta') &\leq \widetilde{(\mu' * \nu)}(\eta') = \sum_{j=0}^r \mu'(\sigma_j \cap \eta') \tilde{\nu}(\eta' \parallel (\sigma_j \cap \eta')) \\ &\leq \sum_{j=0}^r \mu'(\sigma_j \cap \eta') \tilde{\nu}(\eta \parallel \sigma_j) = \sum_{j=0}^r \left( \sum_{\sigma_{j-1} \prec \sigma \preceq \sigma_j} \mu(\sigma) \right) \tilde{\nu}(\eta \parallel \sigma_j), \end{aligned} \quad (4.59)$$

where  $\sigma_{j-1} \prec \sigma$  means  $\sigma_{j-1} \preceq \sigma$  and  $\sigma_{j-1} \neq \sigma$ , and where we set  $\sigma_{-1} := \emptyset$ . But in each term of (4.59), we have  $\tilde{\nu}(\eta \parallel \sigma_j) \leq \tilde{\nu}(\eta \parallel \sigma)$  since  $\nu \in \mathcal{P} \subseteq \mathcal{R}$ , hence

$$\begin{aligned} \widetilde{(\mu * \nu)}(\eta') &\leq \sum_{\sigma \preceq \sigma_r} \mu(\sigma) \tilde{\nu}(\eta \parallel \sigma) \\ &\leq \sum_{\sigma \preceq \eta} \mu(\sigma) \tilde{\nu}(\eta \parallel \sigma) = \widetilde{(\mu * \nu)}(\eta). \end{aligned} \quad (4.60)$$

□

For any cloud  $\mu \geq 0$  and any marker  $\eta$ , let us define

$$M_\mu(\eta) := \sup_{\eta' : \eta \preceq \eta'} \tilde{\mu}(\eta') \quad (4.61)$$

(this may possibly be  $+\infty$ ); it is the supremum of the sums of  $\mu$  over infinite ascending branches passing through  $\eta$ . Obviously  $\tilde{\mu}(\eta) \leq M_\mu(\eta) \leq \|\mu\|$ , and  $M_\mu(\eta)$  is a decreasing function of  $\eta$  with respect to the partial order  $\preceq$ .

The clouds belonging to the class  $\mathcal{R}$  have a remarkable property:

**Proposition 4.8.** *If  $\mu \geq 0$  belongs to the class  $\mathcal{R}$  characterized in Proposition 4.4, then  $M_\mu(\eta)$  takes the same value for all markers  $\eta$ .*

**Corollary 4.9.** *If  $\mu \geq 0$  belongs to the class  $\mathcal{R}$  and is supported on levels  $\leq N$ , then  $\tilde{\mu}(\eta)$  takes the same value for all markers  $\eta$  of level  $\geq N$ .*

*Proof of Proposition 4.8.* Let  $\eta_1$  and  $\eta_2$  be any two markers. We can always find a marker  $\sigma$  satisfying  $\eta_1 \rightarrow \sigma \rightarrow \eta_2$ . Then, for each marker  $\eta'$  such that  $\eta_2 \preceq \eta'$ , we have [using property (c) of Proposition 4.4]

$$\tilde{\mu}(\eta') \leq \tilde{\mu}(\eta_1 * \sigma * \eta') \leq M_\mu(\eta_1), \quad (4.62)$$

hence

$$M_\mu(\eta_2) \leq M_\mu(\eta_1). \quad (4.63)$$

Reversing the roles of  $\eta_1$  and  $\eta_2$ , we conclude that  $M_\mu(\eta_1) = M_\mu(\eta_2)$ .  $\square$

**Definition 4.4.** For each (finite) real number  $a \geq 0$ , we denote by  $\mathcal{S}_a$  the class of clouds  $\mu \geq 0$  satisfying  $M_\mu(\eta) = a$  for all markers  $\eta$ . We write  $\mathcal{S} = \bigcup_{a \geq 0} \mathcal{S}_a$ .

The class  $\mathcal{S}_1$  will play a major role in the sequel (Sections 4.5–4.8).

We have just shown that  $\mathcal{R} \subseteq \mathcal{S}$ ; let us now show that this inclusion is strict:

**Example 4.4.** Let  $X = \{x, y\}$ , and let  $\mu$  be the cloud that takes the value 1 on the markers  $xx, xyx, xyy$  and  $y$ , and takes the value 0 elsewhere (Figure 3). Then we have  $M_\mu(\eta) = 1$  for every marker  $\eta$ , but  $\rho^1 * \mu \not\leq \mu$  (consider the cumulative distributions on the marker  $xy$ ). So  $\mu \in \mathcal{S}_1$  but  $\mu \notin \mathcal{R}$ .

**Proposition 4.9.** *The class  $\mathcal{S}$  forms a multiplicative convex cone. More specifically:*

- (a) If  $\mu \in \mathcal{S}_a$  and  $\nu \in \mathcal{S}_b$  and  $s, t \geq 0$ , then  $s\mu + t\nu \in \mathcal{S}_{sa+tb}$ .
- (b) If  $\mu \in \mathcal{S}_a$  and  $\nu \in \mathcal{S}_b$ , then  $\mu * \nu \in \mathcal{S}_{ab}$ .

*Proof.* (a) Given  $\mu \in \mathcal{S}_a$ ,  $\nu \in \mathcal{S}_b$  and  $s, t \geq 0$ , we clearly have

$$M_{s\mu+t\nu}(\eta) \leq sa + tb$$

for any marker  $\eta$ . The reverse inequality is easily obtained by choosing  $\eta' \succcurlyeq \eta$  such that  $\sum_{\sigma \preccurlyeq \eta'} s\mu(\sigma) \geq sa - \varepsilon$  and then choosing  $\eta'' \succcurlyeq \eta'$  such that  $\sum_{\sigma \preccurlyeq \eta''} t\nu(\sigma) \geq tb - \varepsilon$ .

(b) Equation (4.29) gives, for any marker  $\eta$ ,

$$\widetilde{(\mu * \nu)}(\eta) \leq b \sum_{\sigma \preccurlyeq \eta} \mu(\sigma) \leq ab. \quad (4.64)$$

For the reverse inequality, we first choose  $\bar{\eta} \succcurlyeq \eta$  such that  $\sum_{\sigma \preccurlyeq \bar{\eta}} \mu(\sigma) \geq a - \varepsilon$ . Observe next using (4.29) that for any  $\eta' \succcurlyeq \bar{\eta}$ , we have

$$\widetilde{(\mu * \nu)}(\eta') \geq \sum_{\sigma \preccurlyeq \bar{\eta}} \mu(\sigma) \tilde{\nu}(\eta' \parallel \sigma). \quad (4.65)$$

Now, if  $\text{level}(\bar{\eta}) = k$ , we construct inductively a sequence  $\bar{\eta} \preccurlyeq \eta'_0 \preccurlyeq \eta'_1 \preccurlyeq \dots \preccurlyeq \eta'_k$  such that  $\tilde{\nu}(\eta'_r \parallel \sigma) \geq b - \varepsilon$  for any  $\sigma \preccurlyeq \bar{\eta}$  with  $\text{level}(\sigma) \leq r$ . Taking  $\eta' = \eta'_k$  in (4.65) gives the desired result.  $\square$

In summary, we have introduced four natural classes of clouds, which are multiplicative convex cones and satisfy  $\mathcal{B} \subsetneq \mathcal{P} \subsetneq \mathcal{R} \subsetneq \mathcal{S}$ .

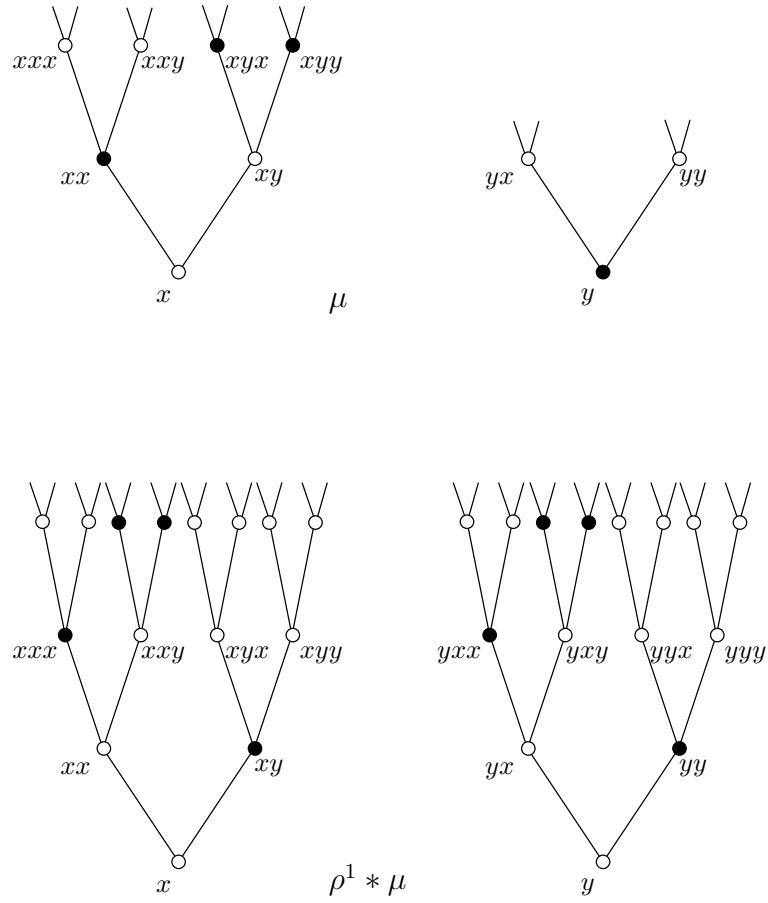


Figure 3. The clouds  $\mu$  and  $\rho^1 * \mu$  in Example 4.4 take the value 1 (resp. 0) on the markers indicated by a full (resp. empty) circle. The cloud  $\mu$  belongs to  $\mathcal{S}_1$  but not to  $\mathcal{R}$ .

#### 4.4. A fundamental inequality

Let us now substantiate our assertion that the partial ordering  $\leq$  is related to efficiency of cleaning.

**Proposition 4.10 (Fundamental comparison inequality).** (FH) *Let  $\mu$  and  $\nu$  be clouds of finite norm. If  $\mu \leq \nu$ , then*

$$T_\mu w \leq T_\nu w. \quad (4.66)$$

If in addition  $\mu, \nu \geq 0$ , then one has

$$\|T_\mu\|_{w \rightarrow w} \leq \|T_\nu\|_{w \rightarrow w} \quad (4.67)$$

and, for every vector  $c \geq 0$ ,

$$\|cT_\mu\|_w \leq \|cT_\nu\|_w. \quad (4.68)$$

*Proof.* Because  $\mu$  and  $\nu$  have finite norm,  $T_\mu w$  (resp.  $T_\nu w$ ) is the pointwise limit of  $T_{\mu_N} w$  (resp.  $T_{\nu_N} w$ ) as  $N \rightarrow \infty$ , where for each  $N \geq 0$  the cloud  $\mu_N$  (resp.  $\nu_N$ ) is supported on levels  $\leq N$  and coincides with  $\mu$  (resp.  $\nu$ ) on these levels. Thus it is enough to prove the result when both  $\mu$  and  $\nu$  are supported on levels  $\leq N$ . We shall do this by induction on  $N$ .

We clearly have (4.66) if  $N = 0$ . So let  $N \geq 1$ , and assume that (4.66) holds whenever  $\mu$  and  $\nu$  are supported on levels  $\leq N - 1$ . Now let  $\mu$  and  $\nu$  be clouds supported on levels  $\leq N$  with  $\mu \leq \nu$ . We can suppose that

$$\nu(\eta) = 0 \quad \text{whenever} \quad \text{level}(\eta) = N; \quad (4.69)$$

for if this is not the case, we can simply replace  $\mu$  and  $\nu$  with the clouds obtained by subtracting  $\nu(\eta)$  from both  $\mu(\eta)$  and  $\nu(\eta)$  for each marker  $\eta$  of level  $N$ . Furthermore, we can also suppose that

$$\mu(\eta) \geq 0 \quad \text{whenever} \quad \text{level}(\eta) = N; \quad (4.70)$$

otherwise, for each marker  $\eta$  of level  $N$  such that  $\mu_\eta < 0$ , we can replace  $\mu_\eta$  with 0: the new cloud obtained in this way is still  $\leq \nu$ .

Next, we consider the cloud  $\mu'$  defined by

$$\mu'(\eta) = \begin{cases} 0 & \text{if level}(\eta) = N, \\ \mu(\eta) + \sup_{x_N \in X} \mu(\eta \circ x_N) & \text{if level}(\eta) = N - 1, \\ \mu(\eta) & \text{otherwise.} \end{cases} \quad (4.71)$$

Note that the definition of  $\mu'$  is the same as (4.14) in the proof of Proposition 4.1, and a similar calculation using the Fundamental Hypothesis gives

$$T_\mu w \leq T_{\mu'} w. \quad (4.72)$$

Now,  $\mu'$  and  $\nu$  are supported on levels  $\leq N - 1$ , and the induction hypothesis gives

$$T_{\mu'}w \leq T_{\nu}w. \quad (4.73)$$

Together these prove (4.66).

Finally, (4.67) and (4.68) are easy consequences of (4.66) in the case of nonnegative clouds  $\mu, \nu$  and, for the latter, nonnegative dirt vectors  $c$ .  $\square$

#### 4.5. Clouds carried by $\Lambda$

The following definition will play a fundamental role in our analysis:

**Definition 4.5.** Let  $\Lambda \subseteq X$ . We say that a cloud  $\mu$  is *carried by*  $\Lambda$  in case  $\mu_{\eta} = 0$  for every marker  $\eta = (x_0, x_1, \dots, x_k)$  having at least one index  $j < k$  with  $x_j \in \Lambda^c$ .

Physically, this means that  $\mu$ , interpreted as a cleaning operator (i.e. as acting by convolution on the right), never sends dirt back into  $\Lambda$  from outside  $\Lambda$ .

Please note that for clouds  $\mu$  carried by  $\Lambda$ , we have  $M_{\mu}(\eta) = \tilde{\mu}(\eta)$  whenever  $\eta$  has at least one entry outside  $\Lambda$ .

**Lemma 4.5.** *Let  $\Lambda \subseteq X$ . Then:*

- (a) *All clouds  $I_f$  are carried by  $\Lambda$ .*
- (b) *The cloud  $\beta_f$  is carried by  $\Lambda$  if and only if  $\text{supp}(f) \subseteq \Lambda$ .*
- (c) *The cloud  $\pi_{\Lambda'}$  is carried by  $\Lambda$  if and only if either  $\Lambda' \subseteq \Lambda$  or  $\Lambda' = X$ .*
- (d) *If  $\mu$  and  $\nu$  are carried by  $\Lambda$  and  $a, b \in \mathbb{R}$ , then  $a\mu + b\nu$  is carried by  $\Lambda$ .*
- (e) *If  $\mu$  and  $\nu$  are carried by  $\Lambda$ , then  $\mu * \nu$  is carried by  $\Lambda$ .*

*Proof.* (a) The cloud  $I_f$  is nonvanishing only on markers of level 0, and so is manifestly carried by  $\Lambda$  for any  $\Lambda$ .

(b) The cloud  $\beta_f$  is nonvanishing only on markers of levels 0 and 1. For the latter we have  $(\beta_f)_{(x_0, x_1)} = f(x_0)$ . It follows that  $\beta_f$  is carried by  $\Lambda$  if and only if  $f(x) = 0$  for all  $x \in \Lambda^c$ .

(c) follows easily from the definitions.

(d) is trivial.

(e) Suppose that  $(\mu * \nu)_{(x_0, \dots, x_k)} \neq 0$ . Then there must exist an index  $j$  ( $0 \leq j \leq k$ ) such that  $\mu_{(x_0, \dots, x_j)} \neq 0$  and  $\nu_{(x_j, \dots, x_k)} \neq 0$ . But since  $\mu$  and  $\nu$  are carried by  $\Lambda$ , we must have  $x_0, \dots, x_{j-1} \in \Lambda$  and  $x_j, \dots, x_{k-1} \in \Lambda$ .  $\square$

Clouds carried by  $\Lambda$  satisfy an identity analogous to Lemma 3.1:

**Lemma 4.6.** *Let  $\Lambda \subseteq X$ , and let  $\mu_1, \dots, \mu_n$  be clouds carried by  $\Lambda$ . Then the cloud  $\mu_1 * \dots * \mu_n * I_\Lambda$  is supported on markers having all their entries in  $\Lambda$ . Moreover, if  $h_i: X \rightarrow \mathbb{R}$  are functions satisfying  $h_i \upharpoonright \Lambda \equiv 1$ , then*

$$\mu_1 * \dots * \mu_n * I_\Lambda = I_{h_1} * \mu_1 * \dots * I_{h_n} * \mu_n * I_\Lambda. \quad (4.74)$$

*Proof.* By Lemma 4.5, the cloud  $\mu_1 * \dots * \mu_n$  is carried by  $\Lambda$ . It follows that  $\mu_1 * \dots * \mu_n * I_\Lambda$  can be nonzero only on markers  $\eta = (x_0, \dots, x_k)$  with all  $x_i \in \Lambda$  ( $0 \leq i \leq k$ ). For such  $\eta$ , the presence of the factors  $I_{h_i}$  changes nothing.  $\square$

We also have the following property concerning convolutions with  $\pi_\Lambda$ :

**Lemma 4.7.** *Let  $\Lambda \subseteq X$ , and let  $\mu \geq 0$  be a cloud. Then the following are equivalent:*

- (a)  $\mu$  is carried by  $\Lambda$ .
- (b)  $\mu * \pi_\Lambda$  is carried by  $\Lambda$ .

*Proof.* (a)  $\implies$  (b) follows from Lemma 4.5(c,e).

(b)  $\implies$  (a): Suppose that  $\mu$  is not carried by  $\Lambda$ , i.e. that there exists a marker  $\eta = (x_0, x_1, \dots, x_k)$  and an index  $j < k$  such that  $x_j \in \Lambda^c \neq \emptyset$  and  $\mu(\eta) > 0$ . If  $x_k \in \Lambda^c$ , we have  $(\mu * \pi_\Lambda)(\eta) \geq \mu(\eta) > 0$ ; if  $x_k \in \Lambda$ , we have  $(\mu * \pi_\Lambda)(\eta') \geq \mu(\eta) > 0$  for  $\eta' = (x_0, x_1, \dots, x_k, y)$  with any  $y \in \Lambda^c$ ; either way we conclude that  $\mu * \pi_\Lambda$  is not carried by  $\Lambda$ .  $\square$

*Remark.* The implication (b)  $\implies$  (a) is false if  $\mu$  is not assumed nonnegative. To see this, consider  $X = \{x, y\}$  and  $\Lambda = \{x\}$ , and set  $\mu(yx) = 1$ ,  $\mu(yx^k y) = -1$  for all  $k \geq 1$ , and  $\mu = 0$  on all other markers. Then it is not hard to verify that  $\mu * \pi_\Lambda = 0$ . But  $\mu$  is not carried by  $\Lambda$ .

For certain pairs of clouds carried by  $\Lambda$ , we can prove an inequality going in the *opposite* direction to Proposition 4.10, provided that we look only at markers ending *outside*  $\Lambda$ ; moreover, this inequality holds pointwise. Let us recall that  $\partial\Lambda$  denotes the set of markers of the form  $\eta = (x_0, \dots, x_k)$  [ $k \geq 0$ ] for which  $x_0, \dots, x_{k-1} \in \Lambda$  and  $x_k \in \Lambda^c$ .

**Proposition 4.11.** *Let  $\Lambda \subseteq X$ , and let  $\mu, \nu \geq 0$  be clouds carried by  $\Lambda$ . Suppose further that  $\tilde{\mu}(\eta) \leq \tilde{\nu}(\eta)$  for all markers  $\eta \in \partial\Lambda$ . Then the following are equivalent:*

- (a)  $\mu \leq \nu$ .
- (b)  $\mu * I_f \leq \nu * I_f$  for every  $f$  satisfying  $\chi_\Lambda \leq f \leq 1$ .
- (c)  $\mu * I_\Lambda \leq \nu * I_\Lambda$ .

(d)  $\tilde{\mu}(\eta) \leq \tilde{\nu}(\eta)$  for every marker  $\eta$  having all its elements inside  $\Lambda$ .

Moreover, if  $\Lambda \neq X$  and  $\tilde{\mu}(\eta) = \tilde{\nu}(\eta)$  for all  $\eta \in \partial\Lambda$ , then (a)–(d) are equivalent to

(e)  $\mu * I_{\Lambda^c} \geq \nu * I_{\Lambda^c}$ .

*Proof.* Since  $\mu$  and  $\nu$  are carried by  $\Lambda$ , the hypothesis implies that  $\tilde{\mu}(\eta) \leq \tilde{\nu}(\eta)$  for every marker  $\eta$  having at least one element outside  $\Lambda$ . It follows that (a) is equivalent to (d).

On the other hand, the inequalities expressing (a)–(c) are all identical when evaluated on markers  $\eta$  having all their elements in  $\Lambda$ : they simply assert (d). So we have (b)  $\implies$  (c)  $\implies$  (d), and it suffices now to show that (a) implies the inequality (b) when evaluated on markers  $\eta$  having at least one element outside  $\Lambda$ . So let  $\eta$  be such a marker, and let  $\sigma$  be its unique ancestor in  $\partial\Lambda$ . Then

$$\begin{aligned} \widetilde{(\mu * I_f)}(\eta) &= \widetilde{(\mu * I_f)}(\sigma) = \tilde{\mu}(\sigma^-) + f(\text{last}(\sigma))\mu(\sigma) \\ &= [1 - f(\text{last}(\sigma))]\tilde{\mu}(\sigma^-) + f(\text{last}(\sigma))\tilde{\mu}(\sigma) \end{aligned}$$

and likewise for  $\nu$ . [Recall that if  $\eta = (x_0, \dots, x_k)$ , then  $\eta^- := (x_0, \dots, x_{k-1})$ .] It follows, by taking a convex combination of inequalities, that (a) implies (b).

Since  $\mu$  is carried by  $\Lambda$ , the cloud  $\mu * I_{\Lambda^c}$  is supported on  $\partial\Lambda$ , and for  $\eta \in \partial\Lambda$  we have

$$(\mu * I_{\Lambda^c})(\eta) = \tilde{\mu}(\eta) - \tilde{\mu}(\eta^-) \quad (4.75)$$

and likewise for  $\nu$ . If  $\tilde{\mu}(\eta) = \tilde{\nu}(\eta)$  for all  $\eta \in \partial\Lambda$ , then (e) is equivalent to

$$\tilde{\mu}(\eta^-) \leq \tilde{\nu}(\eta^-) \quad (4.76)$$

for all  $\eta \in \partial\Lambda$ . But since  $\Lambda \neq X$ , every marker  $\sigma$  with all its elements in  $\Lambda$  is of the form  $\eta^-$  for some  $\eta \in \partial\Lambda$ , so this is equivalent to (d).  $\square$

**Corollary 4.10.** *Let  $\Lambda \subseteq X$ , and let  $\mu, \nu \geq 0$  be clouds carried by  $\Lambda$  and belonging to the class  $\mathcal{S}_a$  (cf. Definition 4.4) for the same constant  $a$ . If  $\mu \leq \nu$ , then*

- (i)  $\mu * I_{\Lambda^c} \geq \nu * I_{\Lambda^c}$ ,
- (ii)  $\mu * I_f \leq \nu * I_f$  for every  $f$  satisfying  $\chi_\Lambda \leq f \leq 1$ .

This is an immediate consequence of Proposition 4.11 together with the definition of  $\mathcal{S}_a$ , which entails that  $M_\mu(\eta) = M_\nu(\eta) = a$  for all markers  $\eta$ . Corollary 4.10 applies in particular (with  $a = 1$ ) if  $\mu$  and  $\nu$  are of the form  $\beta_{h_1} * \dots * \beta_{h_n}$  with all  $\text{supp}(h_i) \subseteq \Lambda$ .



#### 4.6. $\Lambda$ -regular clouds

**Definition 4.6.** We say that cloud  $\mu \geq 0$  is  $\Lambda$ -regular in case it satisfies the following two conditions:

$$\tilde{\mu}(\eta) \leq 1 \quad \text{for all markers } \eta \quad [\text{i.e. } \|\mu\| \leq 1], \quad (4.77)$$

$$\tilde{\mu}(\eta) = 1 \quad \text{for all } \eta \in \partial\Lambda. \quad (4.78)$$

**Lemma 4.8.** Let  $\Lambda \subseteq X$ , and let  $\mu \geq 0$  be a cloud.

- (a) If  $\Lambda \neq X$ , then  $\mu$  is  $\Lambda$ -regular if and only if it is carried by  $\Lambda$  and belongs to the class  $\mathcal{S}_1$ .
- (b) If  $\Lambda = X$ , then  $\mu$  is  $\Lambda$ -regular if and only if  $\|\mu\| \leq 1$ .

*Proof.* (a) If  $\mu \in \mathcal{S}_1$ , then (4.77) clearly holds. If, in addition,  $\mu$  is carried by  $\Lambda$ , we have  $\tilde{\mu}(\eta) = M_\mu(\eta) = 1$  whenever  $\eta$  has at least one element outside  $\Lambda$ , so that (4.78) holds.

Conversely, suppose that  $\mu$  is  $\Lambda$ -regular and that  $\Lambda \neq X$ . If  $\eta$  has all its entries in  $\Lambda$ , then it is a proper ancestor of some  $\sigma \in \partial\Lambda$  (since  $\Lambda \neq X$ ), and we have  $M_\mu(\eta) \geq \tilde{\mu}(\sigma) = 1$ . If  $\eta$  has at least one entry outside  $\Lambda$ , then it has a (unique) marker  $\sigma \in \partial\Lambda$  as an ancestor, in which case  $M_\mu(\eta) \geq \tilde{\mu}(\eta) \geq \tilde{\mu}(\sigma) = 1$ . On the other hand, by (4.77) we always have  $M_\mu(\eta) \leq 1$ , hence  $M_\mu(\eta) = 1$  for all  $\eta$  and thus  $\mu \in \mathcal{S}_1$ . Finally, if  $\eta$  has some proper ancestor  $\sigma \in \partial\Lambda$ , then (4.77), (4.78) imply  $1 \geq \tilde{\mu}(\eta) \geq \mu(\eta) + \tilde{\mu}(\sigma) = \mu(\eta) + 1$ , hence  $\mu(\eta) = 0$ , so that  $\mu$  is carried by  $\Lambda$ .

- (b) If  $\Lambda = X$ , the condition (4.78) is empty. □

**Corollary 4.11.** Let  $\Lambda \subseteq X$ . Then:

- (a) The cloud  $\rho^0$  (indicator of level 0) is  $\Lambda$ -regular.
- (b) The cloud  $\beta_f$  is  $\Lambda$ -regular if and only if  $\text{supp}(f) \subseteq \Lambda$ .
- (c) The cloud  $\pi_{\Lambda'}$  is  $\Lambda$ -regular if and only if  $\Lambda' \subseteq \Lambda$ .
- (d) If  $\mu$  and  $\nu$  are  $\Lambda$ -regular, then  $\mu * \nu$  is  $\Lambda$ -regular.

*Proof.* This is an immediate consequence of Lemma 4.8 along with Proposition 4.9 and Lemma 4.5; for part (c) one has to think separately about the cases in which  $\Lambda$  and/or  $\Lambda'$  is equal or not equal to  $X$ . □

**Corollary 4.12.** Let  $\Lambda \subseteq X$ , and let  $\mu$  be a  $\Lambda$ -regular cloud. Then

$$I_{\Lambda^c} * \mu = I_{\Lambda^c}, \quad (4.79)$$

$$I_\Lambda * \mu * I_\Lambda = \mu * I_\Lambda. \quad (4.80)$$

*Proof.* We begin by proving (4.79). The claim is trivial if  $\Lambda = X$ , so assume  $\Lambda \neq X$ . By Lemma 4.8,  $\mu$  is carried by  $\Lambda$ , so it must vanish on all markers of the form  $(x_0, \dots, x_k)$  with  $x_0 \in \Lambda^c$  and  $k \geq 1$ . But then, since  $\mu$  belongs to  $\mathcal{S}_1$ , it must give mass 1 to each marker of the form  $(x_0)$  with  $x_0 \in \Lambda^c$ .

(4.80) follows from (4.79) by convoluting on the right with  $I_\Lambda$  and rearranging.  $\square$

**Corollary 4.13.** *Let  $\Lambda \subseteq X$ , and let  $\mu, \nu$  be  $\Lambda$ -regular clouds. If  $\mu \preceq \nu$ , then*

- (i)  $\mu * I_{\Lambda^c} \geq \nu * I_{\Lambda^c}$ .
- (ii)  $\mu * I_f \preceq \nu * I_f$  for every  $f$  satisfying  $\chi_\Lambda \leq f \leq 1$ .

*Proof.* If  $\Lambda \neq X$ , this is an immediate consequence of Lemma 4.8 and Corollary 4.10. If  $\Lambda = X$ , the claims are trivially true.  $\square$

#### 4.7. Comparison of $T_\mu$ with $\Pi_\Lambda$

We recall that the cloud  $\pi_\Lambda$ , which is the indicator of  $\partial\Lambda$ , is associated to the operator  $\Pi_\Lambda$ . This cloud plays a special role among  $\Lambda$ -regular clouds, by virtue of its minimality with respect to  $\preceq$ :

**Lemma 4.9.** *Let  $\Lambda \subseteq X$ , and let  $\mu$  be a  $\Lambda$ -regular cloud. Then:*

- (a)  $\mu * I_{\Lambda^c} \leq \pi_\Lambda \preceq \mu \preceq \rho^0$ .
- (b) *If  $\Lambda \neq X$  and  $\mu$  is supported on levels  $\leq N$ , then  $\pi_\Lambda \preceq \beta_\Lambda^N \preceq \mu$ .*

*Proof.* (a) Let  $\mu$  be any  $\Lambda$ -regular cloud and let  $\eta$  be any marker. Then either there exists  $\sigma \preceq \eta$  with  $\sigma \in \partial\Lambda$ , in which case  $\tilde{\mu}(\eta)$  and  $\tilde{\pi}_\Lambda(\eta)$  are both equal to 1; or else there does not exist such a  $\sigma$ , in which case  $\tilde{\pi}_\Lambda(\eta) = 0 \leq \tilde{\mu}(\eta)$ . This proves that  $\pi_\Lambda \preceq \mu$ . Then, Corollary 4.13 gives  $\mu * I_{\Lambda^c} \leq \pi_\Lambda * I_{\Lambda^c} = \pi_\Lambda$ . The inequality  $\mu \preceq \rho^0$  is trivial.

(b) Corollary 4.11(b,d) shows that  $\beta_\Lambda^N$  is  $\Lambda$ -regular, so the relation  $\pi_\Lambda \preceq \beta_\Lambda^N$  follows from (a). For the second inequality, observe first that for each marker  $\eta$  of level  $\leq N$ , we have  $\beta_\Lambda^N(\eta) = \pi_\Lambda(\eta)$ , hence

$$\tilde{\beta}_\Lambda^N(\eta) = \tilde{\pi}_\Lambda(\eta) \leq \tilde{\mu}(\eta), \quad (4.81)$$

where the last inequality again uses (a). On the other hand, if  $\text{level}(\eta) > N$ , since  $\beta_\Lambda^N$  and  $\mu$  are  $\Lambda$ -regular clouds supported on levels  $\leq N$ , we always have

$$\tilde{\beta}_\Lambda^N(\eta) = 1 = \tilde{\mu}(\eta), \quad (4.82)$$

where the last inequality uses  $\mu \in \mathcal{S}_1$  from Lemma 4.8(a). This proves that  $\beta_\Lambda^N \preceq \mu$ .  $\square$

**Lemma 4.10.** *Let  $\Lambda \subseteq X$ , and let  $\mu \geq 0$  be a cloud. Consider the following statements:*

- (a)  $\mu$  is  $\Lambda$ -regular.
- (b)  $\mu * \pi_\Lambda = \pi_\Lambda$ .
- (c)  $\pi_\Lambda * \mu = \pi_\Lambda$ .

Then (a) implies both (b) and (c); and if  $\Lambda \neq X$ , then (b) implies (a) as well.

*Proof.* Suppose first that  $\mu$  is  $\Lambda$ -regular. Then by Lemma 4.9, we have  $\pi_\Lambda \leq \mu \leq \rho^0$ . Since  $\pi_\Lambda \in \mathcal{R}$ , we can right-convolute this inequality with  $\pi_\Lambda$  to obtain  $\pi_\Lambda \leq \mu * \pi_\Lambda \leq \pi_\Lambda$  (since  $\pi_\Lambda * \pi_\Lambda = \pi_\Lambda$ ), thus proving (b). Likewise, since  $\pi_\Lambda \geq 0$ , we can left-convolute with  $\pi_\Lambda$  to prove (c).

Now suppose that  $\mu * \pi_\Lambda = \pi_\Lambda$ . By Lemma 4.7 we deduce that  $\mu$  is carried by  $\Lambda$ . Moreover, if  $\eta \in \partial\Lambda$ , we have

$$1 = \pi_\Lambda(\eta) = (\mu * \pi_\Lambda)(\eta) = \sum_{\sigma \preceq \eta} \mu(\sigma) \pi_\Lambda(\eta \setminus \sigma) = \sum_{\sigma \preceq \eta} \mu(\sigma) = \tilde{\mu}(\eta). \quad (4.83)$$

When  $\Lambda \neq X$ , it easily follows that  $\mu$  is  $\Lambda$ -regular.  $\square$

*Remark.* The implication (c)  $\implies$  (a) is false when  $\Lambda \neq \emptyset$ : consider, for instance,  $\mu = I_{\Lambda^c}$ . Indeed, (c) does not even imply that  $\mu$  is carried by  $\Lambda$ : consider, for instance,  $\mu = I_{\Lambda^c} + I_\Lambda * \rho^1 * I_{\Lambda^c} * \rho^1$ .

We now examine the deviation  $\mu - \pi_\Lambda$ , and prove analogues of Proposition 3.6 and Lemma 3.5:

**Lemma 4.11 (Comparison with  $\pi_\Lambda$ ).** (FH) *Let  $\mu$  be a  $\Lambda$ -regular cloud. Then:*

- (a) *The cloud  $\mu - \pi_\Lambda$  can be decomposed in the form*

$$\mu - \pi_\Lambda = (\mu - \pi_\Lambda) * I_\Lambda + (\mu - \pi_\Lambda) * I_{\Lambda^c} \quad (4.84)$$

$$= \mu * I_\Lambda - (\pi_\Lambda - \mu * I_{\Lambda^c}) \quad (4.85)$$

$$= \mu * I_\Lambda - \mu * I_\Lambda * \pi_\Lambda \quad (4.86)$$

where

- (i)  $\mu * I_\Lambda$  is nonnegative and is supported on markers all of whose entries are in  $\Lambda$ ;

and

- (ii)  $\pi_\Lambda - \mu * I_{\Lambda^c}$  is nonnegative, is  $\leq \mu * I_\Lambda$ , and is supported on markers in  $\partial\Lambda$ .

(b) For any vector  $c \geq 0$ , we have

$$\|c(T_\mu - \Pi_\Lambda)\|_w = \|cT_\mu I_\Lambda\|_w + \|c(\Pi_\Lambda - T_\mu I_{\Lambda^c})\|_w \leq 2\|cT_\mu I_\Lambda\|_w. \quad (4.87)$$

(c) We have

$$\|T_\mu - \Pi_\Lambda\|_{w \rightarrow w} \leq 2\|T_\mu I_\Lambda\|_{w \rightarrow w}. \quad (4.88)$$

*Proof.* (a) The equalities (4.84), (4.85) are trivial, and (4.86) follows by using Lemma 4.10:  $\pi_\Lambda - \mu * I_{\Lambda^c} = \mu * \pi_\Lambda - \mu * I_{\Lambda^c} * \pi_\Lambda = \mu * I_\Lambda * \pi_\Lambda$ . The inequalities  $0 \leq \pi_\Lambda - \mu * I_{\Lambda^c} \leq \mu * I_\Lambda$  are an immediate consequence of Lemma 4.9(a).

(b) The equality is an immediate consequence of the identity, sign and support properties from part (a). The inequality comes from  $0 \leq \pi_\Lambda - \mu * I_{\Lambda^c} \leq \mu * I_\Lambda$  together with Proposition 4.10.

(c) For a general vector  $c$ , we can always write

$$\begin{aligned} \|c(T_\mu - \Pi_\Lambda)\|_w &\leq \| |c| |T_\mu - \Pi_\Lambda| \|_w \\ &= \| |c| T_\mu I_\Lambda \|_w + \| |c| (\Pi_\Lambda - T_\mu I_{\Lambda^c}) \|_w, \end{aligned} \quad (4.89)$$

and the result then follows from part (b).  $\square$

**Proposition 4.12 (Comparison of cleaners).** (FH) *Let  $\mu$  and  $\nu$  be  $\Lambda$ -regular clouds, with  $\mu \leq \nu$ . Then:*

(a) We have

$$\mu * I_\Lambda \leq \nu * I_\Lambda, \quad (4.90)$$

$$\mu * I_{\Lambda^c} \geq \nu * I_{\Lambda^c}, \quad (4.91)$$

$$\pi_\Lambda - \mu * I_{\Lambda^c} \leq \pi_\Lambda - \nu * I_{\Lambda^c}. \quad (4.92)$$

(b) For any vector  $c \geq 0$ , we have

$$\|c(T_\mu - \Pi_\Lambda)\|_w \leq \|c(T_\nu - \Pi_\Lambda)\|_w, \quad (4.93)$$

$$cT_\mu I_{\Lambda^c} \geq cT_\nu I_{\Lambda^c}. \quad (4.94)$$

*Proof.* (a) is a restatement of Corollary 4.13. (b) then follows by using the equality in (4.87) together with Proposition 4.10, exploiting the nonnegativity of all the operators in question.  $\square$

**Proposition 4.13 (Comparison of cleaning sequences).** *Let  $(\mu_n)_{n \geq 1}$  and  $(\nu_n)_{n \geq 1}$  be sequences of nonnegative clouds, with  $\mu_n$   $\Lambda$ -regular and  $\mu_n \leq \nu_n$  for all  $n$ . If  $\nu_n$  converges pointwise to  $\pi_\Lambda$  [i.e.  $\lim_{n \rightarrow \infty} \nu_n(\eta) = \pi_\Lambda(\eta)$  for all markers  $\eta$ ], then  $\mu_n$  also converges pointwise to  $\pi_\Lambda$ .*

*Proof.* Since  $\mu_n \leq \nu_n$ , we have  $\widetilde{\mu}_n \leq \widetilde{\nu}_n$ . The pointwise convergence of  $\nu_n$  to  $\pi_\Lambda$  is equivalent to the pointwise convergence of  $\widetilde{\nu}_n$  to  $\widetilde{\pi}_\Lambda$ , where

$$\widetilde{\pi}_\Lambda(\eta) = \begin{cases} 0 & \text{if } \eta \text{ has all its entries in } \Lambda, \\ 1 & \text{otherwise} \end{cases} \quad (4.95)$$

(that is,  $\widetilde{\pi}_\Lambda = \mathbf{1} - \mathbf{1}_\Lambda$ ). Hence  $\widetilde{\mu}_n(\eta) \rightarrow 0$  if  $\eta$  has all its entries in  $\Lambda$ ; and for all other markers  $\eta$  we have  $\widetilde{\mu}_n(\eta) = 1$  for all  $n$  by  $\Lambda$ -regularity. Hence  $\widetilde{\mu}_n \rightarrow \widetilde{\pi}_\Lambda$ , i.e.  $\mu_n \rightarrow \pi_\Lambda$ .  $\square$

#### 4.8. Convergence of cleaning operators

We are now ready to study the convergence of cleaning operators  $\beta_{h_1} \cdots \beta_{h_n}$  to  $\Pi_\Lambda$ , analogously to what was done in Section 3.5. But in the cloud context we can shed additional light on this convergence by dividing our analysis into two parts:

- (1) Let  $(\mu_n)_{n \geq 1}$  be a sequence of clouds that converges, in some suitable topology, to a limiting cloud  $\mu_\infty$ . Under what conditions can we show that the operators  $T_{\mu_n}$  converge, in some correspondingly suitable topology, to  $T_{\mu_\infty}$ ?
- (2) Consider the special case  $\mu_n = \beta_{h_1} * \dots * \beta_{h_n}$ . Under what conditions on the sequence  $(h_n)$  does  $\mu_n$  converge to  $\pi_\Lambda$  in the topology needed in part (1)?

We shall carry out this two-part analysis in two versions:

- (a) Pointwise convergence  $\mu_n \rightarrow \mu_\infty$  entails vector-norm convergence  $T_{\mu_n} \rightarrow T_{\mu_\infty}$ .
- (b) Uniform-on-levels convergence  $\mu_n \rightarrow \mu_\infty$  entails operator-norm convergence  $T_{\mu_n} \rightarrow T_{\mu_\infty}$ .

In each case we shall require, as was done in Section 3.5, that  $(I_\Lambda \alpha I_\Lambda)^k \rightarrow 0$  in a suitable topology.

##### 4.8.1. Pointwise (vector-norm) version

We begin with the pointwise (vector-norm) version of the convergence theorems. So let  $(\mu_n)_{n \geq 1}$  be a sequence of clouds such that, for each marker  $\eta$ , the sequence  $\mu_n(\eta)$  converges to a limit  $\mu_\infty(\eta)$ . In order to control the convergence of the  $T_{\mu_n}$ , we shall assume that all the clouds  $\mu_n$  are  $\Lambda$ -regular (for some fixed set  $\Lambda \subseteq X$ ). It is immediate from Definition 4.6 that the limiting cloud  $\mu_\infty$  is likewise  $\Lambda$ -regular.

**Theorem 4.1. (Convergence theorem, pointwise version).** (FH) *Let  $(\mu_n)_{n \geq 1}$  be a sequence of  $\Lambda$ -regular clouds satisfying, for each marker  $\eta$ ,*

$$\mu_n(\eta) \xrightarrow[n \rightarrow \infty]{} \mu_\infty(\eta). \quad (4.96)$$

*Then, for any vector  $c \geq 0$  with  $cI_\Lambda \in l^1(w)$  and satisfying*

$$\|c(I_\Lambda \alpha I_\Lambda)^k\|_w \xrightarrow[k \rightarrow \infty]{} 0, \quad (4.97)$$

*we have*

$$\|c(T_{\mu_n} - T_{\mu_\infty})\|_w \xrightarrow[n \rightarrow \infty]{} 0. \quad (4.98)$$

The proof of Theorem 4.1 will be based on the following lemma:

**Lemma 4.12.** *Let  $\mu$  and  $\nu$  be  $\Lambda$ -regular clouds. Then*

$$|\mu - \nu| \leq 2|\mu - \nu| * I_\Lambda. \quad (4.99)$$

*Proof.* We shall prove that

$$|\mu - \nu| * I_{\Lambda^c} \leq |\mu - \nu| * I_\Lambda, \quad (4.100)$$

which is obviously equivalent to (4.99). Since  $\mu$  is  $\Lambda$ -regular,  $\mu * I_{\Lambda^c}$  charges only markers in  $\partial\Lambda$ , and for  $\sigma \in \partial\Lambda$  we have

$$(\mu * I_{\Lambda^c})(\sigma) = \mu(\sigma) = 1 - \sum_{\sigma' \prec \sigma} \mu(\sigma') = 1 - \widetilde{(\mu * I_\Lambda)}(\sigma). \quad (4.101)$$

The same result also holds for  $\nu$ , so that any marker  $\sigma \in \partial\Lambda$  we have

$$(|\mu - \nu| * I_{\Lambda^c})(\sigma) \leq \widetilde{(|\mu - \nu| * I_\Lambda)}(\sigma). \quad (4.102)$$

Now, for any marker  $\eta$ , either  $\eta$  has all its entries lying in  $\Lambda$ , in which case

$$\widetilde{(|\mu - \nu| * I_{\Lambda^c})}(\eta) = 0 \leq \widetilde{(|\mu - \nu| * I_\Lambda)}(\eta), \quad (4.103)$$

or else there exists a unique  $\sigma \preceq \eta$  with  $\sigma \in \partial\Lambda$ , and then we have

$$\begin{aligned} \widetilde{(|\mu - \nu| * I_{\Lambda^c})}(\eta) &= (|\mu - \nu| * I_{\Lambda^c})(\sigma) \\ &\leq \widetilde{(|\mu - \nu| * I_\Lambda)}(\sigma) \leq \widetilde{(|\mu - \nu| * I_\Lambda)}(\eta). \end{aligned} \quad (4.104)$$

□

*Proof of Theorem 4.1.* Fix any  $\varepsilon > 0$ , and choose  $k$  so that

$$\|c(I_\Lambda \alpha I_\Lambda)^k\|_w \leq \varepsilon. \quad (4.105)$$

It is easy to see that for all  $j \geq 1$  we have

$$(I_\Lambda * \rho^1 * I_\Lambda)^{j+1} \leq (I_\Lambda * \rho^1 * I_\Lambda)^j \leq \dots \leq I_\Lambda; \quad (4.106)$$

therefore, by Proposition 4.10, the sequence  $(\|c(I_\Lambda \alpha I_\Lambda)^j I_\Lambda\|_w)_{j \geq 0}$  is decreasing. In particular, the hypothesis  $cI_\Lambda \in l^1(w)$  ensures that all these quantities are finite. Therefore we can find a finite subset  $\Lambda' \subseteq \Lambda$  such that for every  $j$  in the interval  $0 \leq j \leq k$  we have

$$\|c(I_\Lambda \alpha I_\Lambda)^j I_{\Lambda \setminus \Lambda'}\|_w \leq \varepsilon/(k+1). \quad (4.107)$$

Using Lemma 4.12 and Proposition 4.10, we get

$$\|c(T_{\mu_n} - T_{\mu_\infty})\|_w \leq \|cT_{|\mu_n - \mu_\infty|}\|_w \leq 2\|cT_{|\mu_n - \mu_\infty|} I_\Lambda\|_w. \quad (4.108)$$

Furthermore, by the  $\Lambda$ -regularity of  $\mu_n$  and  $\mu_\infty$ , the cloud  $|\mu_n - \mu_\infty| * I_\Lambda$  charges only markers with all their entries lying in  $\Lambda$ . Let us now divide the set of such markers into three classes as follows:

Class 1: Markers of level  $\leq k$  with all their entries lying in  $\Lambda'$ .

Class 2: Markers of level  $\leq k$  with at least one entry lying in  $\Lambda \setminus \Lambda'$ .

Class 3: Markers of level  $> k$ .

We can then decompose any cloud  $\mu_* * I_\Lambda$  ( $\mu_* = \mu_n$  or  $\mu_\infty$ ) in the form  $\mu_*^1 + \mu_*^2 + \mu_*^3$  so that  $\mu_*^i$  charges only markers of class  $i$ . The triangle inequality yields

$$\begin{aligned} \|cT_{|\mu_n - \mu_\infty|} I_\Lambda\|_w &\leq \|cT_{|\mu_n^1 - \mu_\infty^1|}\|_w + \|cT_{\mu_n^2}\|_w + \|cT_{\mu_\infty^2}\|_w \\ &\quad + \|cT_{\mu_n^3}\|_w + \|cT_{\mu_\infty^3}\|_w, \end{aligned} \quad (4.109)$$

and we will now bound these contributions separately as follows:

Class 1. Since class 1 is a finite set of markers, condition (4.96) ensures that  $\|cT_{|\mu_n^1 - \mu_\infty^1|}\|_w$  can be bounded by  $\varepsilon$  if  $n$  has been chosen large enough.

Class 2. We compare the cloud  $\mu_*^2$  (which stands for either  $\mu_n^2$  or  $\mu_\infty^2$ ) with the cloud

$$\nu^2 := \sum_{j=0}^k (I_\Lambda * \rho^1 * I_\Lambda)^j * I_{\Lambda \setminus \Lambda'}. \quad (4.110)$$

For each marker  $\eta$ , either there exists some  $\sigma \preccurlyeq \eta$  with  $\sigma \in \partial\Lambda'$  and  $\text{level}(\sigma) \leq k$ , in which case  $\widetilde{\nu}^2(\eta) = 1 \geq \widetilde{\mu}_*^2(\eta)$ , or else  $\widetilde{\mu}_*^2(\eta) = 0$  (remember that  $\mu_*^2$  charges only markers of class 2). Hence, we have

$$\mu_*^2 \trianglelefteq \nu^2. \quad (4.111)$$

Then, Proposition 4.10 and (4.107) together give

$$\|cT_{\mu_*^2}\|_w \leq \sum_{j=0}^k \|c(I_\Lambda \alpha I_\Lambda)^j I_{\Lambda \setminus \Lambda'}\|_w \leq \varepsilon. \quad (4.112)$$

Class 3. The cloud  $\mu_*^3$  (standing for  $\mu_n^3$  or  $\mu_\infty^3$ ) satisfies

$$\mu_*^3 \trianglelefteq (I_\Lambda * \rho^1 * I_\Lambda)^k. \quad (4.113)$$

Applying again Proposition 4.10 gives, by (4.105),

$$\|cT_{\mu_*^3}\|_w \leq \|c(I_\Lambda \alpha I_\Lambda)^k\|_w \leq \varepsilon. \quad (4.114)$$

□

**Example 4.5.** It is natural to ask whether the hypothesis of  $\Lambda$ -regularity in Theorem 4.1 can be weakened to assuming that the  $\mu_n$  are carried by  $\Lambda$ , with  $\|\mu_n\| = 1$  for all  $n$  (or even  $\|\mu_n\|$  uniformly bounded). The answer is no, at least when  $\Lambda^c$  is infinite and  $\Lambda \neq \emptyset$ : it suffices to choose  $x_0 \in \Lambda$  and a sequence of distinct elements  $y_1, y_2, \dots \in \Lambda^c$ , and to let  $\mu_n$  be the cloud taking the value 1 on the marker  $(x_0 y_n)$  and 0 on all other markers. Then  $\mu_n \rightarrow 0$  pointwise; but if we take  $c = \delta_{x_0}$ ,  $w = \mathbf{1}$  and  $\alpha = (1/2)I$ , we have  $\|cT_{\mu_n}\|_w = 1/2$  for all  $n$ . So the role played by Lemma 4.12 in handling markers ending in  $\Lambda^c$  is apparently crucial.

Here is one natural context in which a sequence of clouds  $(\mu_n)$  has a pointwise limit  $\mu_\infty$ : Suppose that  $(\mu_n)$  is a sequence of nonnegative clouds that is decreasing in the sense of  $\trianglelefteq$ , i.e.

$$\dots \trianglelefteq \mu_{n+1} \trianglelefteq \mu_n \trianglelefteq \dots \trianglelefteq \mu_1. \quad (4.115)$$

Then for each marker  $\eta$  the sequence  $\widetilde{\mu}_n(\eta)$  is decreasing and bounded below by 0, hence has a limit. Since  $\mu_n(\eta) = \widetilde{\mu}_n(\eta) - \widetilde{\mu}_n(\eta^-)$ , we deduce that for each  $\eta$  the sequence  $\mu_n(\eta)$  converges as well, to a limit which we call  $\mu_\infty(\eta)$ .

In particular, we get a sequence of  $\Lambda$ -regular clouds satisfying (4.115) by considering

$$\mu_n := \beta_{h_1} * \dots * \beta_{h_n}, \quad (4.116)$$

where  $(h_n)_{n \geq 1}$  is a sequence of functions taking values in  $[0, 1]$  and supported on  $\Lambda$ . We are mainly interested in the case where the limit  $\mu_\infty$  of the sequence (4.116) is equal to  $\pi_\Lambda$ . The following proposition gives a necessary and sufficient condition on the functions  $(h_n)$  for this to happen:



**Proposition 4.14. (Conditions on cleaning sequences, pointwise version).** *Let  $(h_n)_{n \geq 1}$  be a sequence of functions satisfying  $0 \leq h_n \leq \chi_\Lambda$ , and define*

$$\mu_n := \beta_{h_1} * \dots * \beta_{h_n}. \quad (4.117)$$

*Then  $\mu_n$  converges pointwise to  $\pi_\Lambda$  if and only if*

$$\sum_{n=1}^{\infty} h_n = \infty \quad \text{on } \Lambda. \quad (4.118)$$

*Proof.* Let us prove first that the condition (4.118) is necessary. Suppose that there exists  $x \in \Lambda$  such that

$$\sum_{n=1}^{\infty} h_n(x) < \infty, \quad (4.119)$$

and let  $N$  be an integer such that

$$\sum_{n > N} h_n(x) < 1. \quad (4.120)$$

We first claim that we can find a marker  $\eta$  with  $\text{last}(\eta) = x$  such that  $\mu_N(\eta) > 0$ . Indeed, since  $\mu_N$  is supported on finitely many levels, let us choose a marker  $\sigma$  of maximum level with  $\mu_N(\sigma) > 0$ . Then every marker  $\eta$  obtained from  $\sigma$  by changing its last entry — in particular the one with  $\text{last}(\eta) = x$  — satisfies  $\mu_N(\eta) = \mu_N(\sigma) > 0$  (this is a consequence of the fact that  $\mu_N \in \mathcal{S}_1$ ).

Now, for each  $n > N$ , we have

$$\begin{aligned} \mu_n(\eta) &= (\mu_N * \beta_{h_{N+1}} * \dots * \beta_{h_n})(\eta) \\ &\geq \mu_N(\eta)(1 - h_{N+1}(x)) \cdots (1 - h_n(x)) \geq \mu_N(\eta) \left(1 - \sum_{k=N+1}^{\infty} h_k(x)\right). \end{aligned} \quad (4.121)$$

It follows that

$$\mu_n(\eta) \not\xrightarrow[n \rightarrow \infty]{} 0 = \pi_\Lambda(\eta). \quad (4.122)$$

This proves the necessity of (4.118).

Conversely, assume that (4.118) is satisfied, and let us prove that  $\mu_\infty = \pi_\Lambda$ . Since each  $\mu_n$  is  $\Lambda$ -regular, it is enough to verify that for each marker  $\eta = (y_0, \dots, y_k)$  with all  $y_j \in \Lambda$ , we have

$$\mu_n(\eta) \xrightarrow[n \rightarrow \infty]{} 0. \quad (4.123)$$

We fix such a marker and set

$$\Lambda' := \{y_0, \dots, y_k\}. \quad (4.124)$$

Now, let  $N$  be a large enough integer (to be made precise later). Hypothesis (4.118) ensures that if  $n$  is large enough, we can find indices  $0 = n_0 < n_1 < \dots < n_N = n$  such that for each  $j = 1, \dots, N$  and each  $y \in \Lambda'$ , we have

$$1 - \prod_{i=n_{j-1}+1}^{n_j} h_i(y) > \frac{1}{2}. \quad (4.125)$$

Then Corollaries 4.8 and 4.6 imply that

$$\mu_n \trianglelefteq \beta_{1-\prod_{i=n_0+1}^{n_1}(1-h_i)} * \dots * \beta_{1-\prod_{i=n_{N-1}+1}^{n_N}(1-h_i)} \trianglelefteq (\beta_h)^{*N}, \quad (4.126)$$

where

$$h := \frac{1}{2}\chi_{\Lambda'}. \quad (4.127)$$

Hence

$$\tilde{\mu}_n(\eta) \leq \widetilde{(\beta_h)^{*N}}(\eta). \quad (4.128)$$

But an easy induction shows that for each  $j = 0, \dots, k$ , we have

$$(\beta_h)^{*N}(y_0, \dots, y_j) = 2^{-N} \binom{N}{j} \xrightarrow{N \rightarrow \infty} 0. \quad (4.129)$$

Thus,  $\mu_n(\eta)$  can be made arbitrarily small if we take  $N$  large enough.  $\square$

Combining Theorem 4.1 and Proposition 4.14, we obtain Theorem 3.2 as an immediate corollary.

We can also obtain, as an easy corollary of Proposition 4.14, the following generalization of it:

**Corollary 4.14 (Cleaning by clouds).** *Let  $(\nu_n)_{n \geq 1}$  be a sequence of  $\Lambda$ -regular clouds satisfying*

$$\sum_{n=1}^{\infty} [1 - \nu_n((x))] = \infty \quad \text{for all } x \in \Lambda. \quad (4.130)$$

*Then  $\nu_1 * \dots * \nu_n$  converges pointwise to  $\pi_\Lambda$ .*

*Conversely, if  $\nu_1 * \dots * \nu_n$  converges pointwise to  $\pi_\Lambda$  and the  $(\nu_n)_{n \geq 1}$  are all supported on finitely many levels and belong to  $\mathcal{S}_1$  (this last is automatic if  $\Lambda \neq X$ ), then (4.130) holds.*

*Proof.* Define  $h_n(x) = 1 - \nu_n((x))$ . By Lemma 4.3 we have  $\nu_i \trianglelefteq \beta_{h_i}$ . By Corollary 4.7 we have  $\nu_1 * \dots * \nu_n \trianglelefteq \beta_{h_1} * \dots * \beta_{h_n}$ . The result then follows from Propositions 4.14 and 4.13.

The converse holds under the specified conditions, by exactly the same proof as in Proposition 4.14.  $\square$

*Remark.* For the converse, the hypothesis that the  $(\nu_n)_{n \geq 1}$  be supported on finitely many levels (and belong to  $\mathcal{S}_1$  if  $\Lambda = X$ ) is essential: a counterexample is  $\nu_1 = \pi_\Lambda$ ,  $\nu_2 = \nu_3 = \dots = I$ .

We can also weaken the condition on the sequence  $(\nu_n)$  if we look only at the dust that stays inside  $\Lambda$ :

**Corollary 4.15. (Generalized cleaning by clouds, behavior inside  $\Lambda$ ).**

Let  $(\nu_n)_{n \geq 1}$  be a sequence of nonnegative clouds that are carried by  $\Lambda$  and satisfy  $\|\nu_n\| \leq 1$  and

$$\sum_{n=1}^{\infty} [1 - \nu_n((x))] = \infty \quad \text{for all } x \in \Lambda. \quad (4.131)$$

Then  $\nu_1 * \dots * \nu_n * I_\Lambda$  converges pointwise to zero.

*Proof.* Note first that, by Lemma 4.6, the cloud  $\nu_1 * \dots * \nu_n * I_\Lambda$  is supported on markers having all their entries in  $\Lambda$ . Now define  $h_n(x) = 1 - \nu_n((x))$ ; again by Lemma 4.3 we have  $\nu_i \leq \beta_{h_i}$ , and by Corollary 4.7 we have  $\nu_1 * \dots * \nu_n \leq \beta_{h_1} * \dots * \beta_{h_n}$ . By Proposition 4.14, we have  $\beta_{h_1} * \dots * \beta_{h_n} \rightarrow \pi_\Lambda$  pointwise, so that

$$(\beta_{h_1} * \dots * \beta_{h_n})(\eta) \rightarrow 0 \quad (4.132)$$

for all markers  $\eta$  having all their entries in  $\Lambda$ . Since  $\nu_1 * \dots * \nu_n \leq \beta_{h_1} * \dots * \beta_{h_n}$ , it follows that, for such markers,

$$(\nu_1 * \dots * \nu_n)(\eta) \rightarrow 0 \quad (4.133)$$

and hence also  $(\nu_1 * \dots * \nu_n)(\eta) \rightarrow 0$ .  $\square$

*Remark.* Here there is no converse: one could take, for instance,  $\nu_1 = I_{\Lambda^c}$  and  $\nu_2 = \nu_3 = \dots = I$ .

#### 4.8.2. Uniform (operator-norm) version

Let us now consider the uniform (operator-norm) version of the convergence theorems. We shall assume that  $\mu_n$  converges to  $\mu_\infty$  uniformly on all markers of a given level, i.e.

$$\sup_{\text{level}(\eta)=\ell} |\mu_n(\eta) - \mu_\infty(\eta)| \xrightarrow{n \rightarrow \infty} 0 \quad (4.134)$$

for each  $\ell \geq 0$ . In this case it is not necessary to assume that the  $\mu_n$  are  $\Lambda$ -regular; it suffices to assume that they are carried by  $\Lambda$  and uniformly bounded in norm.

**Theorem 4.2. (Convergence theorem, uniform version).** (FH) *Let  $(\mu_n)_{n \geq 1}$  be a sequence of clouds carried by  $\Lambda$  and satisfying a uniform bound  $\|\mu_n\| \leq M < \infty$ . Suppose further that*

$$\mu_n(\eta) \xrightarrow[n \rightarrow \infty]{} \mu_\infty(\eta) \quad \text{uniformly for all } \eta \text{ of a given level} \quad (4.135)$$

and that

$$\|(I_\Lambda \alpha I_\Lambda)^k\|_{w \rightarrow w} \xrightarrow[k \rightarrow \infty]{} 0. \quad (4.136)$$

Then

$$\|T_{\mu_n} - T_{\mu_\infty}\|_{w \rightarrow w} \xrightarrow[n \rightarrow \infty]{} 0. \quad (4.137)$$

*Proof.* We proceed as in the proof of Theorem 4.1, but here the reasoning is much simpler since the uniform-convergence hypothesis (4.135) allows us to avoid the introduction of the subset  $\Lambda'$ . We begin by observing that  $\mu_\infty$  is carried by  $\Lambda$  and satisfies

$$\|\mu_\infty\| \leq \liminf_{n \rightarrow \infty} \|\mu_n\| \leq M. \quad (4.138)$$

Now fix any  $\varepsilon > 0$ , and choose  $k$  so that

$$\|(I_\Lambda \alpha I_\Lambda)^k\|_{w \rightarrow w} \leq \varepsilon/M. \quad (4.139)$$

We divide the set of all markers into two classes:

Class 1: Markers of level  $\leq k$ .

Class 2: Markers of level  $> k$ .

We decompose any cloud  $\mu_*$  ( $\mu_* = \mu_n$  or  $\mu_\infty$ ) in the form  $\mu_*^1 + \mu_*^2$  so that  $\mu_*^i$  charges only markers of class  $i$ . The triangle inequality yields

$$\|T_{\mu_n - \mu_\infty}\|_{w \rightarrow w} \leq \|T_{|\mu_n^1 - \mu_\infty^1|}\|_{w \rightarrow w} + \|T_{|\mu_n^2|}\|_{w \rightarrow w} + \|T_{|\mu_\infty^2|}\|_{w \rightarrow w}, \quad (4.140)$$

and these contributions may be bounded separately as follows:

Class 1. The cloud  $|\mu_n^1 - \mu_\infty^1|$  charges only markers of level  $\leq k$ . By (4.135), if  $n$  is large enough we have

$$|\mu_n^1 - \mu_\infty^1| \leq \frac{\varepsilon}{k+1} (\rho^0 + \rho^1 + \dots + \rho^k), \quad (4.141)$$

so that  $\| |\mu_n^1 - \mu_\infty^1| \| \leq \varepsilon$ . Therefore, by Proposition 4.1,  $\|T_{|\mu_n^1 - \mu_\infty^1|}\|_{w \rightarrow w} \leq \varepsilon$  as well.

Class 2. Since  $\mu_*$  ( $= \mu_n$  or  $\mu_\infty$ ) is carried by  $\Lambda$ , any marker  $\eta = (x_0, \dots, x_\ell)$  charged by  $\mu_*^2$  must have  $\ell > k$  and  $x_0, \dots, x_k \in \Lambda$ . Since  $\|\mu_*\| \leq M$ , it follows that

$$|\mu_*^2| \leq M(I_\Lambda * \rho^1 * I_\Lambda)^k. \quad (4.142)$$

By Proposition 4.10 we have  $\|T_{|\mu_*^2|}\|_{w \rightarrow w} \leq M\|(I_\Lambda \alpha I_\Lambda)^k\|_{w \rightarrow w} \leq \varepsilon$  by (4.139).  $\square$

In the uniform case we have the following analogue of Proposition 4.14:

**Proposition 4.15. (Conditions on cleaning sequences, uniform version).** *Let  $(h_n)_{n \geq 1}$  be a sequence of functions satisfying  $0 \leq h_n \leq \chi_\Lambda$ , and define*

$$\mu_n := \beta_{h_1} * \dots * \beta_{h_n}. \quad (4.143)$$

Then the following are equivalent:

- (a)  $\mu_n$  converges uniformly-on-levels to  $\pi_\Lambda$ .
- (b)  $\sum_{i=1}^n h_i$  converges uniformly (as  $n \rightarrow \infty$ ) to  $+\infty$  on  $\Lambda$ .
- (b') There exist  $\delta > 0$  and a sequence  $0 = n_0 < n_1 < n_2 < \dots$  satisfying

$$\sum_{i=n_{j-1}+1}^{n_j} h_i \geq \delta \chi_\Lambda \quad (4.144)$$

for all  $j$ .

- (b'') For all  $\delta < \infty$ , there exists a sequence  $0 = n_0 < n_1 < n_2 < \dots$  satisfying (4.144) for all  $j$ .

*Proof.* Since the  $h_i$  are bounded, it is easy to see that (b), (b') and (b'') are all equivalent.

The proof of (b'')  $\implies$  (a) is a straightforward adaptation of the corresponding proof in Proposition 4.14. Indeed, we want to show here that the convergence (4.123) holds uniformly for all markers of a given level. With the hypothesis (b''), we can repeat the end of the proof of Proposition 4.14 with  $\Lambda' = \Lambda$  instead of (4.124), which gives the result.

Conversely, let us prove that (a)  $\implies$  (b). Suppose that  $\sum_{i=1}^n h_i$  does not converge uniformly to  $+\infty$  on  $\Lambda$ . Then there exists  $M < \infty$  such that, for any  $n \geq 1$ , we can find  $x \in \Lambda$  with

$$\sum_{i=1}^n h_i(x) < M. \quad (4.145)$$

For such  $n$  and  $x$ , we consider the markers  $\eta$  whose entries are all equal to  $x$ , and we denote by  $\eta_k$  the only such marker that lies on level  $k$ . For  $0 \leq j \leq n$  and  $k \geq 0$ , let us define

$$p_j(k) := \mu_j(\eta_k), \quad (4.146)$$

where  $\mu_0 := \rho^0$ . We have  $p_0(0) = 1$ ,  $p_0(k) = 0$  for each  $k > 0$ , and for  $j \geq 1$

$$p_j(k) = [1 - h_j(x)]p_{j-1}(k) + h_j(x)p_{j-1}(k-1). \quad (4.147)$$

(The preceding equation holds also for  $k = 0$  if for each  $j$  we set  $p_j(-1) := 0$ .) Observe that for each  $j$  we always have  $\sum_k p_j(k) = 1$ . Let us introduce the barycenter

$$G_j := \sum_k k p_j(k). \quad (4.148)$$

We obviously have  $G_0 = 0$ , and a direct calculation gives, for  $j \geq 1$ ,

$$G_j = G_{j-1} + h_j(x). \quad (4.149)$$

Therefore, we have

$$G_n = \sum_{j=1}^n h_j(x) < M. \quad (4.150)$$

By the Markov inequality, this implies that

$$\sum_{k=0}^{2M} \mu_n(\eta_k) = \sum_{k=0}^{2M} p_n(k) \geq \frac{1}{2}, \quad (4.151)$$

which clearly prevents  $\mu_n$  from converging uniformly to  $\pi_\Lambda$ .  $\square$

Combining Theorem 4.2 and Proposition 4.15, we obtain Theorem 3.1 as an immediate corollary.

### 4.8.3. Some final remarks

There are three natural topologies for convergence of clouds:  $\|\cdot\|$  norm; uniform-on-levels; and pointwise. Norm convergence of the  $\mu_n$  implies operator-norm convergence of the  $T_{\mu_n}$  without any hypothesis on  $\alpha$  beyond the Fundamental Hypothesis (cf. Proposition 4.1). Uniform-on-levels convergence is weaker than norm convergence in that the mass of  $\mu_n - \mu_\infty$  can run “upwards to infinity”, as in the example  $\beta_\Lambda^n \rightarrow \pi_\Lambda$ . This is handled in Theorem 4.2 by assuming that  $\|(I_\Lambda \alpha I_\Lambda)^k\|_{w \rightarrow w} \rightarrow 0$ ; then the  $T_{\mu_n}$  converge in operator norm. Finally, pointwise convergence is still weaker than uniform-on-levels convergence because mass can also run “outwards to infinity” (when  $X$  is infinite), as in the example  $\beta_{\Lambda_n} \rightarrow \beta_\Lambda$  with  $\Lambda_n \uparrow \Lambda$  and all  $\Lambda_n \subsetneq \Lambda$ . This is handled in Theorem 4.1 by working on a fixed vector  $c \in l^1(w)$  and demanding convergence only in vector norm (not operator norm); the upwards-running mass is handled by assuming that  $\|c(I_\Lambda \alpha I_\Lambda)^k\|_w \rightarrow 0$ .

### 4.9. Some further identities and inequalities

Let us now present some further identities and inequalities that will play an important role in Section 6. Our first result is a cloud analogue and extension of Lemmas 3.8 and 3.9:

**Lemma 4.13.**

(a) Let  $h: X \rightarrow [0, 1]$  and set  $\Lambda := \text{supp } h$ . Then we have the identity

$$\sum_{k=0}^{\infty} \beta_h^{*k} * I_h = \mathbf{1}_\Lambda. \quad (4.152)$$

[Recall that  $\mathbf{1}_\Lambda$  is the cloud that takes the value 1 on markers having all their entries in  $\Lambda$ , and 0 elsewhere.]

(b) Let  $\nu \geq 0$  be a cloud satisfying  $\|\nu\| \leq 1$ . Set  $h(x) := 1 - \nu(x)$ . Then we have the inequality

$$\sum_{k=0}^{\infty} \nu^{*k} * I_h \leq \mathbf{1}. \quad (4.153)$$

*Proof.* (a) The proof of Lemma 3.8 is in fact a proof of the cloud identity (4.152): no more need be said. But just for completeness, here is an alternate proof: Let  $\mu$  be any cloud; then by (4.36) we have

$$\mu * I_h = \widetilde{\mu} - \widetilde{\mu * \beta_h}. \quad (4.154)$$

Now set  $\mu = \beta_h^{*k}$  and sum from  $k = 0$  to  $N$ : we get

$$\sum_{k=0}^N \beta_h^{*k} * I_h = \mathbf{1} - \widetilde{\beta_h^{*(N+1)}} \quad (4.155)$$

(since  $\widetilde{\beta_h^{*0}} = \widetilde{\rho^0} = \mathbf{1}$ ). Now take  $N \rightarrow \infty$ : by Proposition 4.14,  $\beta_h^{*(N+1)}$  converges pointwise to  $\pi_\Lambda$ , hence  $\widetilde{\beta_h^{*(N+1)}}$  converges pointwise to  $\widetilde{\pi_\Lambda} = \mathbf{1} - \mathbf{1}_\Lambda$ . This proves (4.152).

(b) By Corollary 4.4 we have, for any cloud  $\mu \geq 0$ ,

$$\widetilde{(\mu * \nu)}(\eta) \leq \widetilde{\mu}(\eta^-) + \mu(\eta)\nu(\text{last}(\eta)) = \widetilde{\mu}(\eta) - [1 - \nu(\text{last}(\eta))] \mu(\eta) \quad (4.156)$$

and hence

$$\mu * I_h \leq \widetilde{\mu} - \widetilde{\mu * \nu}. \quad (4.157)$$

Now take  $\mu = \nu^{*k}$  and sum from  $k = 0$  to  $N$ :

$$\sum_{k=0}^N \nu^{*k} * I_h \leq \mathbf{1} - \widetilde{\nu^{*(N+1)}} \leq \mathbf{1} \quad (4.158)$$

(since  $\widetilde{\nu^{*0}} = \widetilde{\rho^0} = \mathbf{1}$ ). Taking  $N \rightarrow \infty$  gives the result.  $\square$

We also have the following curious collection of inequalities and identity involving the cloud  $\mu := \mathbf{1} - \tilde{\nu}$ . These too will play a central role in Section 6; what makes them so powerful is that the inequalities go in *both* directions.

**Lemma 4.14.** *Let  $\nu \geq 0$  be a cloud satisfying  $\|\nu\| \leq 1$ . Define  $h(x) := 1 - \nu(x)$  and  $\mu := \mathbf{1} - \tilde{\nu}$ .*

(a) *We have  $I_h \leq \mu \leq \mathbf{1}$  and*

$$\sum_{n=0}^{\infty} \nu^{*n} * \mu \leq \mathbf{1}. \quad (4.159)$$

(b) *If  $\nu$  is carried by  $\Lambda$  and satisfies  $\nu(x) < 1$  for all  $x \in \Lambda$  [i.e.  $\text{supp } h = \Lambda$ ], then*

$$\sum_{n=0}^{\infty} \nu^{*n} * \mu \geq \mathbf{1}_{\Lambda}. \quad (4.160)$$

(c) *If  $\nu$  is  $\Lambda$ -regular and satisfies  $\nu(x) < 1$  for all  $x \in \Lambda$ , then  $\mu$  is supported on markers having all their entries in  $\Lambda$  (so that  $I_h \leq \mu \leq \mathbf{1}_{\Lambda}$ ) and*

$$\sum_{n=0}^{\infty} (I_{\Lambda} * \nu * I_{\Lambda})^{*n} * \mu = \sum_{n=0}^{\infty} \nu^{*n} * \mu = \mathbf{1}_{\Lambda}. \quad (4.161)$$

(d) *If  $\nu$  belongs to  $\mathcal{S}_1$  and is supported on levels  $\leq K$ , then  $\mu$  is supported on levels  $\leq K - 1$ .*

Please note that since  $\mu \geq I_h$ , Lemma 4.14(a) is a strengthening of Lemma 4.13(b). Furthermore, if  $\nu = \beta_h$ , then  $\mu = I_h$ , so that Lemma 4.14(c) is a direct generalization of Lemma 4.13(a).

*Proof of Lemma 4.14.* (a) Since  $\|\nu\| \leq 1$ , we have  $\mu \geq 0$ . Since  $\nu \geq 0$ , we have  $\mu \leq \mathbf{1}$ . Finally, since  $\mu(x) = 1 - \nu(x) = h(x)$ , we have  $\mu \geq I_h$ .

Let us now observe that  $\mu := \mathbf{1} - \tilde{\nu} = (\rho^0 - \nu) * \mathbf{1}$  and perform the following calculation:

$$\sum_{n=0}^{N-1} \nu^{*n} * \mu = \sum_{n=0}^{N-1} \nu^{*n} * (\rho^0 - \nu) * \mathbf{1} = (\rho^0 - \nu^{*N}) * \mathbf{1} = \mathbf{1} - \widetilde{\nu^{*N}} \quad (4.162)$$

$$\leq \mathbf{1}. \quad (4.163)$$

Now take  $N \rightarrow \infty$ : since  $\mu, \nu \geq 0$ , the left-hand side increases pointwise to a limiting cloud  $\sum_{n=0}^{\infty} \nu^{*n} * \mu$  whose values lie in  $[0, 1]$  by (4.163).

(b) Now suppose that  $\nu$  is carried by  $\Lambda$  and that  $\nu(x) < 1$  for all  $x \in \Lambda$ . Then Corollary 4.15 implies that  $\nu^{*N}(\sigma) \rightarrow 0$  for all markers  $\sigma$  ending in  $\Lambda$ ,



hence  $\widetilde{\nu^{*N}}(\eta) \rightarrow 0$  for all markers  $\eta$  having all their entries in  $\Lambda$ . By (4.162) this implies (4.160).

(c) Since  $\nu$  is  $\Lambda$ -regular, we have  $\widetilde{\nu}(\eta) = 1$  whenever  $\eta$  has at least one entry outside  $\Lambda$ ; hence  $\mu$  is supported on markers having all their entries in  $\Lambda$ . Since in addition  $\nu((x)) < 1$  for all  $x \in \Lambda$ , we can apply Corollary 4.14 to conclude that  $\nu^{*N} \rightarrow \pi_\Lambda$  pointwise, i.e.  $\widetilde{\nu^{*N}} \rightarrow \widetilde{\pi_\Lambda}$  pointwise. But

$$\widetilde{\pi_\Lambda}(\eta) = \begin{cases} 0 & \text{if } \eta \text{ has all its entries in } \Lambda, \\ 1 & \text{otherwise,} \end{cases} \quad (4.164)$$

so that  $\widetilde{\pi_\Lambda} = \mathbf{1} - \mathbf{1}_\Lambda$ . By (4.162) this proves the second equality in (4.161). Since  $\mu = I_\Lambda * \mu * I_\Lambda$ , we have  $\nu^{*n} * \mu = (I_\Lambda * \nu * I_\Lambda)^{*n} * \mu$  by Lemma 4.6, which proves the first equality.

(d) If  $\nu \in \mathcal{S}_1$  is supported on levels  $\leq K$ , it follows that  $\widetilde{\nu}(\eta) = 1$  whenever  $\text{level}(\eta) \geq K$ , hence  $\mu$  is supported on levels  $\leq K - 1$ .  $\square$

*Remark.* At a formal level, the computation (4.162)–(4.163) is inspired by the (admittedly meaningless) “identity”

$$\sum_{n=0}^{\infty} \nu^{*n} = (\rho^0 - \nu)^{-1}.$$

Another, more physical, way of expressing the intuition behind (4.162)–(4.163) is to observe that

$$\nu^{*n} * \mu = \widetilde{\nu^{*n}} - \widetilde{\nu^{*(n+1)}}, \quad (4.165)$$

so that  $(\nu^{*n} * \mu)(\eta)$  measures the mass that lies  $\preceq \eta$  after  $n$  steps but gets pushed above (or out) at the  $(n + 1)$ st step. Summing over  $n$ , we should get 1 whenever  $\eta$  has all its entries in  $\Lambda$ , since by Corollary 4.15 all the mass should eventually be pushed out.

## 5. Alternate sufficient conditions for cleanability

*In this section we do not assume the Fundamental Hypothesis.* Rather, our goal is to examine briefly the conditions under which cleaning can be assured even in the absence of the Fundamental Hypothesis. Please note that we are entitled to use here those results of Section 4 that refer only to clouds. However, we must scrupulously avoid using those results which, like Propositions 4.1 and 4.10, refer to operators and therefore depend on the Fundamental Hypothesis.

First, we need a few definitions. Let  $\mathbf{a} = (a_{xy})_{x,y \in X}$  be any nonnegative matrix (in practice, we will take  $\mathbf{a}$  to be either  $\alpha$  or  $I_\Lambda \alpha I_\Lambda$ ). For any marker  $\eta = (x_0, \dots, x_k)$ , we define

$$\mathbf{a}^\eta := a_{x_0 x_1} a_{x_1 x_2} \cdots a_{x_{k-1} x_k}. \quad (5.1)$$

We then define the spaces of clouds  $l^1(\mathbf{a})$  and  $l^\infty(\mathbf{a})$  by the norms

$$\|\nu\|_{1,\mathbf{a}} := \sum_{\eta} \mathbf{a}^{\eta} |\nu(\eta)|, \quad (5.2)$$

$$\|\nu\|_{\infty,\mathbf{a}} := \sup_{\eta: \mathbf{a}^{\eta} > 0} |\nu(\eta)|. \quad (5.3)$$

If  $\mu$  and  $\nu$  are clouds, we write

$$\langle \mu, \nu \rangle_{\mathbf{a}} := \sum_{\eta} \mathbf{a}^{\eta} \mu(\eta) \nu(\eta) \quad (5.4)$$

whenever this sum has an unambiguous meaning; in particular, if  $\mu, \nu \geq 0$ , then  $\langle \mu, \nu \rangle_{\mathbf{a}}$  is well-defined though it may be  $+\infty$ .

If  $\nu$  is a cloud, we would like to define an operator  $T_{\nu,\mathbf{a}}$  by the usual formula

$$T_{\nu,\mathbf{a}} := \sum_{\eta=(x_0,\dots,x_k)} \nu(\eta) I_{\{x_0\}} \mathbf{a} I_{\{x_1\}} \mathbf{a} \cdots \mathbf{a} I_{\{x_{k-1}\}} \mathbf{a} I_{\{x_k\}}. \quad (5.5)$$

[When  $\mathbf{a} = \alpha$ , we write simply  $T_{\nu}$ .] The trouble is that, in the absence of the Fundamental Hypothesis or its equivalent for  $\mathbf{a}$ , it is difficult to guarantee that this sum is well-defined (compare Proposition 4.1). We therefore restrict ourselves in this section to *nonnegative* clouds  $\nu$  and consider (5.5) as defining a matrix  $T_{\nu,\mathbf{a}}$  whose elements lie in  $[0, +\infty]$ . Of course, when necessary we shall verify *a posteriori* that the elements of  $T_{\nu,\mathbf{a}}$  are finite.

Finally, if  $c \geq 0$  and  $w \geq 0$  are vectors, we define a nonnegative cloud  $\mu_{c,w}$  by

$$\mu_{c,w}((x_0, \dots, x_k)) = c_{x_0} w_{x_k}. \quad (5.6)$$

It is easily checked that for any nonnegative cloud  $\nu$  we have

$$c T_{\nu,\mathbf{a}} w = \langle \mu_{c,w}, \nu \rangle_{\mathbf{a}}. \quad (5.7)$$

(We define  $0 \cdot \infty = 0$ ; this may be needed to interpret the inner product  $c T_{\nu,\mathbf{a}} w$ .) In particular, taking  $\nu = \rho^k$ , we have

$$c \mathbf{a}^k w = \sum_{\text{level}(\eta)=k} \mathbf{a}^{\eta} \mu_{c,w}(\eta) \quad (5.8)$$

and hence that

$$\sum_{k=0}^{\infty} c \mathbf{a}^k w = \|\mu_{c,w}\|_{1,\mathbf{a}}. \quad (5.9)$$

The following theorem and its corollary give a sufficient condition for a dust vector  $c$  to be cleanable in  $l^1(w)$  sense by every cleaning process that visits each site infinitely many times. The proofs are almost trivial, now that we have the results of Section 4 in hand.

**Theorem 5.1.** *Let  $\mathbf{a}$  be a nonnegative matrix, let  $\mu$  be a nonnegative cloud belonging to  $l^1(\mathbf{a})$ , and let  $(\nu_n)_{n \geq 1}$  be a sequence of nonnegative clouds satisfying a uniform bound  $\|\nu_n\|_{\infty, \mathbf{a}} \leq M < \infty$  and tending pointwise  $\mathbf{a}$ -a.e. to zero [i.e.  $\lim_{n \rightarrow \infty} \nu_n(\eta) = 0$  for each marker  $\eta$  having  $\mathbf{a}^\eta > 0$ ]. Then*

$$\lim_{n \rightarrow \infty} \langle \mu, \nu_n \rangle_{\mathbf{a}} = 0. \quad (5.10)$$

*Proof.* An immediate consequence of Lebesgue's dominated convergence theorem.  $\square$

**Corollary 5.1 (Sufficient condition for universal cleaning).** *Let  $\Lambda \subseteq X$ , and let  $c \geq 0$  and  $w \geq 0$  be vectors satisfying*

$$\sum_{k=0}^{\infty} c(I_{\Lambda} \alpha I_{\Lambda})^k w < \infty. \quad (5.11)$$

*If  $(f_n)_{n \geq 1}$  are functions  $X \rightarrow [0, 1]$ , supported in  $\Lambda$  and satisfying  $\sum_n f_n = +\infty$  everywhere on  $\Lambda$ , we have*

$$\lim_{n \rightarrow \infty} c \beta_{f_1} \cdots \beta_{f_n} I_{\Lambda} w = 0. \quad (5.12)$$

*More generally, if  $(\omega_n)_{n \geq 1}$  are nonnegative clouds carried by  $\Lambda$  that satisfy  $\|\omega_n\| \leq 1$  and  $\sum_n [1 - \omega_n((x))] = +\infty$  for all  $x \in \Lambda$ , we have*

$$\lim_{n \rightarrow \infty} c T_{\omega_1} \cdots T_{\omega_n} I_{\Lambda} w = 0. \quad (5.13)$$

*Proof.* This is simply the special case of Theorem 5.1 in which  $\mathbf{a} = I_{\Lambda} \alpha I_{\Lambda}$ ,  $\mu = \mu_{c,w}$  and  $\nu_n = \omega_1 * \cdots * \omega_n * I_{\Lambda}$ : by (5.9), the hypothesis (5.11) guarantees that  $\mu_{c,w} \in l^1(\mathbf{a})$ . [Using Lemma 4.6 to rewrite  $\nu_n$  as  $I_{\Lambda} * \omega_1 * I_{\Lambda} * \cdots * I_{\Lambda} * \omega_n * I_{\Lambda}$ , we see that  $\langle \mu_{c,w}, \nu_n \rangle_{\alpha} = \langle \mu_{c,w}, \nu_n \rangle_{I_{\Lambda} \alpha I_{\Lambda}}$ .] We can take  $M = 1$ . The fact that  $\nu_n(\eta) \rightarrow 0$  for all  $\eta$  is an immediate consequence of Corollary 4.15.  $\square$

In view of Theorem 4.1, it is natural to wonder whether the hypothesis  $\sum_{k=0}^{\infty} c \alpha^k w < \infty$  in Corollary 5.1 can be weakened to assuming that  $c \alpha^k w$  is finite for every  $k$  and tends to zero as  $k \rightarrow \infty$ . It turns out that this is not the case; indeed, the following example shows that Corollary 5.1 is in a certain sense sharp:

**Example 5.1.** Let  $(\rho_k)_{k \geq 0}$  be any sequence of positive numbers satisfying  $\sum_{k=0}^{\infty} \rho_k = \infty$ . Then, on a countably infinite state space  $X$ , we can construct a matrix  $\alpha \geq 0$  and vectors  $c, w \geq 0$  such that

- (a)  $c \alpha^k w = \rho_k$  for every  $k \geq 0$ .
- (b) There exists a sequence  $x_1, x_2, \dots \in X$  in which each element of  $X$  occurs infinitely often, and for which we have  $\lim_{n \rightarrow \infty} c \beta_{x_1} \cdots \beta_{x_n} w = \infty$ .

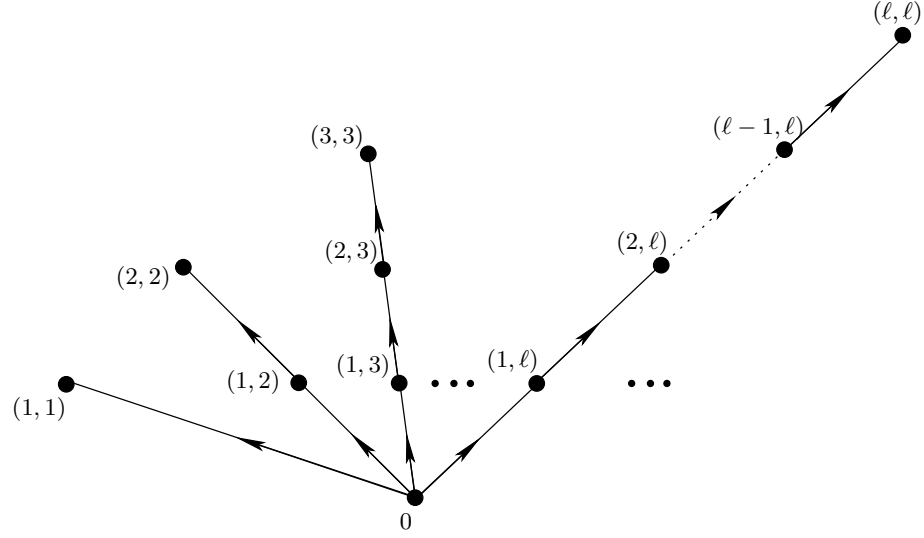


Figure 4. The directed graph associated to the matrix  $\alpha$  in Example 5.1. The vertices of this graph correspond to the points of  $X$ , and the edges correspond to the nonzero matrix elements of  $\alpha$ .

To see this, take  $X = \{0\} \cup \{(k, \ell) : 1 \leq k \leq \ell\}$ , and define the matrix  $\alpha$  by

$$\alpha_{0,(1,\ell)} = \rho_1 / (2^\ell \rho_0), \quad (5.14)$$

$$\alpha_{(k,\ell),(k+1,\ell)} = 2\rho_{k+1} / \rho_k \quad \text{for } 1 \leq k \leq \ell - 1, \quad (5.15)$$

with all other coefficients of  $\alpha$  set to 0. The state space  $X$  can be visualized as a tree with root 0 and branches numbered  $\ell = 1, 2, \dots$  (see Figure 4). Now consider the vector  $c = \rho_0 \delta_0$ , which puts a mass  $\rho_0$  of dirt on site 0 and nothing elsewhere, and the vector  $w = \mathbf{1}$ . It is easy to see that, for each  $k \geq 1$ , the vector  $c\alpha^k$  is supported on sites  $(k, \ell)$  [ $\ell \geq k$ ] and takes there the values  $\rho_k / 2^{\ell-k+1}$ . Hence  $c\alpha^k w = \rho_k$  for all  $k$ . In order to construct the sequence  $x_1, x_2, \dots$ , let us first consider the following two sequences of sites:

$$\begin{aligned} A_L &:= (1, 2), (1, 3), (2, 3), (1, 4), (2, 4), (3, 4), (1, L), \dots, (L-1, L), \\ B_L &:= (1, 1), (2, 2), (3, 3), \dots, (L, L) \end{aligned}$$

so that  $A_L$  (resp.  $B_L$ ) sweeps, in order, the non-summit (resp. summit) sites of the branches 1 through  $L$ . Now let us define our sweeping process  $x_1, x_2, \dots$  as follows:

$$0, A_{L_1}, B_1, A_{L_2}, B_2, A_{L_3}, B_3, \dots$$

where the indices  $L_1 < L_2 < \dots$  will be chosen in a moment. After the first step ( $x_1 = 0$ ), we have a mass  $\rho_1 / 2^\ell$  of dust on each site  $(1, \ell)$ . After the

sequence  $A_{L_1}$ , we end up with a mass  $\rho_\ell/2$  on each site  $(\ell, \ell)$  for  $1 \leq \ell \leq L_1$ . The step  $B_1$  then destroys the mass  $\rho_1/2$  that sat on the site  $(1, 1)$ . After the sequence  $A_{L_2}$ , we end up with a mass  $\rho_\ell/2$  on each site  $(\ell, \ell)$  for  $2 \leq \ell \leq L_2$ . The step  $B_2$  then destroys the mass  $\rho_2/2$  that sat on the site  $(2, 2)$ . And so forth. If we choose each  $L_r$  so that

$$\sum_{\ell=r}^{L_r} \frac{\rho_\ell}{2} \geq 10^r + \frac{\rho_r}{2}, \quad (5.16)$$

then we are assured of having a total mass at least  $10^r$  on the summit sites at each step after the operation  $A_{L_r}$  has been completed. Hence  $c\beta_{x_1} \cdots \beta_{x_n} w \rightarrow \infty$ .

This example shows that  $c\alpha^k w$  being finite for every  $k$  and tending to zero as  $k \rightarrow \infty$  is not sufficient to ensure that *every* cleaning process that visits each site infinitely many times will succeed in removing the dirt in  $l^1(w)$  sense. Nevertheless, this hypothesis turns out to be sufficient to ensure that there *exists* a successful cleaning process:

**Theorem 5.2 (Sufficient condition for existence of cleaning).** *Let  $\Lambda \subseteq X$ , and let  $c \geq 0$  and  $w \geq 0$  be vectors satisfying*

$$(a) \quad c(I_\Lambda \alpha I_\Lambda)^k w < \infty \text{ for all } k \geq 0$$

and

$$(b) \quad \lim_{k \rightarrow \infty} c(I_\Lambda \alpha I_\Lambda)^k w = 0.$$

*Then it is possible to find a sequence of sites  $x_1, x_2, \dots \in \Lambda$  and a sequence of numbers  $\varepsilon_1, \varepsilon_2, \dots \in (0, 1]$  such that  $\lim_{n \rightarrow \infty} c\beta_{\varepsilon_1 \delta_{x_1}} \cdots \beta_{\varepsilon_n \delta_{x_n}} I_\Lambda w = 0$ .*

**Question 5.1.** Can we always take  $\varepsilon_i = 1$ ?

The proof of this theorem will be based on a simple lemma:

**Definition 5.1.** We say that a cloud  $\nu' \geq 0$  arises from a cloud  $\nu \geq 0$  by a *single-marker update acting at the marker  $\eta$*  if there exists a real number  $\kappa \in [0, \nu(\eta)]$  such that

- $\nu'(\eta) = \nu(\eta) - \kappa$ ,
- $\nu'(\sigma) = \nu(\sigma) + \kappa$  for each child  $\sigma$  of  $\eta$ ,
- $\nu'(\tau) = \nu(\tau)$  for every other marker  $\tau$ .

**Lemma 5.1 (Imitation Lemma).** *Suppose that the cloud  $\nu' \geq 0$  arises from the cloud  $\nu \geq 0$  by a single-marker update acting at the marker  $\eta$ . Let  $y = \text{last}(\eta)$ . Then, for any vector  $c \geq 0$ , we can find a number  $\varepsilon \in [0, 1]$  (depending on  $c, \nu$  and  $\nu'$ ) such that*

$$cT_{\nu'} = cT_\nu \beta_{\varepsilon \delta_y}. \quad (5.17)$$

*Proof.* Let  $\kappa = \nu(\eta) - \nu'(\eta)$ . If  $\kappa = 0$ , we can take  $\varepsilon := 0$ ; so let us assume henceforth that  $\nu(\eta) \geq \kappa > 0$ . We then have, for all  $x \in X$ ,

$$(cT_\nu \beta_{\varepsilon \delta_y})_x = \begin{cases} (cT_\nu)_x + \varepsilon (cT_\nu)_y \alpha_{yx} & \text{if } x \neq y, \\ (1 - \varepsilon)(cT_\nu)_y + \varepsilon (cT_\nu)_y \alpha_{yy} & \text{if } x = y, \end{cases} \quad (5.18)$$

whereas

$$(cT_\nu)_x = \begin{cases} (cT_\nu)_x + \kappa (cT_\eta)_y \alpha_{yx} & \text{if } x \neq y, \\ (cT_\nu)_y - \kappa (cT_\eta)_y + \kappa (cT_\eta)_y \alpha_{yy} & \text{if } x = y. \end{cases} \quad (5.19)$$

Therefore, to satisfy (5.17), we proceed as follows: if  $(cT_\eta)_y = 0$  we set  $\varepsilon := 0$ ; if  $(cT_\eta)_y > 0$ , we set  $\varepsilon := \kappa \{(cT_\eta)_y / (cT_\nu)_y\}$ . In the latter case we have  $(cT_\nu)_y \geq \nu(\eta)(cT_\eta)_y \geq \kappa (cT_\eta)_y > 0$ , so that  $\varepsilon \leq 1$ .  $\square$

*Proof of Theorem 5.2.* Since  $\beta_{\varepsilon_1 \delta_{x_1}} \cdots \beta_{\varepsilon_n \delta_{x_n}} I_\Lambda = I_\Lambda \beta_{\varepsilon_1 \delta_{x_1}} I_\Lambda \cdots I_\Lambda \beta_{\varepsilon_n \delta_{x_n}} I_\Lambda$  by Lemma 3.1, everything in both hypothesis and conclusion takes place within  $\Lambda$ , so we can assume for notational simplicity that  $\Lambda = X$ .

By virtue of the Imitation Lemma, it suffices to find a sequence  $(\nu_n)_{n \geq 0}$  of nonnegative clouds, with  $\nu_0 = \rho^0$ , such that each  $\nu_{n+1}$  arises from  $\nu_n$  by a single-marker update, and which satisfy

$$cT_{\nu_n} w \xrightarrow[n \rightarrow \infty]{} 0. \quad (5.20)$$

We shall construct these clouds to be  $\{0, 1\}$ -valued and lie in  $\mathcal{S}_1$ . In particular, updating the marker  $\eta$  at step  $n$  will mean that  $\nu_n(\eta) = 1$  and  $\nu_{n+1}(\eta) = 0$ , and for each child  $\sigma$  of  $\eta$ ,  $\nu_n(\sigma) = 0$  and  $\nu_{n+1}(\sigma) = 1$ .

Since for each  $k \geq 0$  we have

$$c\alpha^k w = \sum_{\text{level}(\eta)=k} \alpha^\eta \mu_{c,w}(\eta) < \infty, \quad (5.21)$$

for each  $\varepsilon > 0$  we can find a finite subset  $M_{\varepsilon,k}$  of the set of all markers of level  $k$ , such that

$$\sum_{\substack{\text{level}(\eta)=k \\ \eta \notin M_{\varepsilon,k}}} \alpha^\eta \mu_{c,w}(\eta) < \frac{\varepsilon}{2^{k+1}}. \quad (5.22)$$

Furthermore, we can arrange for these subsets to be nested, i.e.  $M_{\varepsilon,k} \subseteq M_{\varepsilon',k}$  whenever  $\varepsilon > \varepsilon'$ .

We now fix a sequence  $(\varepsilon_j)$  of positive numbers decreasing to 0, and we choose a sequence of integers  $N_1 < N_2 < \dots$  such that  $c\alpha^k w \leq \varepsilon_j$  whenever  $k > N_j$ . Let  $S_{j,\ell}$  denote the set of all markers of level  $\ell$  which are ancestors of any marker in  $\cup_{k=0}^{N_j} M_{\varepsilon_j,k}$ . By construction, all these sets are finite. We then define the sequence  $(\nu_n)_{n \geq 1}$  of clouds by a sequence of single-marker updates starting from  $\nu_0 = \rho^0$ , visiting the markers in the following order:

$$S_{1,0}, S_{1,1}, \dots, S_{1,N_1}, S_{2,0}, S_{2,1}, \dots, S_{2,N_2}, \dots$$

(the markers within each  $S_{j,\ell}$  being updated in an arbitrary order).

Let now  $j \geq 1$ , and consider any stage in this process while we are updating the markers in  $\cup_{\ell=0}^{N_{j+1}} S_{j+1,\ell}$ . At such a stage, the cloud  $\nu_n$  is always supported on three kinds of markers:

- Markers lying on levels  $k \leq N_j$  but outside  $M_{\varepsilon_j,k}$ . [Since at stage  $j$  we updated all the ancestors of markers in  $\cup_{k=0}^{N_j} M_{\varepsilon_j,k}$ , it is impossible for subsequent operations to place mass on any of those markers.]
- Markers lying on two successive levels  $N$  and  $N + 1$  strictly above  $N_j$ . [This happens while we are updating markers in  $S_{j+1,N}$ .]
- Markers lying on levels  $N_j < k \leq N_{j+1}$  but outside  $M_{\varepsilon_{j+1},k}$ .

The first kind of markers contributes to  $cT_{\nu_n}w$  at most  $\varepsilon_j$ , the second at most  $2\varepsilon_j$ , and the third at most  $\varepsilon_{j+1}$ . This proves that (5.20) holds.  $\square$

The following example shows that there cannot be any converse to Theorem 5.2 that refers only to the behavior of the sequence  $(c\alpha^k w)_{k \geq 0}$ :

**Example 5.2.** Let  $(\rho_k)_{k \geq 0}$  be *any* sequence of positive numbers (in particular, it can tend to  $+\infty$ , or oscillate, or whatever). Then, on a countably infinite state space  $X$ , we can construct a matrix  $\alpha \geq 0$  and vectors  $c, w \geq 0$  such that

- (a)  $c\alpha^k w = \rho_k$  for every  $k \geq 0$ .
- (b) There exists a sequence  $y_1, y_2, \dots \in X$  for which  $\lim_{n \rightarrow \infty} c\beta_{y_1} \cdots \beta_{y_n} w = 0$ .

We shall use the same state space  $X = \{0\} \cup \{(k, \ell) : 1 \leq k \leq \ell\}$  as in Example 5.1 (cf. Figure 4) but shall make a slightly different choice of the matrix  $\alpha$ .

Start by choosing any sequence of positive numbers  $(\sigma_\ell)_{\ell \geq 1}$  satisfying  $\lim_{\ell \rightarrow \infty} \sigma_\ell = 0$  and  $\sum_{\ell=1}^{\infty} \sigma_\ell = \infty$  (e.g.  $\sigma_\ell = 1/\ell$  will do). Then it is not hard to see that we can find positive numbers  $(\gamma_{k\ell})_{1 \leq k \leq \ell}$  such that

- (i)  $\gamma_{k\ell} \leq \sigma_\ell$  for all  $k, \ell$ .
- (ii)  $\sum_{\ell=k}^{\infty} \gamma_{k\ell} = \rho_k$  for all  $k$ .

(For instance, for each  $k$  we could choose inductively

$$\gamma_{k\ell} = \min \left[ \sigma_\ell, \frac{1}{2} \left( \rho_k - \sum_{\ell'=k}^{\ell-1} \gamma_{k\ell'} \right) \right]$$

for  $\ell \geq k$ .) Now define the matrix  $\alpha$  by

$$\alpha_{0,(1,\ell)} = \gamma_{1\ell} / \rho_0, \tag{5.23}$$

$$\alpha_{(k,\ell),(k+1,\ell)} = \gamma_{k+1,\ell} / \gamma_{k\ell} \quad \text{for } 1 \leq k \leq \ell - 1, \tag{5.24}$$

with all other coefficients of  $\alpha$  set to 0. Choose once again  $c = \rho_0 \delta_0$  and  $w = \mathbf{1}$ . Now choose the sequence  $y_1, y_2, \dots$  to be

$$0, (1, 1), (1, 2), (2, 2) \quad (1, 3), (2, 3), (3, 3), \dots$$

It is easy to see that, at any stage after the site  $(\ell, \ell)$  has been swept and before the site  $(\ell + 1, \ell + 1)$  has been swept, the total quantity of dirt  $c\beta_{y_1} \cdots \beta_{y_n} w$  does not exceed  $\sigma_{\ell+1} + \sum_{\ell'=\ell+2}^{\infty} \gamma_{1,\ell'}$ ; and this tends to zero as  $\ell \rightarrow \infty$ .

It is worth remarking that if (and only if)  $\sum_{k=0}^{\infty} \rho_k = \infty$ , then it is possible to choose first the  $(\sigma_\ell)$  and then the  $(\gamma_{k\ell})$  so that, in addition to properties (i) and (ii), we have

$$(iii) \quad \sum_{\ell=1}^{\infty} \gamma_{\ell\ell} = \infty.$$

Indeed, let  $(\sigma_k)_{k \geq 1}$  be *any* sequence of positive numbers satisfying  $\lim_{k \rightarrow \infty} \sigma_k = 0$  and  $\sum_{k=1}^{\infty} \min(\rho_k, \sigma_k) = \infty$ .<sup>7</sup> Then the preceding construction yields  $\gamma_{kk} = \min(\sigma_k, (1/2)\rho_k)$ , so that  $\sum_{k=1}^{\infty} \gamma_{kk} = \infty$ . In this way we can find a matrix  $\alpha$  that serves simultaneously for Example 5.1 and the present example: that is, there exists *both* a sequence  $x_1, x_2, \dots \in X$  in which each element of  $X$  occurs infinitely often and such that  $\lim_{n \rightarrow \infty} c\beta_{x_1} \cdots \beta_{x_n} w = \infty$ , and a sequence  $y_1, y_2, \dots \in X$  such that  $\lim_{n \rightarrow \infty} c\beta_{y_1} \cdots \beta_{y_n} w = 0$ .

## 6. Converse results

In this section we do **not** assume the *Fundamental Hypothesis*. Rather, our goal is to study what happens in case the Fundamental Hypothesis fails. As before, we are entitled to use here those results of Section 4 that refer only to clouds, but must avoid using those results that refer to operators.

### 6.1. General result

Our main result is the following:

**Theorem 6.1.** *Let  $X$  be a finite or countably infinite set, let  $\Lambda \subseteq X$ , and let  $c \geq 0$  and  $w \geq 0$  be vectors that are strictly positive on  $\Lambda$ . Consider the following conditions on a matrix  $\alpha$ :*

- (a)  $\sum_{k=0}^{\infty} c(I_\Lambda \alpha I_\Lambda)^k w < \infty$ .
- (b) For all  $h: X \rightarrow [0, 1]$  with  $\text{supp } h = \Lambda$  such that  $h \geq \varepsilon \chi_\Lambda$  for some  $\varepsilon > 0$ , we have  $\sum_{k=0}^{\infty} c(I_\Lambda \beta_h I_\Lambda)^k w < \infty$ .

---

<sup>7</sup>Given any sequence  $(\rho_k)_{k \geq 1}$  of positive numbers satisfying  $\sum_{k=1}^{\infty} \rho_k = \infty$ , we can always find a sequence  $(\sigma_k)_{k \geq 1}$  such that  $0 < \sigma_k \leq \rho_k$  for all  $k$ ,  $\lim_{k \rightarrow \infty} \sigma_k = 0$  and  $\sum_{k=1}^{\infty} \sigma_k = \infty$ . To see this, note first that  $\rho'_k := \min(\rho_k, 1)$  also satisfies  $\sum_{k=1}^{\infty} \rho'_k = \infty$ . Now let  $\sigma_k := \rho'_k / (1 + \sum_{j=1}^k \rho'_j)^{1/2}$ . Clearly  $\lim_{k \rightarrow \infty} \sigma_k = 0$ , and  $\sum_{k=1}^N \sigma_k \geq (\sum_{k=1}^N \rho'_k) / (1 + \sum_{k=1}^N \rho'_k)^{1/2} \rightarrow \infty$  as  $N \rightarrow \infty$ .



- (b') *There exists  $h: X \rightarrow [0, 1]$  with  $\text{supp } h = \Lambda$  such that  $\sum_{k=0}^{\infty} c(I_{\Lambda} \beta_h \times I_{\Lambda})^k w < \infty$ .*
- (c) *For every finite sequence  $f_1, \dots, f_m$  of functions  $X \rightarrow [0, 1]$  with  $\text{supp}(f_i) \subseteq \Lambda$  such that  $\sum_i f_i \geq \varepsilon \chi_{\Lambda}$  for some  $\varepsilon > 0$ , we have  $\sum_{k=0}^{\infty} c(I_{\Lambda} \times \beta_{f_1} \cdots \beta_{f_m} I_{\Lambda})^k w < \infty$ .*
- (c') *There exists a finite sequence  $f_1, \dots, f_m$  of functions  $X \rightarrow [0, 1]$  with  $\text{supp}(f_i) \subseteq \Lambda$  such that  $\sum_{k=0}^{\infty} c(I_{\Lambda} \beta_{f_1} \cdots \beta_{f_m} I_{\Lambda})^k w < \infty$ .*
- (d) *For every cloud  $\nu \geq 0$  that is carried by  $\Lambda$  and satisfies  $\|\nu\| \leq 1$  and  $\nu((x)) \leq 1 - \varepsilon$  for every  $x \in \Lambda$  (for some  $\varepsilon > 0$ ), we have  $\sum_{k=0}^{\infty} c(I_{\Lambda} T_{\nu} \times I_{\Lambda})^k w < \infty$ .*
- (d') *There exists a cloud  $\nu \in \mathcal{S}_1$  that is carried by  $\Lambda$  and supported on finitely many levels, such that  $\sum_{k=0}^{\infty} c(I_{\Lambda} T_{\nu} I_{\Lambda})^k w < \infty$ .*

Then (a)  $\iff$  (b)  $\iff$  (b')  $\iff$  (c)  $\iff$  (d)  $\implies$  (c')  $\implies$  (d'); and for matrices  $\alpha$  satisfying the additional hypothesis

$$\text{There exists a constant } C < \infty \text{ such that } I_{\Lambda} \alpha I_{\Lambda} w \leq C w \quad (6.1)$$

all seven conditions are equivalent.

*Proof.* (d)  $\implies$  (c)  $\implies$  (b)  $\implies$  (b')  $\implies$  (c')  $\implies$  (d') is trivial. So it suffices to prove (a)  $\implies$  (d) and (b')  $\implies$  (a), as well as to prove (d')  $\implies$  (a) under the hypothesis (6.1).

(a)  $\implies$  (d): Since  $\nu$  is carried by  $\Lambda$ , for  $k \geq 1$ ,  $(I_{\Lambda} * \nu * I_{\Lambda})^k$  charges only markers with all their entries in  $\Lambda$ . Lemma 4.13(b) then implies that

$$\sum_{k=0}^{\infty} (I_{\Lambda} * \nu * I_{\Lambda})^k \leq \frac{1}{\varepsilon} \mathbf{1}_{\Lambda} + I_{\Lambda^c} \leq \frac{1}{\varepsilon} \sum_{k=0}^{\infty} (I_{\Lambda} * \rho^1 * I_{\Lambda})^k \quad (6.2)$$

(here  $\varepsilon \leq 1$ ). Passing from clouds to operators, we have

$$\sum_{k=0}^{\infty} c(I_{\Lambda} T_{\nu} I_{\Lambda})^k w \leq \frac{1}{\varepsilon} \sum_{k=0}^{\infty} c(I_{\Lambda} \alpha I_{\Lambda})^k w < \infty. \quad (6.3)$$

(b')  $\implies$  (a) and (d')  $\implies$  (a): Let us start with hypothesis (d'), and observe that hypothesis (b') is merely the special case in which  $\nu = \beta_h$ .

Note first that we must have  $\nu((x)) < 1$  for every marker  $(x)$  of level 0 with  $x \in \Lambda$ , since otherwise  $c(I_{\Lambda} T_{\nu} I_{\Lambda})^k w$  could not tend to zero as  $k \rightarrow \infty$ . (Here we have used the strict positivity of  $c$  and  $w$  on  $\Lambda$ .)

Now define  $\mu := \mathbf{1} - \tilde{\nu}$ . By Lemma 4.14(c),  $\mu$  takes values in  $[0, 1]$  and is supported on markers having all their entries in  $\Lambda$ . By Lemma 4.14(d), if  $\nu$  is

supported on levels  $\leq K$ , then  $\mu$  is supported on levels  $\leq K - 1$ . Putting these facts together, we obtain

$$0 \leq \mu \leq I_\Lambda * \left( \sum_{k=0}^{K-1} (I_\Lambda * \rho^1 * I_\Lambda)^k \right) * I_\Lambda. \quad (6.4)$$

By Lemma 4.14(c) we then have

$$\mathbf{1}_\Lambda = \left( \sum_{n=0}^{\infty} (I_\Lambda * \nu * I_\Lambda)^n \right) * \mu \leq \left( \sum_{n=0}^{\infty} (I_\Lambda * \nu * I_\Lambda)^n \right) * \left( \sum_{k=0}^{K-1} (I_\Lambda * \rho^1 * I_\Lambda)^k \right). \quad (6.5)$$

Passing now to operators, we have

$$\sum_{k=0}^{\infty} (I_\Lambda \alpha I_\Lambda)^k \leq \left( \sum_{n=0}^{\infty} (I_\Lambda T_\nu I_\Lambda)^n \right) * \left( \sum_{k=0}^{K-1} (I_\Lambda \alpha I_\Lambda)^k \right). \quad (6.6)$$

Now sandwich this between nonnegative vectors  $c$  and  $w$ . If  $\nu = \beta_h$  [case (b')], then  $K = 1$ , so that the second large parenthesis on the right-hand side of (6.6) is the identity operator. In case (d'), we use hypothesis (6.1). Either way, we find that

$$\sum_{k=0}^{\infty} c(I_\Lambda \alpha I_\Lambda)^k w \leq C' \sum_{n=0}^{\infty} c(I_\Lambda T_\nu I_\Lambda)^n w \quad (6.7)$$

for a finite constant  $C'$ . This completes the proof.  $\square$

Please note that hypothesis (6.1) is automatic whenever  $\Lambda$  is a *finite* set. On the other hand, Examples 3.2 and 3.3 show that, when  $\Lambda$  is infinite, hypothesis (6.1) cannot be dispensed with in proving that (c')  $\implies$  (a) [or, *a fortiori*, (d')  $\implies$  (a)].

The following example shows that the hypothesis that  $\nu$  is supported on finitely many levels cannot be dispensed with in proving that (d')  $\implies$  (a), even in the presence of the Fundamental Hypothesis [which is much stronger than (6.1)]:

**Example 6.1.** Let  $\Lambda = X = \{1, 2, 3, \dots\}$ . Let  $\alpha$  be right shift (acting on dirt vectors), i.e.  $\alpha_{i,i+1} = 1$  for all  $i \geq 1$  and all other matrix elements of  $\alpha$  are 0. Set  $c = \mathbf{1}$  and let  $w_j = 1/j^2$ . Note that  $\alpha w \leq w$ . Then  $(\alpha^k w)_i = w_{i+k} = 1/(i+k)^2$ , so that

$$\sum_{k=0}^{\infty} (\alpha^k w)_i = \sum_{k=0}^{\infty} \frac{1}{(i+k)^2} \sim \frac{1}{i} \quad \text{as } i \rightarrow \infty, \quad (6.8)$$

and hence  $\sum_{k=0}^{\infty} c\alpha^k w = +\infty$ . On the other hand, let us take  $\nu$  to be the cloud

$$\nu(\eta) = \begin{cases} 1 & \text{if level}(\eta) = \text{first}(\eta), \\ 0 & \text{otherwise.} \end{cases} \quad (6.9)$$

(In other words, for dirt starting at site  $i$ ,  $\nu$  sends it upwards  $i$  levels.) For this choice of  $\alpha$  we have  $(T_\nu^k w)_i = w_{2^k i} = 4^{-k}/i^2$ , so that  $\sum_{k=0}^{\infty} cT_\nu^k w < \infty$ .

In view of Theorem 3.2, which guarantees cleaning whenever  $h$  is strictly positive on  $\Lambda$  — without any need for uniformity — one might be tempted to remove the uniformity hypothesis in condition (b), i.e. to replace (b) by

(b\*) For all  $h: X \rightarrow [0, 1]$  with  $\text{supp } h = \Lambda$ , we have  $\sum_{k=0}^{\infty} c(I_\Lambda \beta_h I_\Lambda)^k w < \infty$ .

But neither (a) nor any conceivable stronger hypothesis can possibly imply this, as Example 3.1 shows. The upshot is that, without a uniformity hypothesis on  $h$ , one can conclude in some cases that  $c(I_\Lambda \beta_h I_\Lambda)^k w$  tends to zero as  $k \rightarrow \infty$  [cf. Theorem 3.2 and Corollary 5.1], but one cannot conclude anything about the rate of convergence — in particular, one cannot conclude that this sequence is summable in  $k$ .

## 6.2. Finite $\Lambda$

If  $\Lambda$  is a *finite* subset of  $X$  (and in particular if  $X$  is finite), then all choices of strictly positive vectors  $c$  and  $w$  are equivalent, and Theorem 6.1 can be rephrased in a simpler form, in terms of the spectral radii of the various matrices. The main idea is that if the spectral radius of  $I_\Lambda \alpha I_\Lambda$  is  $\geq 1$ , then there is no way to clean the set  $\Lambda$  completely.

Let us recall that the spectral radius (= largest absolute value of an eigenvalue) of a finite matrix  $A$  satisfies  $\text{spr}(A) = \lim_{n \rightarrow \infty} \|A^n\|^{1/n}$ ; in particular, the latter limit is independent of the choice of norm. It is well known (and easy to prove using  $\|A^{m+n}\| \leq \|A^m\| \|A^n\|$ ) that  $\text{spr}(A) < 1 \iff \lim_{n \rightarrow \infty} \|A^n\| = 0 \iff \sum_{n=0}^{\infty} \|A^n\| < \infty$ . Finally, on a finite set the convergence of sequences of vectors or matrices can be understood either pointwise or in norm; the two notions are equivalent.

Since the uniformity conditions on  $h$ , etc. in Theorem 6.1 are trivially satisfied when  $\Lambda$  is finite, we can state the following immediate corollary of Theorem 6.1:

**Corollary 6.1.** *Let  $X$  be a finite or countably infinite set, and let  $\Lambda \subseteq X$  be a finite subset. Then the following conditions on a matrix  $\alpha$  are equivalent:*

- (a)  $\text{spr}(I_\Lambda \alpha I_\Lambda) < 1$ .
- (b) For all  $h: X \rightarrow [0, 1]$  with  $\text{supp } h = \Lambda$ , we have  $\text{spr}(I_\Lambda \beta_h I_\Lambda) < 1$ .

- (b') *There exists  $h: X \rightarrow [0, 1]$  with  $\text{supp } h \subseteq \Lambda$  such that  $\text{spr}(I_\Lambda \beta_h I_\Lambda) < 1$ .*
- (c) *For every finite sequence  $f_1, \dots, f_m$  of functions  $X \rightarrow [0, 1]$  with  $\text{supp}(\sum_{i=1}^m f_i) = \Lambda$ , we have  $\text{spr}(I_\Lambda \beta_{f_1} \cdots \beta_{f_m} I_\Lambda) < 1$ .*
- (c') *There exists a finite sequence  $f_1, \dots, f_m$  of functions  $X \rightarrow [0, 1]$  with  $\text{supp}(f_i) \subseteq \Lambda$  such that  $\text{spr}(I_\Lambda \beta_{f_1} \cdots \beta_{f_m} I_\Lambda) < 1$ .*
- (d) *For every cloud  $\nu \geq 0$  that is carried by  $\Lambda$  and satisfies  $\|\nu\| \leq 1$  and  $\nu((x)) < 1$  for every  $x \in \Lambda$ , we have  $\text{spr}(I_\Lambda T_\nu I_\Lambda) < 1$ .*
- (d') *There exists a cloud  $\nu \in \mathcal{S}_1$  that is carried by  $\Lambda$  and supported on finitely many levels, such that  $\text{spr}(I_\Lambda T_\nu I_\Lambda) < 1$ .*
- (e) *For every infinite sequence  $f_1, f_2, \dots$  of functions  $X \rightarrow [0, 1]$  with  $\text{supp}(f_i) \subseteq \Lambda$  such that  $\sum_i f_i = \infty$  everywhere on  $\Lambda$ , we have  $\lim_{n \rightarrow \infty} I_\Lambda \times \beta_{f_1} \cdots \beta_{f_n} I_\Lambda = 0$ .*
- (f) *For every infinite sequence  $\nu_1, \nu_2, \dots$  of nonnegative clouds carried by  $\Lambda$  and satisfying  $\|\nu_i\| \leq 1$  and  $\sum_i [1 - \nu_i((x))] = \infty$  for all  $x \in \Lambda$ , we have  $\lim_{n \rightarrow \infty} I_\Lambda T_{\nu_1} \cdots T_{\nu_n} I_\Lambda = 0$ .*

*Proof.* The equivalence of (a), (b), (b'), (c), (c'), (d) and (d') is an immediate consequence of Theorem 6.1 specialized to  $\Lambda$  finite. The implication (a)  $\implies$  (f) follows from Corollary 5.1, and (f)  $\implies$  (e) is trivial. Finally, (e)  $\implies$  (c) [or (f)  $\implies$  (d)] is easy: just consider the sequence  $f_1, \dots, f_m$  repeated infinitely many times and use Lemma 4.6.  $\square$

The following example shows that, when  $X$  is infinite, the behavior of the matrix  $\alpha$  is not completely controlled by that of the matrices  $I_\Lambda \alpha I_\Lambda$  for all finite  $\Lambda$ :

**Example 6.2.** We shall exhibit a nonnegative matrix  $\alpha$  satisfying  $\text{spr}(I_\Lambda \alpha I_\Lambda) < 1$  for all finite  $\Lambda \subset X$  — so that in particular the Fundamental Hypothesis holds for all the matrices  $I_\Lambda \alpha I_\Lambda$  — but for which the Fundamental Hypothesis does not hold, i.e. we cannot find a vector  $w > 0$  satisfying  $\alpha w \leq w$ .

Take  $X = \{1, 2, \dots\}$  and  $0 \leq \varepsilon < 1$ , and set  $\alpha_{1,1} := \varepsilon$ ,  $\alpha_{1,j} := 1$  for all  $j \geq 2$ ,  $\alpha_{i+1,i} := 2$  for all  $i \geq 2$ , and  $\alpha_{i,j} := 0$  elsewhere:

$$\alpha = \begin{pmatrix} \varepsilon & 1 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 2 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \quad (6.10)$$

If we could find a  $w > 0$  such that  $\alpha w \leq w$ , we would have  $w_{i+1} \geq 2w_i$  for all  $i \geq 2$ . Since  $w_2 > 0$ , this would imply  $w_i \rightarrow +\infty$  as  $i \rightarrow \infty$ . But  $\alpha w \leq w$  implies also that  $w_1 \geq \sum_{i=2}^{\infty} w_i$ , which is impossible.

Now let us show why  $\text{spr}(I_{\Lambda} \alpha I_{\Lambda}) < 1$  for all finite  $\Lambda \subset X$ . Set  $m := \max \Lambda$ . Then, for any vector  $v$  and any  $j \geq m - 1$ , we have  $(I_{\Lambda} \alpha I_{\Lambda})^j v \in \text{span}(e_1)$  where  $e_1 := (1, 0, 0, \dots)$ . Since  $\alpha e_1 = \varepsilon e_1$ , it follows that

$$\text{spr}(I_{\Lambda} \alpha I_{\Lambda}) = \begin{cases} \varepsilon & \text{if } 1 \in \Lambda, \\ 0 & \text{if } 1 \notin \Lambda. \end{cases} \quad (6.11)$$

In particular, by taking  $\varepsilon = 0$  we can even arrange to have  $\text{spr}(I_{\Lambda} \alpha I_{\Lambda}) = 0$  [i.e.,  $I_{\Lambda} \alpha I_{\Lambda}$  is nilpotent] for every finite  $\Lambda$ .

### Acknowledgments

We wish to thank Claude Dellacherie for valuable conversations on probabilistic potential theory; Uriel Rothblum for a valuable correspondence concerning Perron–Frobenius theorems; and Victor Blondel, Kousha Etessami and Alex Scott for valuable correspondence and/or conversations concerning computational complexity and decidability. Last but not least, we wish to thank Marina Papa Sokal for emphasizing to us the importance of the *experimental* study of the floor-cleaning problem.

This research was supported in part by U.S. National Science Foundation grants PHY-0099393 and PHY-0424082. One of us (A.D.S.) also wishes to thank the Université de Rouen for kind hospitality and financial support during several visits.

This paper is dedicated to our daughters Serena Xu Xiaoling Papa-Sokal and Fanny Danayit de la Rue; it could not have been written with them.

### References

- [1] M. AIZENMAN AND E.H. LIEB (1984) Dobrushin revisited. Unpublished manuscript (October 1984).
- [2] A. BERMAN AND R.J. PLEMMONS (1979) *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York.
- [3] V.D. BLONDEL AND V. CANTERINI (2003) Undecidable problems for probabilistic automata of fixed dimension. *Theory Comput. Systems* **36**, 231–245.
- [4] V.D. BLONDEL AND J.N. TSITSIKLIS (1997) When is a pair of matrices mortal? *Inform. Process. Lett.* **63**, 283–286.
- [5] V.D. BLONDEL AND J.N. TSITSIKLIS (2000) A survey of computational complexity results in systems and control. *Automatica* **36**, 1249–1274.

- [6] V.D. BLONDEL AND J.N. TSITSIKLIS (2005) Stable matrix in a semigroup generated by nonnegative matrices is undecidable. Unpublished manuscript (March 2005).
- [7] R.L. DOBRUSHIN (1968) The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory Probab. Appl.* **13**, 197–224.
- [8] R.L. DOBRUSHIN (1970) Prescribing a system of random variables by conditional distributions. *Theory Probab. Appl.* **15**, 458–486.
- [9] R.L. DOBRUSHIN AND S.B. SHLOSMAN (1985) Constructive criterion for the uniqueness of Gibbs field. In: *Statistical Physics and Dynamical Systems* (Köszeg, 1984), J. Fritz, A. Jaffe and D. Szász (eds.), Birkhäuser, Boston MA, 347–370.
- [10] H. FÖLLMER (1982) A covariance estimate for Gibbs measures. *J. Func. Anal.* **46**, 387–395.
- [11] H.-O. GEORGI (1988) *Gibbs Measures and Phase Transitions*, Section 8.1. De Gruyter, Berlin–New York.
- [12] J.G. KEMENY, J.L. SNELL AND A.W. KNAPP (1966) *Denumerable Markov Chains*, especially Chapters 7–9.. Van Nostrand, Princeton.
- [13] O.E. LANFORD III (1973) Entropy and equilibrium states in classical statistical mechanics. In: *Statistical Mechanics and Mathematical Problems*, A. Lenard (ed.), Lect. Notes Phys. **20**, Springer-Verlag, Berlin, 1–113.
- [14] E. NUMMELIN (1984) *General Irreducible Markov Chains and Non-Negative Operators*, especially Chapter 3. Cambridge University Press, Cambridge.
- [15] D. REVUZ (1975) *Markov Chains*, especially Chapters 1 and 2. North-Holland, Amsterdam.
- [16] U. ROTHBLUM (2004) Private communication (April 2004).
- [17] E. SENETA (1981) *Non-Negative Matrices and Markov Chains*. Springer-Verlag, New York–Heidelberg–Berlin.
- [18] B. SIMON (1983) *The Statistical Mechanics of Lattice Gases*, vol. I, Sections V.1 and V.3.. Princeton University Press, Princeton.
- [19] F.L. SPITZER (1976) *Principles of Random Walk*. Springer-Verlag, New York.
- [20] R.P. STANLEY (1999) *Enumerative Combinatorics*, vols. 1 and 2. Cambridge University Press, Cambridge–New York.
- [21] L.N. VASERSHTEIN (1969) Markov processes over denumerable products of spaces describing large system of automata. *Problems Inform. Transmission* **5** (3), 47–52.
- [22] D. WEITZ (2005) Combinatorial criteria for uniqueness of Gibbs measures. *Random Struct. and Algorithms* **27**, 445–475.