The Annals of Probability
2015, Vol. 43, No. 4, 1892–1943
DOI: 10.1214/14-AOP923

BRANCHING RANDOM TESSELLATIONS WITH INTERACTION: A THERMODYNAMIC VIEW

BY HANS-OTTO GEORGII, TOMASZ SCHREIBER[1] AND CHRISTOPH THÄLE

Ludwig-Maximilians University Munich, Nicolaus Copernicus University Toruń and Ruhr University Bochum

A branching random tessellation (BRT) is a stochastic process that transforms a coarse initial tessellation of $\mathbb{R}^d$ into a finer tessellation by means of random cell divisions in continuous time. This concept generalises the so-called STIT tessellations, for which all cells split up independently of each other. Here, we allow the cells to interact, in that the division rule for each cell may depend on the structure of the surrounding tessellation. Moreover, we consider coloured tessellations, for which each cell is marked with an internal property, called its colour. Under a suitable condition, the cell interaction of a BRT can be specified by a measure kernel, the so-called division kernel, that determines the division rules of all cells and gives rise to a Gibbsian characterisation of BRTs. For translation invariant BRTs, we introduce an "inner" entropy density relative to a STIT tessellation. Together with an inner energy density for a given "moderate" division kernel, this leads to a variational principle for BRTs with this prescribed kernel, and further to an existence result for such BRTs.

1. Introduction. A central object of stochastic geometry and spatial stochastics are tessellations of $\mathbb{R}^d$ (with $d \geq 1$), that is, locally finite families of d-dimensional convex polytopes that cover $\mathbb{R}^d$ and have pairwise disjoint interiors. They are used in many practical applications. For example, random tessellations serve as models for cellular or polycrystalline materials, plant cells or influence zones, for instance, in the modelling of telecommunication networks or animal territories; see [19, 29] for an overview.

The standard random tessellations usually considered in the literature are the Poisson hyperplane tessellations, the Poisson–Voronoi and the Poisson–Delaunay tessellations; cf. [22] for definitions. These have the property of being facet-to-facet (or side-to-side in the planar case), which is to say that the intersection of any

Received April 2013; revised November 2013.

[1]This work has been initiated by the second author, Tomasz Schreiber. Shortly before his untimely death of cancer at the age of 35, he entrusted to the first author a draft manuscript that introduced the general setup of this paper and proposed a study of relative entropy densities for BRTs. To keep his memory, we have tried to realise his intentions with suitable modifications of concepts and results. HOG, CT.

MSC2010 subject classifications. Primary 60D05, 60K35; secondary 28D20, 60G55, 82B21.

Key words and phrases. Branching tessellation, coloured tessellation, free energy, Gibbs measure, relative entropy, STIT tessellation, stochastic geometry, variational principle.

two of its cells is either empty or a common face of both cells. However, there are numerous applications for which models of this kind are inappropriate, for example, network models for telecommunication systems or models for crack structures in geology. Hence, there is a growing demand for mathematically tractable models of nonfacet-to-facet tessellations, which may serve as idealised reference models. Only some years ago, the class of iteration-stable random tessellations (called STIT tessellations for short) was introduced by Nagel and Weiß in [18]. These tessellations are constructed by means of a temporal random process of cell division, and thus live in space–time. They have attracted considerable interest because of its analytical tractability; see, for example, [20, 23–28] or [30].

Our objects of study here generalise the STIT models in two respects. On the one hand, we consider coloured tessellations, for which each cell is equipped with an individual colour. For example, the colour of a cell could represent its nutrient content, its genotype, age, or whatever else might be relevant to describe the state of a cell. (In a different context, coloured tessellations have been studied by Arak and Surgailis [1, 2], e.g.) On the other hand, and more importantly, we allow for an *interaction of cells* during their division process. That is, our objects of interest can be viewed in two ways that are equivalent but deal differently with space–time: either

– as Gibbsian spatial systems of interacting branching processes of coloured cells, or
– as temporal processes of tessellations in space.

The latter viewpoint can informally be described as follows. At time zero, one starts with an initial random tessellation of $\mathbb{R}^d$ into coloured cells. Each cell lives for a random time, which is determined by an interactive competition of cells. Namely, the survival rate of a cell c at any time $s > 0$ may not only depend on the cell's geometry and colour, but in fact on the whole tessellation including its past evolution. When the lifetime has run out, a hyperplane with coloured half-spaces is chosen randomly according to some rule that may again depend on the cell's geometry, colour and the past evolution of the surrounding tessellation, and is used to cut c into two polyhedral sub-cells c^+ and c^-, which inherit their colours from the respective half-spaces of the cutting hyperplane. The daughter cells c^+ and c^- then replace c in the collective division game, which is continued until time 1, say. The resulting tessellation of $\mathbb{R}^d$ at a deterministic time $s \in [0, 1]$ is denoted by T_s, and the tessellation-valued stochastic process $(T_s)_{s\in[0,1]}$ is what we call a *branching random tessellation* or BRT for short. The rule determining the splitting of cells is given by a measure kernel, which will be called the associated *division kernel*.

In the special case when (i) the distribution of lifetimes is exponential with parameter proportional to the mean width of the cells, and (ii) the bi-coloured hyperplanes are chosen at random according to the motion-invariant hyperplane measure and some reference measure on the colour space, $(T_s)_{s\in[0,1]}$ is a coloured STIT tessellation of $\mathbb{R}^d$ and its distribution is invariant under rigid motions whenever so is the initial random tessellation. The coloured STIT tessellations play an important role in the background of our theory, in a way which is conceptually similar to that of the Poisson point processes in the theory of Gibbsian point processes.

Let us note that the Gibbsian viewpoint, for which the BRTs are considered as interacting branching processes of coloured cells, parallels the Gibbsian treatment of interacting particle systems and interacting diffusions developed in [5–7, 10], for example. Let us also mention that different tessellation models with cell interaction, namely Delaunay or Voronoi tessellations of Gibbsian type (which undergo no time evolution), are studied in [3, 8, 9].

The main results of this paper are the following.

– To begin, we discuss how the intuitive concept of "cell interaction" that governs a BRT $\mathbf{P}$ can be specified by a so-called *division kernel* Φ. We show that such a Φ can equivalently be used in two different ways: either as the collection of instantaneous splitting rates of all cells during their joint time evolution, or in the Gibbsian way, as a means to determine the conditional distribution of the behaviour of all cells within any bounded window when that of all other cells is given. A third equivalent use of Φ involves a Campbell-like formula for the jump intensity measure of $\mathbf{P}$. We show further that a measure kernel Φ as above exists as soon as $\mathbf{P}$ satisfies a condition of local absolute continuity (LAC) relative to a STIT model.
– We then turn to a kind of thermodynamic formalism for BRTs $\mathbf{P}$ that are invariant under spatial translations. The basic quantity is an inner entropy density $h^{\mathrm{in}}(\mathbf{P})$, which is defined as the limit of a conditional entropy per unit volume of $\mathbf{P}$ relative to a reference STIT model. The adjective "inner" refers to the fact that only the cells completely inside the respective window are taken into account, rather than all cells that hit the window. The functional h^{in} will be shown to share some familiar properties of the entropy densities for the standard models of statistical mechanics, at least with some natural adaptations.
– Finally, we consider an arbitrary division kernel Ψ that satisfies some mild regularity conditions, which roughly require that Ψ is not too far from a STIT kernel; such a Ψ will be called moderate. We introduce an associated inner energy density $u^{\mathrm{in}}(\mathbf{P};\Psi)$ as well as some sort of pressure $v^{\mathrm{in}}(\mathbf{P};\Psi)$. The resulting inner excess free energy density $h^{\mathrm{in}}(\mathbf{P};\Psi)$ gives rise to a variational principle, which states that the minimisers of $h^{\mathrm{in}}(\cdot;\Psi)$ are precisely the translation invariant BRTs that admit Ψ as their division kernel. It is further shown that such minimisers do exist, for any prescribed distribution P of the time-zero tessellation. This proves the existence of a BRT $\mathbf{P}$ for any given initial distribution P and any moderate division kernel Ψ. For general Ψ, such a $\mathbf{P}$ is not necessarily unique.

The paper is organised as follows: Section 2 introduces the setup and recalls some necessary facts. Besides tessellations and BRTs, the main concepts are division kernels and local conditional BRTs of Gibbsian type. This section also includes some examples of division kernels to which our theory applies. The main results together with their framework are stated in Section 3. These are Theorems 3.1 and 3.3 on the significance and existence of global division kernels, Theorems 3.5 and 3.6 on the existence of the inner entropy density and its properties, and Theorems 3.9 and 3.10 on the variational characterisation and the existence of invariant BRTs with given moderate division kernels. All proofs are collected in the final Section 4.

2. Preliminaries.

2.1. *Tessellations*.

2.1.1. *Polytopes and tessellations*. Consider the Euclidean space $\mathbb{R}^d$ of arbitrary dimension $d \geq 1$. We shall deal with certain random processes of coloured tessellations of $\mathbb{R}^d$ into (coloured) convex polytopes. Let us specify these terms. First, a *polytope* p in $\mathbb{R}^d$ is the closed convex hull of a finite set of points and is always assumed to have nonempty interior; the set of all such polytopes is denoted by $\mathbb{P}$. Each polytope $p \in \mathbb{P}$ is equipped with a translation covariant selector $m(p)$, called its "*centre*" or "*midpoint*", for example, its barycentre, its Steiner point or its circumcentre. We write $r(p) = \max_{x\in p} |x - m(p)|$ for its *radius* and ∂p and $\mathrm{int}(p)$ for its topological *boundary*, respectively, *interior*.

More generally, we will assume that each polytope is marked with some internal property, called its *colour*. So, we fix an arbitrary Polish space Σ, which we call the *colour space*. A coloured polytope, called *cell* in the sequel, is a pair $c = (p, \sigma)$ with $p \in \mathbb{P}$ and $\sigma \in \Sigma$. Let us denote by $\mathrm{sp}(c) := p$ and $\mathrm{col}(c) := \sigma$, respectively, the spatial part and the colour of c. The space of cells is thus $\mathbb{C} := \mathbb{P} \times \Sigma$. To simplify notation, *we adopt the general convention that spatial operations on cells* (*and also on coloured tessellations defined below*), *such as intersections with subsets of* $\mathbb{R}^d$ *and translations*, *solely refer to the spatial part and do not affect their colours*. For example, $m(c) := m(\mathrm{sp}(c))$, $r(c) := r(\mathrm{sp}(c))$, $\mathrm{int}(c) := \mathrm{int}(\mathrm{sp}(c))$, $c \cap W := (\mathrm{sp}(c) \cap W, \mathrm{col}(c))$ for $W \subset \mathbb{R}^d$, and $c - x := (\mathrm{sp}(c) - x, \mathrm{col}(c))$ when $x \in \mathbb{R}^d$. Finally, $\mathrm{vol}(c) := \mathrm{vol}_d(\mathrm{sp}(c))$ is the (d-dimensional) volume of the spatial part of c. Let us also define the space $\mathbb{C}_0 = \{c \in \mathbb{C} : m(c) = 0\}$ of cells having their midpoint at the origin.

The cells are the constituents of the coloured tessellations which we introduce now; for brevity we will omit the adjective "coloured" in the following. (Note that letting Σ be a singleton one recovers the uncoloured case usually considered in the literature; cf. [22, 29].)

DEFINITION 2.1. A (coloured) *tessellation* T of $\mathbb{R}^d$ is a countable subset of $\mathbb{C}$ such that:

- T is locally finite, in that any bounded subset of $\mathbb{R}^d$ only hits a finite number of cells from T,
- two distinct cells of T have disjoint interiors, that is, $\mathrm{int}(c) \cap \mathrm{int}(c') = \varnothing$ for all $c, c' \in T$ with $c \neq c'$,
- the cells cover the whole space, which is to say that $\bigcup_{c\in T} c = \mathbb{R}^d$.

The space of all tessellations of $\mathbb{R}^d$ will henceforth be denoted by $\mathbb{T}$.

Besides tessellations of $\mathbb{R}^d$, we will also consider tessellations in local windows $W \subset \mathbb{R}^d$, which will generally be chosen to be polytopes, or sometimes also finite unions of polytopes. So, we write $\mathbb{P}_\cup$ for the set of all finite, not necessarily connected unions of polytopes, and for $W \in \mathbb{P}_\cup$ we let $\mathbb{C}_W$ be the set of cells that are contained in W. We finally write $\mathbb{T}_W$ for the set of all tessellations of W, that is, of all finite collections $\{c_1, \dots, c_n\}$ of cells with pairwise disjoint interiors and such that $c_1 \cup \dots \cup c_n = W$.

2.1.2. *Measurability.* We need measurable structures on all spaces introduced above. We start with the space $\mathbb{P}$ of polytopes. As the sets in $\mathbb{P}$ are compact and nonempty, the natural metric on $\mathbb{P}$ is the usual Hausdorff distance d_H; cf. [22], Chapter 12.3. Hence, the space $\mathbb{P}$ can be equipped with the Borel σ-field $\mathcal{B}(\mathbb{P})$ induced by d_H. In fact, $\mathcal{B}(\mathbb{P})$ is generated by the sets $\{p \in \mathbb{P} : p \cap B \neq \varnothing\}$ with $B \in \mathcal{B}(\mathbb{R}^d)$, the Borel σ-field on $\mathbb{R}^d$; see [22], Chapters 12.2–12.3. The coloured counterpart $\mathbb{C}$ is endowed with the product σ-field $\mathcal{B}(\mathbb{C}) = \mathcal{B}(\mathbb{P}) \otimes \mathcal{B}(\Sigma)$, where $\mathcal{B}(\Sigma)$ is the Borel σ-field on Σ. The space $\mathbb{C}_0$ of centred cells receives the trace σ-field.

We next need to introduce a suitable σ-field on $\mathbb{T}$. As is usual in point process theory, we let $\mathcal{B}(\mathbb{T})$ be the σ-field generated by the counting variables

$$N_A : \mathbb{T} \to \mathbb{N} \cup \{+\infty\}, \qquad T \mapsto |T \cap A|, \qquad A \in \mathcal{B}(\mathbb{C}), \tag{2.1}$$

where $|\cdot|$ stands for the cardinality of the argument set, that is, N_A counts how many cells of T belong to A. In view of the structure of $\mathcal{B}(\mathbb{C})$, $\mathcal{B}(\mathbb{T})$ is also generated by the random variables

$$N_{B,S} : \mathbb{T} \ni T \mapsto \bigl|\{c \in T : c \cap B \neq \varnothing, \mathrm{col}(c) \in S\}\bigr|,$$

with B a bounded Borel set in $\mathbb{R}^d$ and $S \in \mathcal{B}(\Sigma)$. Moreover, $\mathcal{B}(\mathbb{T})$ is the Borel σ-field for the vague topology on $\mathbb{T}$, which is generated by the functions

$$e_g : \mathbb{T} \to [0, \infty), \qquad T \mapsto \sum_{c\in T} g(c),$$

where $g \geq 0$ is a continuous function on $\mathbb{C}$ with a bounded support in the spatial coordinate; see [15], Appendix 15.7, or [16], Theorem A2.3.

To deal with local properties of tessellations, we will often restrict a tessellation to a local window $W \in \mathbb{P}$. We thus define the *projection* to such a W by

$$\pi_W : \mathbb{T} \to \mathbb{T}_W, \qquad T \mapsto T_W := \bigl\{c \cap W : c \in T, \mathrm{int}(c \cap W) \neq \varnothing\bigr\}. \tag{2.2}$$

In the same manner as above, we may introduce a σ-field $\mathcal{B}(\mathbb{T}_W)$ on $\mathbb{T}_W$. One can then easily check that the mapping π_W is measurable.

The culminating concept of this subsection is the following.

DEFINITION 2.2. A probability measure P on $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ satisfying the first-moment condition $\int P(\mathrm{d}T)|T_W| < \infty$ for all windows $W \in \mathbb{P}$ is called a *random tessellation*. The set of all such P is denoted by $\mathscr{P}(\mathbb{T})$.

2.2. *Branching tessellations.*

2.2.1. *Cutting cells by hyperplanes.* We now turn to the main objects of our investigation: tessellations which arise from a given initial tessellation by a successive splitting of cells into two pieces by means of suitable hyperplanes. Recall that a hyperplane η with unit normal $u \in \mathbb{S}_+^{d-1}$ (upper unit half-sphere) and signed distance $r \in \mathbb{R}$ to the origin can be written in the form $\eta = \{x \in \mathbb{R}^d : \langle x, u\rangle = r\}$, where $\langle \cdot, \cdot \rangle$ stands for the usual scalar product. So, the space of hyperplanes can be identified with $\mathbb{S}_+^{d-1} \times \mathbb{R}$. For η as above, we write $\eta^+ = \{x \in \mathbb{R}^d : \langle x, u\rangle \geq r\}$ and $\eta^- = \{x \in \mathbb{R}^d : \langle x, u\rangle \leq r\}$ for the associated half-spaces. More generally, we consider *bi-coloured hyperplanes* $H = (\eta, \sigma^+, \sigma^-) \in \mathbb{H} := \mathbb{S}_+^{d-1} \times \mathbb{R} \times \Sigma^2$, for which each of the half-spaces $\eta^\pm$ is equipped with a colour $\sigma^\pm$. We write $\mathrm{sp}(H) := \eta$ and $\mathrm{col}^\pm(H) := \sigma^\pm$, respectively, for the spatial part and the colours of H and again *adopt the convention that spatial operations with bi-coloured hyperplanes only refer to the spatial part*, for example, $c \cap H := \mathrm{sp}(c) \cap \mathrm{sp}(H)$ or $c \cap H^\pm := (\mathrm{sp}(c) \cap \mathrm{sp}(H)^\pm, \mathrm{col}^\pm(H))$ for any $c \in \mathbb{C}$. Moreover, for such a cell c, we let

$$\langle c\rangle = \{H \in \mathbb{H} : H \cap \mathrm{int}(c) \neq \varnothing\} \tag{2.3}$$

be the set of all bi-coloured hyperplanes which hit the interior of (the spatial part of) c. Each bi-coloured hyperplane H defines a cell division operation $\oslash$ on tessellations. Namely, let $T \in \mathbb{T}$, $c \in T$ and $H \in \langle c\rangle$. Then $\oslash$ is defined by

$$\oslash_{c,H}(T) := (T \setminus \{c\}) \cup \{c \cap H^+, c \cap H^-\} \tag{2.4}$$

with $c \cap H^\pm$ as above. Branching tessellations are now defined as follows. For simplicity, the time interval will mostly be the unit interval [0, 1].

DEFINITION 2.3. (a) Let $W \in \mathbb{P}_\cup$ be a finite union of polytopes. A *branching tessellation in the window* W with bounded time interval $I = [a, b)$ or $[a, b]$ is a family $\mathbf{T} = (T_s)_{s \in I}$ of tessellations in W such that:

- the function $s \mapsto T_s$ from I to $\mathbb{T}_W$ is piecewise constant, right-continuous and has only a finite number of jumps,

- at each point s of discontinuity (so that $T_s \neq T_{s-} := \lim_{r\uparrow s} T_r$), there exists a unique cell $c \in T_{s-}$ and a bi-coloured hyperplane $H \in \langle c\rangle$ such that

$$T_s = \oslash_{c,H}(T_{s-}).$$

Further, T_a is called the initial tessellation. We write $\mathbb{BT}_W$ for the set of all such branching tessellations in W.

(b) A family $\mathbf{T} = (T_s)_{0\le s\le 1}$ is called a *branching tessellation in* $\mathbb{R}^d$ if for each window $W \in \mathbb{P}$ the restricted process $\mathbf{T}_W = \boldsymbol{\pi}_W(\mathbf{T}) := (\pi_W(T_s))_{0\le s\le 1}$ is a branching tessellation in W. Again, T_0 is then called the initial tessellation of $\mathbf{T}$. The set of all branching tessellations in $\mathbb{R}^d$ is denoted by $\mathbb{BT}$.

The following remark provides a further way of describing the time evolution of a branching tessellation.

REMARK 2.4. (a) Let $\mathbf{T}$ be a branching tessellation in a window $W \in \mathbb{P}_\cup$ with time interval $I = [0, 1]$. (The case of other time intervals is similar.) Keeping record of all jump times of $\mathbf{T}$ together with the associated cells that are divided and the respective cutting hyperplanes, one arrives at the set

$$\begin{aligned}\mathscr{D}(\mathbf{T}) = \{&(s, c, H) \in (0, 1] \times \mathbb{C} \times \mathbb{H} : T_{s-} \neq T_s,\\ &c \in T_{s-}, H \in \langle c\rangle, T_s = \oslash_{c,H}(T_{s-})\}\end{aligned} \tag{2.5}$$

of all "division events". There is a one-to-one correspondence between $\mathbf{T}$ and the pair $(T_0, \mathscr{D}(\mathbf{T}))$, in that $\mathbf{T}$ can be recovered from the initial tessellation T_0 and the set $\mathscr{D}(\mathbf{T})$ of division events. Indeed, labelling the elements of $\mathscr{D}(\mathbf{T})$ with the indices $1, \dots, n := |\mathscr{D}(\mathbf{T})|$ according to the order of their time coordinates so that $0 =: s_0 < s_1 < \cdots < s_n \le s_{n+1} := 1$, one has the recursion $T_s = T_0$ for $s \in [0, s_1)$ and

$$T_s = \oslash_{c_i,H_i}(T_{s_{i-1}}) \qquad \text{for } s \in [s_i, s_{i+1}),\ i = 1, \dots, n.$$

Finally, $T_1 = T_{s_n}$.

This description also gives rise to a convenient way of visualising $\mathbf{T}$ as a graph in $[0, 1] \times \mathbb{C}_W$; see Figure 1. The set of vertices is

$$\mathscr{V}(\mathbf{T}) = \{(0, c) : c \in T_0\} \cup \{(s, c \cap H^\pm) : (s, c, H) \in \mathscr{D}(\mathbf{T})\}.$$

Moreover, each $(s, c) \in \mathscr{V}(\mathbf{T})$ is equipped with a "lifeline" $[s, s^*) \times \{c\}$, where $s^* = s'$ if $(s', c, H) \in \mathscr{D}(\mathbf{T})$ for some $s' > s$ and $H \in \langle c\rangle$, and $s^* = 1$ otherwise. If $s^* < 1$, this lifeline is augmented by the lines from (s^*, c) to the two children $(s^*, c \cap H^\pm)$ of (s, c). If $s^* = 1$, the half-open line $[s, 1) \times \{c\}$ is replaced by the closed line $[s, 1] \times \{c\}$. In this way, one obtains a finite forest of binary "family" trees in $\mathbb{C}_W$ that evolve from the cells of T_0. So, these cells are the roots, or ancestors, and the $|T_0| + |\mathscr{D}(\mathbf{T})|$ leaves form the tessellation T_1. This branching mechanism is strongly reminiscent of the fragmentation processes considered in [4].

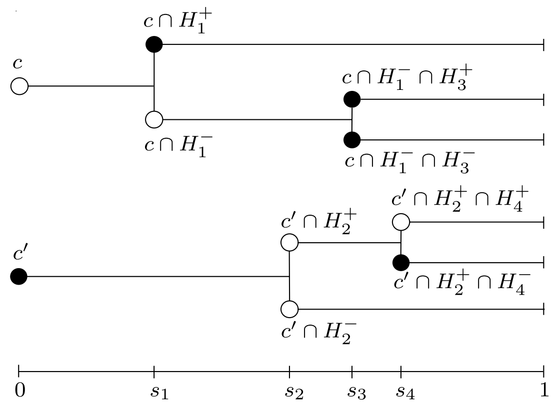

FIG. 1. *Representation of a two-coloured branching tessellation in a finite window with initial tessellation $T_0 = \{c, c'\}$. The cells living at a time s constitute a tessellation T_s. At each time s_i, a cell that lives up to this moment is selected and cut in two by a bi-coloured hyperplane H_i, which impresses its colours onto the cell's pieces.*

(b) Branching tessellations in the whole space $\mathbb{R}^d$ admit a similar description in terms of division events. For each $\mathbf{T} \in \mathbb{BT}$, we can then define

$$\mathscr{D}(\mathbf{T}) = \bigcup_{V \in \mathbb{P}} \bigcap_{W \in \mathbb{P}:\, W \supset V} \mathscr{D}(\mathbf{T}_W). \tag{2.6}$$

Conversely, for each $W \in \mathbb{P}$ one can recover the division events in W from $\mathscr{D}(\mathbf{T})$ via

$$\mathscr{D}(\mathbf{T}_W) = \{(s, c \cap W, H) : (s, c, H) \in \mathscr{D}(\mathbf{T}),\, H \in \langle c \cap W \rangle\}.$$

It follows that $\mathbf{T}$ is uniquely determined by T_0 and $\mathscr{D}(\mathbf{T})$, and $\mathbf{T}$ can be regarded as a forest of infinitely many finite binary family trees of coloured cells, the roots of which correspond to the cells of the initial tessellation T_0 of $\mathbb{R}^d$.

Later on, it will be essential for us to keep track of the past of a branching tessellation. So, instead of considering the evolution $\mathbf{T} = (T_s)_{0 \le s \le 1}$ in $\mathbb{T}$, we will consider the process $(\mathbf{T}_s)_{0 \le s \le 1}$ in $\mathbb{BT}$, which is given by $\mathbf{T}_s = (T_u)_{0 \le u \le s}$. Equivalently, $\mathbf{T}_s$ can be thought of as being obtained from $\mathbf{T}$ by removing from $\mathscr{D}(\mathbf{T})$ all elements with time-coordinate larger than s. In this way, each $\mathbf{T}_s$ can be considered to be an element of $\mathbb{BT}$, which is frozen at time s (and thus remains constant thereafter). The set of all such branching tessellations is denoted by $\mathbb{BT}_s$. In particular, $\mathbb{BT}_1 = \mathbb{BT}$, and $\mathbb{BT}_u \subset \mathbb{BT}_s$ when $u < s$. We write

$$\boldsymbol{\pi}_s : \mathbb{BT} \to \mathbb{BT}_s, \qquad \mathbf{T} \mapsto \mathbf{T}_s, \tag{2.7}$$

for the natural projection that removes the division events after time s. As before, the nonbold T_s stands for the tessellation at time s, whereas a bold $\mathbf{T}_s$ stands for an element of $\mathbb{BT}_s$.

Besides this projection concerning time, we have also the projection to a spatial window $W \in \mathbb{P}$, which is given by

$$\boldsymbol{\pi}_W : \mathbb{BT} \to \mathbb{BT}_W, \qquad \mathbf{T} \mapsto \mathbf{T}_W = (T_{W,s})_{0 \le s \le 1} \quad \text{with } T_{W,s} = \pi_W(T_s), \tag{2.8}$$

where $\mathbb{BT}_W = \boldsymbol{\pi}_W(\mathbb{BT})$ and π_W is as in (2.2). We also write $\boldsymbol{\pi}_{W,s} = \boldsymbol{\pi}_W \circ \boldsymbol{\pi}_s$, $\mathbf{T}_{W,s} = \boldsymbol{\pi}_{W,s}(\mathbf{T})$ and $\mathbb{BT}_{W,s} = \boldsymbol{\pi}_{W,s}(\mathbb{BT})$. So, to obtain $\mathbf{T}_{W,s}$ from $\mathbf{T}$ one has to remove from $\mathscr{D}(\mathbf{T})$ all division events with a time coordinate exceeding s or a hyperplane not hitting the cell's intersection with W.

2.2.2. *Branching random tessellations*. Our main objects of interest are probability measures on $\mathbb{BT}$. So, we need to equip $\mathbb{BT}$ with a σ-field. We know from Remark 2.4 that each $\mathbf{T} \in \mathbb{BT}$ is uniquely determined by its initial tessellation T_0 together with the set $\mathscr{D}(\mathbf{T})$ of division events as given by (2.5) and (2.6). Since $\mathscr{D}(\mathbf{T})$ is a locally finite subset of $(0, 1] \times \mathbb{C} \times \mathbb{H}$, one can proceed as usually in point process theory by defining $\mathcal{B} = \mathcal{B}(\mathbb{BT})$ as the smallest σ-field for which the counting variables

$$\mathbf{N}_{A,B} : \mathbf{T} \mapsto |T_0 \cap A| + |\mathscr{D}(\mathbf{T}) \cap B| \tag{2.9}$$

with $A \in \mathcal{B}(\mathbb{C})$ and $B \in \mathcal{B}((0, 1]) \otimes \mathcal{B}(\mathbb{C}) \otimes \mathcal{B}(\mathbb{H})$ are measurable; here $\mathcal{B}((0, 1])$ denotes the Borel σ-field on $(0, 1]$. By standard theory, $(\mathbb{BT}, \mathcal{B})$ is a Borel space. For any window $W \in \mathbb{P}$, we define a σ-field $\mathcal{B}_W = \mathcal{B}(\mathbb{BT}_W)$ on $\mathbb{BT}_W$ in the same way. To simplify notation, we will not distinguish between the σ-field $\mathcal{B}_W$ on $\mathbb{BT}_W$ and its pre-image $\boldsymbol{\pi}_W^{-1}\mathcal{B}_W$ on $\mathbb{BT}$, which will be denoted by the same symbol. Anyway, with these definitions it is clear that both the projection $\boldsymbol{\pi}_W$ in (2.8) and the time restriction map $\boldsymbol{\pi}_\bullet : (s, \mathbf{T}) \mapsto \mathbf{T}_s$ of (2.7) are measurable.

DEFINITION 2.5. A *branching random tessellation* (*BRT*) of $\mathbb{R}^d$ is a probability measure $\mathbf{P}$ on $(\mathbb{BT}, \mathcal{B})$ satisfying the first-moment condition

$$\int \mathbf{P}(\mathrm{d}\mathbf{T}) |T_{W,1}| < \infty \qquad \text{for all windows } W \in \mathbb{P}. \tag{2.10}$$

The set of all such BRTs of $\mathbb{R}^d$ is denoted by $\mathscr{P} = \mathscr{P}(\mathbb{BT})$. BRTs within a window $W \in \mathbb{P}_\cup$ are defined analogously.

For every $\mathbf{P} \in \mathscr{P}$ and any of the projections $\boldsymbol{\pi}_*$ in (2.7) and (2.8), we write $\mathbf{P}_* = \mathbf{P} \circ \boldsymbol{\pi}_*^{-1}$ for the image of $\mathbf{P}$ under $\boldsymbol{\pi}_*$. In particular, each $\mathbf{P}_s$ is a BRT. In fact, one can achieve that $\mathbf{P}_s$ depends measurably on s, in that the mapping $[0, 1] \times \mathcal{B} \ni (s, A) \mapsto \mathbf{P}_s(A)$ is a probability kernel, as will be assumed throughout the following. This can be seen by disintegrating the measure

$$\overline{\mathbf{P}} := \int_0^1 \mathrm{d}s \int \mathbf{P}(\mathrm{d}\mathbf{T}) \delta_{(s,\mathbf{T}_s)} \tag{2.11}$$

on $\overline{\mathbb{BT}} := \{(s, \mathbf{T}_s) : s \in [0, 1], \mathbf{T}_s \in \mathbb{BT}_s\}$; cf. [15], Appendix 15.3. Later on, we will also consider the projections $\bar{\boldsymbol{\pi}}_W = \mathrm{id} \otimes \boldsymbol{\pi}_W$ that act on the second coordinate of $\overline{\mathbb{BT}}$ as in (2.8) and leave the first coordinate untouched, and the projection images $\overline{\mathbf{P}}_W = \overline{\mathbf{P}} \circ \bar{\boldsymbol{\pi}}_W^{-1}$, where $W \in \mathbb{P}$. We also introduce the notation $\overline{\mathbb{BT}}_W := \bar{\boldsymbol{\pi}}_W(\overline{\mathbb{BT}})$.

2.3. *Division kernels.* Consider a random element $\mathbf{T}$ of $\mathbb{BT}_W$ for a window $W \in \mathbb{P}_\cup$. The process $(\mathbf{T}_s)_{0\le s\le 1}$ is then automatically Markovian because its "past" is part of the "present". In this paper, we will focus on the "nice" case in which the evolution of this Markov process is described by a rate kernel that specifies the jump times and transitions of $(\mathbf{T}_s)_{0\le s\le 1}$. Since the only transitions are single-cell divisions by bi-coloured hyperplanes, this means that the rate kernels take the following form.

DEFINITION 2.6. A *division kernel* is a measure kernel Φ from the set

$$\{(s, \mathbf{T}_s, c) \in \overline{\mathbb{BT}} \times \mathbb{C} : c \in T_s\}$$

to $\mathbb{H}$ such that each $\Phi(s, \mathbf{T}_s, c, \cdot)$ is a finite measure supported on $\langle c\rangle \subset \mathbb{H}$. If Φ is only defined for arguments in $\overline{\mathbb{BT}}_W \times \mathbb{C}_W$, Φ is called a *division kernel for the window* $W \in \mathbb{P}_\cup$.

In the following, it will be convenient to work also with the *cumulative division kernel*

$$\widehat{\Phi}(s, \mathbf{T}_s, \cdot) = \sum_{c\in T_s} \delta_c \otimes \Phi(s, \mathbf{T}_s, c, \cdot) \tag{2.12}$$

from $\overline{\mathbb{BT}}$ to $\mathbb{C} \times \mathbb{H}$. Note that, conversely, $\Phi(s, \mathbf{T}_s, c, \cdot) = \widehat{\Phi}(s, \mathbf{T}_s, \{c\} \times \cdot)$.

The next remark describes how a division kernel determines the evolution of a BRT within a bounded window.

REMARK 2.7 (Local BRTs with prescribed division kernels). Let $W \in \mathbb{P}_\cup$ be a fixed window, Φ_W be a division kernel for W, and

$$\hat{\phi}_W(s, \mathbf{T}_s) := \widehat{\Phi}_W\bigl(s, \mathbf{T}_s, T_s \times \langle W\rangle\bigr)$$

the finite total mass of the cumulative kernel $\widehat{\Phi}_W(s, \mathbf{T}_s, \cdot)$. We construct a random element $\mathbf{T}$ of $\mathbb{BT}_W$ as follows:

(I) Pick an initial tessellation $T_0 \in \mathbb{T}_W$ according to some probability law P_W on $\mathbb{T}_W$, and let $s_0 = 0$ and $\mathbf{T}_0 := T_0$. Also, let $i = 1$ and proceed with the following random recursion over the number i.

(R) Suppose that $i \ge 1$ and both a random time $s_{i-1} \in [0, 1]$ and a BRT $\mathbf{T}_{s_{i-1}} \in \mathbb{BT}_{s_{i-1}}$ are already realised. Then take a random time $s_i \in (s_{i-1}, \infty]$ with "survival" probability

$$\mathrm{Prob}(s_i > s) = \exp\biggl[-\int_{s_{i-1}}^{s} \hat{\phi}_W(u \wedge 1, \mathbf{T}_{s_{i-1}})\, \mathrm{d}u\biggr] \tag{2.13}$$

for $s > s_{i-1}$. If $s_i \leq 1$, proceed to define an extension $\mathbf{T}_{s_i} \in \mathbb{BT}_{s_i}$ of $\mathbf{T}_{s_{i-1}}$ as follows: pick a random cell $c_i \in T_{s_{i-1}}$ and a bi-coloured hyperplane H_i according to the law

$$\widehat{\Phi}_W(s_i, \mathbf{T}_{s_{i-1}}, \cdot)/\hat{\phi}_W(s_i, \mathbf{T}_{s_{i-1}}).$$

(Note that the denominator does not vanish for each possible choice of s_i.) Then let $T_s = T_{s_{i-1}}$ for $s \in (s_{i-1}, s_i)$ and $T_{s_i} = \oslash_{c_i, H_i}(T_{s_{i-1}})$, that is,

$$\mathscr{D}(\mathbf{T}_{s_i}) = \mathscr{D}(\mathbf{T}_{s_{i-1}}) \cup \{(s_i, c_i, H_i)\}.$$

Next, let $i := i + 1$ and go to (R). In the case $s_i > 1$, let $T_s = T_{s_{i-1}}$ for $s \in (s_{i-1}, 1]$, set $n = i - 1$, and stop.

One needs to ensure that this algorithm terminates after finitely many steps. It is not difficult to show that this is the case if

$$\sup_{s, \mathbf{T}_s, c} \Phi_W(s, \mathbf{T}_s, c, \langle c \rangle) =: \phi < \infty; \tag{2.14}$$

see the proof of Lemma 4.3 below. This lemma shows further that the process $(\mathbf{T}_s)_{0 \leq s \leq 1}$ can be characterised as the unique, in general time-inhomogeneous pure jump (i.e., piecewise constant) Markov process in $\mathbb{BT}_W$ with initial distribution P_W and generator

$$\mathbb{L}_{W,s}^{\Phi_W} g(\mathbf{T}_s) = \int_{T_s \times \langle W \rangle} \widehat{\Phi}_W(s, \mathbf{T}_s, \mathrm{d}(c, H)) [g(\oslash_{s,c,H}(\mathbf{T}_s)) - g(\mathbf{T}_s)] \tag{2.15}$$

at time $s \in [0, 1]$. Here, $\oslash_{s,c,H}(\mathbf{T}_s) \in \mathbb{BT}_s$ is the branching tessellation that coincides with $\mathbf{T}_s$ for times less than s and equals $\oslash_{c,H}(T_s)$ at time s, and g is any bounded measurable function on $\mathbb{BT}_W$. The distribution of $\mathbf{T}$ is a BRT $\mathbf{P}_W$ in $W \in \mathbb{P}$, and this $\mathbf{P}_W$ is called the *BRT in W with division kernel Φ_W and initial distribution P_W*.

The main objects of this paper are BRTs on the full space $\mathbb{R}^d$ that can be characterised in a similar way as the local BRTs in the remark above. Namely, for any division kernel Φ and $0 \leq s \leq 1$ we define an operator $\mathbb{L}_s^\Phi$ by

$$\mathbb{L}_s^\Phi g(\mathbf{T}_s) = \int \widehat{\Phi}(s, \mathbf{T}_s, \mathrm{d}(c, H)) [g(\oslash_{s,c,H}(\mathbf{T}_s)) - g(\mathbf{T}_s)]. \tag{2.16}$$

Here, $\oslash_{s,c,H}$ is as in the preceding remark, and g is any bounded local function on $\mathbb{BT}$, where *local* means that g is $\mathcal{B}_W$-measurable for some $W \in \mathbb{P}$.

DEFINITION 2.8. For a given division kernel Φ, we will say that a BRT $\mathbf{P} \in \mathscr{P}$ *evolves according to* Φ if the Markov process $\mathbf{T} = (\mathbf{T}_s)_{0 \leq s \leq 1}$ in $\mathbb{BT}$ with distribution $\mathbf{P}$ satisfies the forward equation with generators $\mathbb{L}_s^\Phi$, in that

$$\int_0^t \mathrm{d}s \int \mathrm{d}\mathbf{P}_s \mathbb{L}_s^\Phi g = \int g \, \mathrm{d}\mathbf{P}_t - \int g \, \mathrm{d}\mathbf{P}_0 \tag{2.17}$$

for all $t \in [0, 1]$ and all bounded local functions g on $\mathbb{BT}$.

Obviously, this definition refers to a BRT $\mathbf{P}$ as a process evolving in time, by saying that the Markov process with distribution $\mathbf{P}$ evolves just as the local processes in Remark 2.7, in that a cell c in environment $\mathbf{T}_s$ at time s is split by a bi-coloured hyperplane H with instantaneous intensity $\widehat{\Phi}(s, \mathbf{T}_s, c, \mathrm{d}H)\,\mathrm{d}s$. Later we will study the spatial aspects of $\mathbf{P}$.

2.4. *Examples of division kernels*. This section contains a few examples of division kernels; two simulation pictures are shown in Figure 2. The first is (by now) classical and will be used as a reference model throughout the following.

EXAMPLE 2.9 (STIT tessellations). Let Λ be a locally finite measure on the set $\mathbb{H}$ of all bi-coloured hyperplanes, which is invariant under all translations. That is, under the identification of $H \in \mathbb{H}$ with $(u, r, \sigma^+, \sigma^-) \in \mathbb{S}_+^{d-1} \times \mathbb{R} \times \Sigma^2$, Λ can be written in the form

$$\Lambda(\mathrm{d}H) = \lambda(\mathrm{d}u)\,\mathrm{d}r\,\mu\bigl(u, \mathrm{d}\sigma^+, \mathrm{d}\sigma^-\bigr). \tag{2.18}$$

Here, λ is a measure on $\mathbb{S}_+^{d-1}$, and μ is a probability kernel from $\mathbb{S}_+^{d-1}$ to Σ^2. (The translation invariance is expressed by the fact that the r-marginal is Lebesgue measure and μ does not depend on r.) A natural choice is the motion-invariant measure Λ_{iso} for which λ is the normalised surface measure λ_{iso} on $\mathbb{S}_+^{d-1}$ and $\mu(u, \cdot) = \nu \otimes \nu$ for a reference probability measure ν on Σ. Then a STIT tessellation with driving measure Λ is a BRT for the division kernel

$$\Lambda^*(s, \mathbf{T}_s, c, \cdot) := \Lambda\bigl(\cdot \cap \langle c\rangle\bigr). \tag{2.19}$$

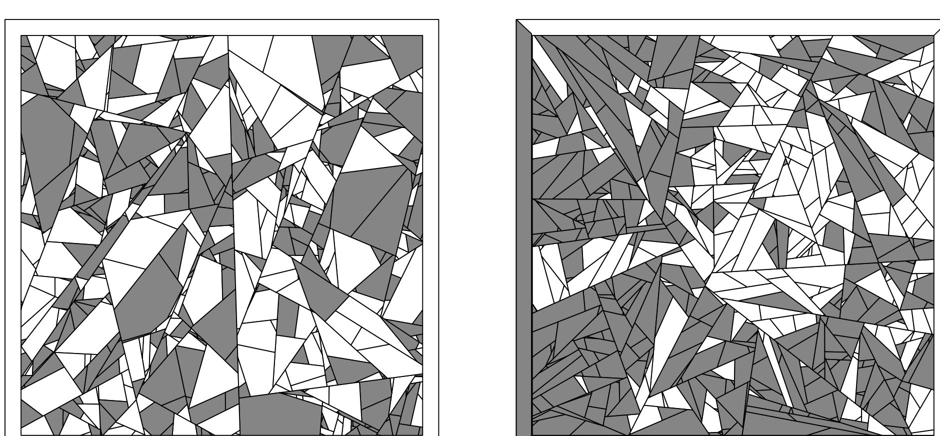

FIG. 2. *Simulations of two BRTs with isotropic selection of lines*, *two colours and the full window as single initial cell*. Left: *A STIT tessellation*; *colours are chosen at random*. Right: *Colour mutation and size balancing as in Example* 2.10, *but without aging*. *Here*, $\varepsilon = 0.025$, $\beta(\mathsf{s}) = (1 + \mathsf{s})/2$ *and a mixed boundary condition as indicated*.

In the uncoloured case, this model has been introduced by Mecke, Nagel and Weiß [17, 18]. Since Λ^* does not depend on the time s, the random holding times $s_i - s_{i-1}$ in Remark 2.7 above are exponentially distributed and can be understood as minima over $c \in T_{s_{i-1}}$ of independent exponential times with parameter $\Lambda(\langle c\rangle)$, which are associated to the presently existing cells. [In the isotropic case $\Lambda = \Lambda_{\mathrm{iso}}$, the parameter $\Lambda(\langle c\rangle)$ is precisely the mean width of c.] In other words, the tessellations evolve according to a continuous-time branching process on $\mathbb{C}_W$, $W \in \mathbb{P}$, in which all cells c behave independently of each other, live for an exponential time with parameter $\Lambda(\langle c\rangle)$ and then split into two parts according to the conditional distribution $\Lambda(\cdot|\langle c\rangle)$. In particular, this implies that smaller cells live stochastically longer.

In view of this independence of the evolution in different cells, it is clear that for each $T_0 \in \mathbb{T}$ there exists a unique whole-space BRT $\mathbf{\Pi}^\Lambda(T_0, \cdot)$, called STIT tessellation of $\mathbb{R}^d$, with initial tessellation T_0 and driving measure Λ. In fact, if the support of λ contains a linear basis of $\mathbb{R}^d$, one can also construct a unique BRT $\mathbf{\Pi}^{\Lambda,\infty} = \mathbf{\Pi}^\Lambda(T_0^\infty, \cdot)$ for the degenerate initial tessellations T_0^∞ that consist of the single "cell" $\mathbb{R}^d$ with any colour σ; see [17], Theorem 1, and [18], Theorem 1. [By (2.19), $\mathbf{\Pi}^{\Lambda,\infty}$ does not depend on σ.]

Formally, $\mathbf{\Pi}^\Lambda$ is a probability kernel from $\mathbb{T}$ to $\mathbb{BT}$. So, for each $P \in \mathscr{P}(\mathbb{T})$, $P\mathbf{\Pi}^\Lambda = \int P(\mathrm{d}T_0)\mathbf{\Pi}^\Lambda(T_0, \cdot)$ is the unique BRT for Λ with initial distribution P. Its projections to arbitrary windows $W \in \mathbb{P}$ are given by

$$(P\mathbf{\Pi}^\Lambda) \circ \boldsymbol{\pi}_W^{-1} = P_W \mathbf{\Pi}_W^\Lambda \tag{2.20}$$

for the restricted STIT kernel $\mathbf{\Pi}_W^\Lambda(T_{W,0}, \cdot)$ from $\mathbb{T}_W$ to $\mathbb{BT}_W$ with the restricted driving measure $\Lambda(\cdot \cap \langle W\rangle)$. The abbreviation STIT stands for stability under the operation of iteration of tessellations. An explanation and further remarkable properties can be found in [17, 18, 20, 23–28] and [30].

A generalisation of the STIT models, which still keeps the independence of the division process for distinct cells, are the *cell-driven BRTs*, which have division kernels of the form

$$\Phi(s, \mathbf{T}_s, c, \mathrm{d}H) = \varphi(c, H)\Lambda(\mathrm{d}H) \tag{2.21}$$

with a density function $\varphi(c, H)$ on $\mathbb{C} \times \mathbb{H}$ which vanishes except when $H \in \langle c\rangle$. A special case are the *shape-driven BRTs* investigated in [26]; see also the examples therein.

The next example demonstrates the flexibility of modelling in the present setting: it combines an interaction between the colours of the cells with a geometric homogenisation mechanism and an aging effect. The last feature takes advantage of the fact that division kernels may also depend on the past.

EXAMPLE 2.10 (Contact-induced mutations with size balancing and aging). Let the colour space be $\Sigma = \{-1, 1\}$ and consider a division kernel of the form

$$\Phi(s, \mathbf{T}_s, c, \mathrm{d}H) = \varphi(c, \eta)\lambda_{\mathrm{iso}}(\mathrm{d}u)\,\mathrm{d}r\,\mu\big(s, \mathbf{T}_s, c, \mathrm{d}\sigma^+\big)\mu\big(s, \mathbf{T}_s, c, \mathrm{d}\sigma^-\big),$$

where $H = (\eta, \sigma^+, \sigma^-)$ with spatial part $\eta = (u, r) \in \mathbb{S}_+^{d-1} \times \mathbb{R}$ and colours $\sigma^\pm \in \Sigma$. A special choice of the geometric pre-factor is

$$\varphi(c, \eta) = \varepsilon \mathbb{1}_{\langle c\rangle}(\eta) + \varepsilon^{-1}\mathbb{1}_{\langle \varepsilon\star c\rangle}(\eta),$$

for some small $\varepsilon > 0$; here, $\varepsilon \star c = m(c) + \varepsilon(c - m(c))$ is the ε-retraction of c. This choice has the effect that the cutting hyperplane will typically pass close to the midpoint $m(c)$ of c, so that its two daughter cells have comparable size. One can further choose the colouring rule

$$\mu(s, \mathbf{T}_s, c, \cdot) = \delta_{\mathrm{col}(c)} + \beta(\mathsf{a}_{s,c,\mathbf{T}_s}, \mathsf{s}_{c,T_s})\delta_{-\mathrm{col}(c)},$$

where $\mathsf{a}_{s,c,\mathbf{T}_s} = s - \min\{u \in [0, s] : c \in T_u\}$ is the age of c at time s,

$$\mathsf{s}_{c,T_s} = \sum_{c'\in T_s :\, \mathrm{col}(c')=-\mathrm{col}(c)} \mathrm{vol}_{d-1}\big(c \cap c'\big)/\,\mathrm{vol}_{d-1}(\partial c)$$

is the opposite-type surface fraction (measured by the Hausdorff measure of dimension $d - 1$), and $\beta : [0, 1]^2 \to (0, \infty)$ is a suitable positive function. For instance, β can be taken to be decreasing in a so that increasing age reduces the willingness of splitting and mutating. One can further let β be increasing in s. Then the larger a cell's surface fraction is in contact with cells of opposite type, the more the cell gets "nervous" and hurries to divide, and the more likely it is that its daughter cells mutate to adapt their type to that of the neighbours.

Our third example may seem somewhat exotic. It will be used in Remark 3.12 to demonstrate that a BRT on the full space $\mathbb{R}^d$ is not necessarily uniquely determined by its initial distribution and its division kernel.

EXAMPLE 2.11 (Directional infinite-range interaction). This is an uncoloured model, for which Σ is a singleton. We further confine ourselves to the planar case $d = 2$. Let $\Lambda_{\mathrm{hor}}(\mathrm{d}H) = \delta_{(0,1)}(\mathrm{d}u)\,\mathrm{d}r$ and $\Lambda_{\mathrm{vert}}(\mathrm{d}H) = \delta_{(1,0)}(\mathrm{d}u)\,\mathrm{d}r$ be the measures on $\mathbb{H} = \mathbb{S}_+^1 \times \mathbb{R}$ for which all lines are horizontal, respectively, vertical. For any cell $c \in \mathbb{C}$ let $\mathrm{diam}_{\mathrm{hor}}(c) = \max_{x,y\in c} |x_1 - y_1|$ and $\mathrm{diam}_{\mathrm{vert}}(c) = \max_{x,y\in c} |x_2 - y_2|$ be the horizontal and vertical diameters of c, where x_i stands for the ith coordinate of x. Also, let

$$\mathbb{C}_{\mathrm{hor}} = \big\{c \in \mathbb{C} : \mathrm{diam}_{\mathrm{hor}}(c) > \mathrm{diam}_{\mathrm{vert}}(c)\big\}$$

be the set of all "horizontal" cells. Finally, writing $[n]$ for the centred square of area n^2, let

$$\rho_{\mathrm{hor}}(T) = \limsup_{n\to\infty} n^{-2}\big|\big\{c \in T \cap \mathbb{C}_{\mathrm{hor}} : m(c) \in [n]\big\}\big|$$

be the upper density of horizontal cells for a tessellation $T \in \mathbb{T}$, and define $\rho_{\mathrm{vert}}(T)$ analogously. Then let

$$\mathbb{T}_{\mathrm{hor}} = \{T \in \mathbb{T} : \rho_{\mathrm{hor}}(T) > \rho_{\mathrm{vert}}(T)\}$$

be the set of tessellations with a dominating fraction of horizontal cells, and $\mathbb{T}_{\mathrm{vert}} = \mathbb{T} \setminus \mathbb{T}_{\mathrm{hor}}$. Consider the division kernel

$$\Phi(s, \mathbf{T}_s, c, \cdot) = \mathbb{1}_{\mathbb{T}_{\mathrm{hor}}}(T_s)\Lambda_{\mathrm{hor}}(\cdot \cap \langle c\rangle) + \mathbb{1}_{\mathbb{T}_{\mathrm{vert}}}(T_s)\Lambda_{\mathrm{vert}}(\cdot \cap \langle c\rangle). \tag{2.22}$$

Since $\mathbb{T}_{\mathrm{hor}}$ is invariant under translations and tail-measurable, this Φ looks at the actual tessellation "at infinity" in order to decide whether the cutting line should be horizontal or vertical.

2.5. *Gibbsian BRTs.* In this section, we introduce a Gibbsian perspective on BRTs. As is standard in the theory of Gibbs measures, one aims at describing a macroscopic system by means of its local conditional distributions that describe the behaviour inside a bounded region when the remaining system is fixed. We first define such conditional distributions in the context of BRTs. This will allow us then to introduce Gibbsian BRTs. Let $W \in \mathbb{P}$ be a fixed window.

2.5.1. *Inner and outer projections.* Recall from (2.2) and (2.8) that the projections π_W and $\boldsymbol{\pi}_W$ are defined by intersecting the cells with W, and thus wipes off much information on the cell geometry (such as, e.g., the location of midpoints). To avoid this, we introduce the "*inner*" *projection*

$$\pi_W^{\mathrm{in}} : \mathbb{T} \to \mathbb{T}_W^{\mathrm{in}}, \qquad T \mapsto T_W^{\mathrm{in}} := \{c \in T : c \subset \mathrm{int}(W)\}, \tag{2.23}$$

which removes all cells which are not completely contained in the interior of W. It takes values in the set $\mathbb{T}_W^{\mathrm{in}}$ of all possibly empty, not necessarily connected collections of cells inside W with pairwise disjoint interiors. The counting variables N_A in (2.1) are even defined on $\mathbb{T}_W^{\mathrm{in}}$ and generate a σ-field $\mathcal{B}(\mathbb{T}_W^{\mathrm{in}})$, for which π_W^{in} is measurable. As the cells of T_W^{in} are even required to be contained in the interior of W, T_W^{in} is a measurable function of T_W.

In the same way, we define the *inner projection*

$$\boldsymbol{\pi}_W^{\mathrm{in}} : \mathbb{BT} \ni \mathbf{T} \mapsto \mathbf{T}_W^{\mathrm{in}} = (T_{W,s}^{\mathrm{in}})_{0 \le s \le 1} \tag{2.24}$$

on $\mathbb{BT}$, where $T_{W,s}^{\mathrm{in}} = \pi_W^{\mathrm{in}}(T_s)$. Arguing as in Remark 2.4, one finds that $\mathbf{T}_W^{\mathrm{in}}$ is uniquely determined by $T_{W,0}^{\mathrm{in}}$, $\mathscr{D}(\mathbf{T}_W^{\mathrm{in}})$, and the finite set of all "immigration events" (s, c) with $T_{W,s}^{\mathrm{in}} = T_{W,s-}^{\mathrm{in}} \cup \{c\}$. Consequently, one can generate a σ-field on the range $\mathbb{BT}_W^{\mathrm{in}}$ of $\boldsymbol{\pi}_W^{\mathrm{in}}$ by means of counting variables similar to those in (2.9), so that $\boldsymbol{\pi}_W^{\mathrm{in}}$ becomes measurable. Note also that $\boldsymbol{\pi}_W^{\mathrm{in}} = \boldsymbol{\pi}_W^{\mathrm{in}} \circ \boldsymbol{\pi}_W$.

Complementary to the above, we also introduce an "*outer*" *projection for W* by

$$\pi_W^{\mathrm{out}} : \mathbb{T} \ni T \mapsto T_W^{\mathrm{out}} := T \setminus T_W^{\mathrm{in}} = \{c \in T : c \setminus \mathrm{int}(W) \neq \varnothing\}, \tag{2.25}$$

and a "*boundary*" *projection*

$$\pi_W^{\partial} : \mathbb{T} \ni T \mapsto \pi_W^{\partial}(T_W) = \{c \cap W : c \in T, c \cap \partial W \neq \varnothing\}. \tag{2.26}$$

Likewise, on the level of branching tessellations, we define

$$\boldsymbol{\pi}_W^{\mathrm{out}} : \mathbb{BT} \ni \mathbf{T} \mapsto \mathbf{T}_W^{\mathrm{out}} = \big(\pi_W^{\mathrm{out}}(T_s)\big)_{0 \le s \le 1}, \tag{2.27}$$

$$\boldsymbol{\pi}_W^{\partial} : \mathbb{BT} \ni \mathbf{T} \mapsto \mathbf{T}_W^{\partial} = \big(\pi_W^{\partial}(T_s)\big)_{0 \le s \le 1} = \boldsymbol{\pi}_W^{\mathrm{out}}(\mathbf{T}_W). \tag{2.28}$$

In the forest picture of Figure 1, each $\mathbf{T}_W^{\mathrm{out}}$ in the range $\mathbb{BT}_W^{\mathrm{out}}$ of $\boldsymbol{\pi}_W^{\mathrm{out}}$ corresponds to a forest of binary trees from which all cells within W are erased. So, one can use the counting variables in (2.9) to generate a σ-field on $\mathbb{BT}_W^{\mathrm{out}}$, and $\boldsymbol{\pi}_W^{\mathrm{out}}$ is then evidently measurable. The same applies to $\boldsymbol{\pi}_W^{\partial}$. Furthermore, to keep the full information on the initial tessellation in $\mathbb{R}^d$, respectively, in W, it will also be convenient to introduce the mappings

$$\boldsymbol{\pi}_W^{0,\mathrm{out}} : \mathbf{T} \mapsto \mathbf{T}_W^{0,\mathrm{out}} := \big(T_{W,0}^{\mathrm{in}}, \mathbf{T}_W^{\mathrm{out}}\big), \tag{2.29}$$

$$\boldsymbol{\pi}_W^{0,\partial} : \mathbf{T} \mapsto \mathbf{T}_W^{0,\partial} := \big(T_{W,0}^{\mathrm{in}}, \mathbf{T}_W^{\partial}\big). \tag{2.30}$$

For each of the projections $\boldsymbol{\pi}_W^{*}$ in (2.24), (2.27), (2.28), (2.29) and (2.30), we write $\mathcal{B}_W^{*} = \sigma(\boldsymbol{\pi}_W^{*})$ for the σ-field on $\mathbb{BT}$ that is generated by this projection. By abuse of notation, we will use the same symbol $\mathcal{B}_W^{*}$ for the σ-field on the range of $\boldsymbol{\pi}_W^{*}$.

2.5.2. *Conditional BRTs*. Let $\mathbf{T} \in \mathbb{BT}$ any branching tessellation. Consider the time-dependent "inner" window

$$\mathrm{in}_W\big(s, \mathbf{T}_W^{\partial}\big) := W \setminus \mathrm{int}\big(\cup\{c : c \in T_{W,s}^{\partial}\}\big) = \cup\{c : c \in T_{W,s}^{\mathrm{in}}\}, \tag{2.31}$$

which is possibly empty and not necessarily connected. It is measurable jointly in both arguments, piecewise constant and right-continuous as a function of s. Let

$$0 < t_1 = t_1\big(\mathbf{T}_W^{\partial}\big) < \cdots < t_n = t_{n(\mathbf{T}_W^{\partial})}\big(\mathbf{T}_W^{\partial}\big) < 1 \tag{2.32}$$

be the jump times of the path $s \mapsto \mathrm{in}_W(s, \mathbf{T}_W^{\partial})$, which depend measurably on $\mathbf{T}_W^{\partial}$. [Note that possibly $n(\mathbf{T}_W^{\partial}) = 0$. For the sake of convenience, we also exclude the case that there is a jump at time 1, which occurs with probability zero.] At each t_i, $\mathbf{T}_W^{\partial}$ creates a new cell c_i inside W, namely

$$c_i = c_i\big(\mathbf{T}_W^{\partial}\big) := \mathrm{cl}\big(\mathrm{in}_W\big(t_i, \mathbf{T}_W^{\partial}\big) \setminus \mathrm{in}_W\big(t_{i-1}, \mathbf{T}_W^{\partial}\big)\big),$$

where $t_0 = 0$. In other words, $\mathbf{T}_W^{\partial}$ induces a process of immigration of cells into W.

DEFINITION 2.12. Let Φ be a division kernel and suppose that the following random process $\mathbf{S} = (S_s)_{0 \le s \le 1}$ with $\mathbf{S} \cup \mathbf{T}_W^{\mathrm{out}} := (S_s \cup T_{W,s}^{\mathrm{out}})_{0 \le s \le 1} \in \mathbb{BT}$ is well defined:

- Let $\mathbf{S}_{[0,t_1)} = (S_s)_{0\le s<t_1}$ be the BRT in the window $\mathrm{in}_W(0, \mathbf{T}_W^{\partial})$ with time interval $[0, t_1)$, initial tessellation $S_0 = T_{W,0}^{\mathrm{in}}$ and division kernel
$$\Phi_W^{\mathrm{in}}(s, \mathbf{S}_s, c, \cdot|\mathbf{T}_W^{\mathrm{out}}) := \Phi(s, \mathbf{S}_s \cup \mathbf{T}_{W,s}^{\mathrm{out}}, c, \cdot)$$
for $c \in S_s$, $s \in [0, t_1)$.
- For $i = 1, \dots, n$ and conditional on $\mathbf{S}_{t_i-}$ let $\mathbf{S}_{[t_i,t_{i+1})} = (S_s)_{t_i\le s<t_{i+1}}$ be the BRT in the window $\mathrm{in}_W(t_i, \mathbf{T}_W^{\partial}) = c_i \cup \mathrm{in}_W(t_{i-1}, \mathbf{T}_W^{\partial})$ with time interval $[t_i, t_{i+1})$, initial tessellation $S_{t_i} = S_{t_i-} \cup \{c_i\}$ and division kernel
$$\Phi_W^{\mathrm{in}}(s, \mathbf{S}_s, c, \cdot|\mathbf{T}_W^{\mathrm{out}}) := \Phi(s, \mathbf{S}_s \cup \mathbf{T}_{W,s}^{\mathrm{out}}, c, \cdot)$$
for $c \in S_s$, $s \in [t_i, t_{i+1})$. Here $t_{n+1} = 1$, and we finally set $S_1 := S_{1-}$.

The distribution of $\mathbf{S}$ on $\mathbb{BT}_W^{\mathrm{in}}$ will be denoted by $\mathbf{G}_W^{\Phi}(\cdot|\mathbf{T}_W^{0,\mathrm{out}})$ and is called the *conditional BRT for* Φ *in* W *with initial tessellation* $T_{W,0}^{\mathrm{in}}$ *and boundary condition* $\mathbf{T}_W^{\mathrm{out}}$.

By construction, $\mathbf{G}_W^{\Phi}$ is a probability kernel from $(\mathbb{BT}_W^{0,\mathrm{out}}, \mathcal{B}_W^{0,\mathrm{out}})$ to $(\mathbb{BT}_W^{\mathrm{in}}, \mathcal{B}_W^{\mathrm{in}})$.

EXAMPLE 2.13 (Conditional STIT tessellations). As in Example 2.9, let Λ be a locally finite measure on $\mathbb{H}$ and Λ^* be the associated division kernel; cf. (2.19). Then $\mathbf{G}_W^{\Lambda}(\cdot|\mathbf{T}_W^{0,\mathrm{out}}) := \mathbf{G}_W^{\Lambda^*}(\cdot|\mathbf{T}_W^{0,\mathrm{out}})$ is simply the distribution of
$$\bigcup_{c\in T_{W,0}^{\mathrm{in}}} \mathbf{S}^{(c)} \cup \bigcup_{i=1}^{n} \mathbf{S}^{(i)}$$
for independent random STIT tessellations $\mathbf{S}^{(c)}$ and $\mathbf{S}^{(i)}$ for Λ. Here, $\mathbf{S}^{(c)}$ evolves in time $[0, 1]$ from the single-cell tessellation $S_0^{(c)} = \{c\}$ of the initial polytope $\mathrm{sp}(c)$, whereas $\mathbf{S}^{(i)}$ evolves in time $[t_i, 1]$ from the single-cell initial tessellation $S_{t_i}^{(i)} = \{c_i\}$ of the "immigrated" polytope $\mathrm{sp}(c_i)$ and is extended to the full interval $[0, 1]$ by setting $S_s^{(i)} = \varnothing$ for $s \in [0, t_i)$. Since Λ^* does not depend on the surrounding tessellation, it follows that the measure $\mathbf{G}_W^{\Lambda}(\cdot|\mathbf{T}_W^{0,\mathrm{out}})$ depends only on $\mathbf{T}_W^{0,\partial}$.

Here is the natural counterpart of the concept of (macroscopic) Gibbs measures in our setup of branching random tessellations.

DEFINITION 2.14. Let Φ be any division kernel. A BRT $\mathbf{P} \in \mathscr{P}$ is called a *Gibbsian BRT for* Φ if, for all $W \in \mathbb{P}$, $\mathbf{G}_W^{\Phi}$ is a regular version of its conditional probability given $\mathcal{B}_W^{0,\mathrm{out}}$. More explicitly, this means that
$$\int f\,\mathrm{d}\mathbf{P} = \int \mathbf{P}(\mathrm{d}\mathbf{T}) \int \mathbf{G}_W^{\Phi}(\mathrm{d}\mathbf{S}|\mathbf{T}_W^{0,\mathrm{out}}) f(\mathbf{S} \cup \mathbf{T}_W^{\mathrm{out}})$$
for all bounded measurable functions f on $\mathbb{BT}$ and all $W \in \mathbb{P}$.

In contrast to Definition 2.8 in which a BRT is considered as a process in time, the preceding definition emphasises the spatial aspects of a BRT, by saying that Φ describes the cell splitting mechanism within an arbitrary local window when the evolution of all other cells is given.

2.6. *Translation invariance.* A main focus of this paper is on BRTs that are invariant under spatial translations. For each $x \in \mathbb{R}^d$, the translation ϑ_x by the vextor $-x$ acts:

– on cells $c \in \mathbb{C}$ via $\vartheta_x : c \mapsto c - x := (\mathrm{sp}(c) - x, \mathrm{col}(c))$,
– on bi-coloured hyperplanes $H \in \mathbb{H}$ via

$$\vartheta_x : H \mapsto H - x := \big(\mathrm{sp}(H) - x, \mathrm{col}^+(H), \mathrm{col}^-(H)\big),$$

– on tessellations $T \in \mathbb{T}$ via

$$\vartheta_x : T \mapsto T - x := \{c - x : c \in T\},$$

– on branching tessellations $\mathbf{T} = (T_s)_{0 \le s \le 1} \in \mathbb{BT}$ via

$$\vartheta_x : \mathbf{T} \mapsto \mathbf{T} - x := (T_s - x)_{0 \le s \le 1}.$$

That is, only the spatial coordinates are shifted, but the colours remain unchanged. Moreover, by abuse of notation we use the same symbol ϑ_x for the translation on each level, and we will also use it for the simultaneous translation of pairs of objects as above.

DEFINITION 2.15. A BRT $\mathbf{P} \in \mathscr{P}$ is called *translation invariant* if it is invariant under the action of the translation group $\Theta = (\vartheta_x)_{x \in \mathbb{R}^d}$ on $\mathbb{BT}$, in that $\mathbf{P} \circ \vartheta_x^{-1} = \mathbf{P}$ for all $x \in \mathbb{R}^d$. We write $\mathscr{P}_\Theta = \mathscr{P}_\Theta(\mathbb{BT})$ for the set of all translation invariant BRTs that satisfy the first-moment condition (2.10), which by translation invariance is equivalent to the requirement that the "hitting intensity"

$$i_1(\mathbf{P}) := \int \mathbf{P}(\mathrm{d}\mathbf{T}) |T_{[1],1}| \tag{2.33}$$

is finite. Here, $[1] := [-1/2, 1/2]^d$ stands for the centred unit cube.

Translation invariance allows to investigate the behaviour of a random tessellation "around a typical cell", which for convenience is located "around the origin". This is formalised by means of Palm calculus as presented in [15], Chapter 12, and [22], Theorem 4.1.1. Let $\mathbf{P} \in \mathscr{P}_\Theta$ be given. Then the *Campbell measure* of $\mathbf{P}$ on $\mathbb{BT} \times \mathbb{C}$ is defined by

$$\mathbf{C}^{\mathbf{P}} = \int \mathbf{P}(\mathrm{d}\mathbf{T}) \sum_{c \in T_1} \delta_{(\mathbf{T}, c)}. \tag{2.34}$$

It captures the joint distribution of the (terminal) cells and the complete history of their surrounding tessellation. The Palm calculus now states that there exists a finite measure $\mathbf{P}^0$ on $\mathbb{BT} \times \mathbb{C}_0$, the so-called *Palm measure* of $\mathbf{P}$, such that the *Palm formula*

$$\int \mathrm{d}\mathbf{C}^{\mathbf{P}}(\mathbf{T}, c) f\big(m(c), c - m(c), \mathbf{T} - m(c)\big) = \int \mathrm{d}x \int \mathrm{d}\mathbf{P}^0(\mathbf{T}, c) f(x, c, \mathbf{T}) \tag{2.35}$$

holds for any nonnegative measurable function f on $\mathbb{R}^d \times \mathbb{C}_0 \times \mathbb{T}$. Its normalised marginal on $\mathbb{C}_0$ is called the *typical cell distribution*.

Later on, we will often consider the integral over time s of the Campbell measure and the Palm measure of the projected BRTs $\mathbf{P}_s$, and it will be convenient to have a shorthand notation for these objects. So, we define the *extended Campbell measure*

$$\overline{\mathbf{C}}^{\mathbf{P}} = \int_0^1 \mathrm{d}s \int \mathbf{P}_s(\mathrm{d}\mathbf{T}_s) \sum_{c \in T_s} \delta_{(s, \mathbf{T}_s, c)} \tag{2.36}$$

and the *extended Palm measure*

$$\overline{\mathbf{P}}^0 = \int_0^1 \mathrm{d}s \int \mathrm{d}\mathbf{P}_s^0(\mathbf{T}_s, c) \delta_{(s, \mathbf{T}_s, c)}. \tag{2.37}$$

For $W \in \mathbb{P}$, we similarly define the *extended local Campbell measure*

$$\overline{\mathbf{C}}^{\mathbf{P}_W} = \int_0^1 \mathrm{d}s \int \mathbf{P}_{W,s}(\mathrm{d}\mathbf{T}_{W,s}) \sum_{c \in T_{W,s}} \delta_{(s, \mathbf{T}_{W,s}, c)}. \tag{2.38}$$

Also, we will often use the *time-integrated* version of the Palm formula (2.35), where the Campbell measure $\mathbf{C}^{\mathbf{P}}$ and the Palm measure $\mathbf{P}^0$ are replaced by their extended relatives $\overline{\mathbf{C}}^{\mathbf{P}}$ and $\overline{\mathbf{P}}^0$, respectively. For example, combining the time-integrated Palm formula with the first-moment condition (2.33) we find that the total mass of $\overline{\mathbf{P}}^0$ can be estimated by

$$\|\overline{\mathbf{P}}^0\| = \int \mathrm{d}\overline{\mathbf{P}}(s, \mathbf{T}_s) \big|\{c \in T_s : m(c) \in [1]\}\big| \le i_1(\mathbf{P}) < \infty. \tag{2.39}$$

We conclude this section with some comments on random, but not branching, tessellations $P \in \mathscr{P}(\mathbb{T})$. These can be considered as BRTs by identifying the space $\mathbb{T}$ with $\mathbb{BT}_0$. In particular, it is then clear what translation invariance means, and we can introduce the set $\mathscr{P}_\Theta(\mathbb{T})$ of all translation invariant random tessellations P that satisfy the first-moment condition

$$i_0(P) := \int P(\mathrm{d}T) |T_{[1]}| < \infty. \tag{2.40}$$

Since $i_0(\mathbf{P} \circ \boldsymbol{\pi}_0^{-1}) \le i_1(\mathbf{P})$, the initial distribution of each $\mathbf{P} \in \mathscr{P}_\Theta$ satisfies (2.40).

3. Results. Most of our results use a STIT tessellation as a reference model. Therefore, we fix throughout a locally finite reference measure Λ on $\mathbb{H}$ which is invariant under translations. Moreover, we write $\mathbf{\Pi}^\Lambda(T_0, \cdot)$ for the associated STIT kernel, as introduced ibidem.

3.1. *The role of division kernels for BRTs.* Definitions 2.8 and 2.14 provide two ways of describing how a BRT $\mathbf{P}$ may depend on a division kernel Φ, by considering either the evolution in time or the division of cells in space. Our first result implies that these two descriptions are equivalent.

THEOREM 3.1. *For each* $\mathbf{P} \in \mathscr{P}$ *and every cell division kernel* Φ, *the following statements are equivalent.*

(a) $\mathbf{P}$ *evolves according to* Φ *as specified in Definition* 2.8.
(b) $\mathbf{P}$ *is Gibbsian for* Φ *in the sense of Definition* 2.14.
(c) *For all nonnegative measurable functions* f *on* $\overline{\mathbb{BT}} \times \mathbb{C} \times \mathbb{H}$,

$$\int \mathbf{P}(\mathrm{d}\mathbf{T}) \sum_{(s,c,H)\in\mathscr{D}(\mathbf{T})} f(s, \mathbf{T}_{s-}, c, H)$$
$$= \int \mathrm{d}\overline{\mathbf{P}}(s, \mathbf{T}_s) \int \widehat{\Phi}\big(s, \mathbf{T}_s, \mathrm{d}(c, H)\big) f(s, \mathbf{T}_s, c, H).$$

If the above properties (a) to (c) hold, we will simply say that $\mathbf{P}$ *admits the division kernel* Φ, or that Φ *is a division kernel for* $\mathbf{P}$. While statements (a) and (b) elucidate the temporal and spatial roles of Φ, the equivalent statement (c) provides a characterisation of the "jump intensity measure" of $\mathbf{P}$ in terms of Φ. In particular, one finds that the division kernel of the (unconditioned) marginal process in a local window W is obtained by a natural averaging over the possible environments outside W. To state this fact, we recall that the extended measure $\overline{\mathbf{P}}$ and the extended projections $\bar{\boldsymbol{\pi}}_W$ have been introduced in and after (2.11). Further, we will need a projection that refers to the cell division procedure. Namely, for $W \in \mathbb{P}$ we introduce the projection

$$\tilde{\pi}_W : (c, H) \mapsto (c \cap W, H) \tag{3.1}$$

on $\mathbb{C} \times \mathbb{H}$, which for each $T \in \mathbb{T}$ maps the set

$$\tilde{\pi}_W^{-1}\Delta_W := \big\{(c, H) : c \in \mathbb{C}, H \in \langle c \cap W\rangle\big\}$$

onto $\Delta_W := \{(c, H) : c \in \mathbb{C}_W, H \in \langle c\rangle\}$.

COROLLARY 3.2. *If a BRT* $\mathbf{P} \in \mathscr{P}$ *admits a cell division kernel* Φ, *its projection* $\mathbf{P}_W$ *to a window* $W \in \mathbb{P}$ *is a BRT in W for the cumulative division kernel* $\widehat{\Phi}_W$, *which is defined as a regular version of the conditional measure*

$$\widehat{\Phi}_W(s, \mathbf{T}_{W,s}, B) := \mathbb{E}_{\overline{\mathbf{P}}}\big[\widehat{\Phi}(\cdot, \cdot, \tilde{\pi}_W^{-1} B) | \bar{\boldsymbol{\pi}}_W = (s, \mathbf{T}_{W,s})\big].$$

Here, B *is any measurable subset of* Δ_W.

Next, we ask for conditions under which a given BRT $\mathbf{P} \in \mathscr{P}$ admits a division kernel Φ. (The converse question of whether a BRT for a given division kernel exists will be addressed in Theorem 3.10.) As we will see, this is the case whenever $\mathbf{P}$ is locally absolutely continuous with respect to the STIT model $P\mathbf{\Pi}^{\Lambda}$ with initial distribution $P = \mathbf{P} \circ \pi_0^{-1}$, in that

$$\text{(LAC)} \qquad \mathbf{P}_W \ll P_W \mathbf{\Pi}_W^{\Lambda} \qquad \text{for all } W \in \mathbb{P};$$

recall that $P_W \mathbf{\Pi}_W^{\Lambda} = (P\mathbf{\Pi}^{\Lambda}) \circ \boldsymbol{\pi}_W^{-1}$ by (2.20).

We note in passing that (LAC) also implies that the realisations of $\mathbf{P}$ almost surely exhibit a "tame" geometry. Namely, in the planar case, they show exactly one type of vertices, the so-called T-vertices, at which an endpoint of a line segment hits an inner point of another line segment (provided this holds already for the initial tessellation); see [18, 20] and the references cited therein.

THEOREM 3.3. *For each $\mathbf{P} \in \mathscr{P}$ satisfying* (LAC) *there exists a division kernel Φ for $\mathbf{P}$. Moreover, if $\mathbf{P}$ is also invariant under translations, one can achieve that Φ is covariant in the sense that*

$$\widehat{\Phi}(s, \mathbf{T}_s, \vartheta_x^{-1} \cdot) = \widehat{\Phi}(s, \mathbf{T}_s - x, \cdot) \tag{3.2}$$

for all $x \in \mathbb{R}^d$ and all $(s, \mathbf{T}_s) \in \overline{\mathbb{BT}}$.

Stated differently, the preceding theorem says that every $\mathbf{P} \in \mathscr{P}$ satisfying (LAC) is Gibbsian for some Φ. This is analogous to similar results in standard Gibbs theory (cf. [13], Theorem 2.30, or [14], Theorem V.2.2a). We note further that, by Corollary 3.2, the covariance property (3.2) implies that also the local division kernels can be chosen to be covariant in the sense that

$$\widehat{\Phi}_W(s, \mathbf{T}_{W,s}, \vartheta_x^{-1} \cdot) = \widehat{\Phi}_{W-x}(s, \mathbf{T}_{W,s} - x, \cdot) \tag{3.3}$$

for all $x \in \mathbb{R}^d$, $(s, \mathbf{T}_{W,s}) \in \overline{\mathbb{BT}}_W$ and $W \in \mathbb{P}$.

3.2. *The inner entropy density.* We now turn to a "thermodynamic" investigation of translation invariant BRTs. Our goal in this subsection is an appropriate notion of entropy. Recall that the relative entropy, or Kullback–Leibler divergence, between two probability measures μ and ν on a common measurable space is defined to be $\mathcal{H}(\mu; \nu) = \int \log f \, d\mu$ if $\mu \ll \nu$ with Radon–Nikodym density f, and $+\infty$ otherwise. It can also be written in the form

$$\mathcal{H}(\mu; \nu) = \int \varrho(f) \, d\nu, \tag{3.4}$$

where ϱ is the nonnegative convex function

$$\varrho : a \mapsto 1 - a + a \log a. \tag{3.5}$$

The formula (3.4) readily shows that $\mathcal{H}(\mu;\nu) \geq 0$ with equality precisely when $\mu = \nu$. We can also take it as the definition of relative entropy in the more general case when μ and ν are finite, not necessarily normalised measures.

Further, if $\mathcal{A}$ is a sub-σ-field of the underlying σ-field then the *conditional relative entropy given* $\mathcal{A}$ is defined as

$$\mathcal{H}(\mu;\nu|\mathcal{A}) = \int \mathcal{H}\big(\mu_{\mathcal{A}}(\cdot|x);\nu_{\mathcal{A}}(\cdot|x)\big)\mu(\mathrm{d}x), \tag{3.6}$$

where $\mu_{\mathcal{A}}(\cdot|x)$ and $\nu_{\mathcal{A}}(\cdot|x)$ are conditional measure kernels given $\mathcal{A}$ for μ and ν, respectively (provided such kernels exist).

In our setup, we take the STIT model for Λ as our reference measure and introduce an "inner" entropy as follows. Recall the definition (2.30) of $\boldsymbol{\pi}_W^{0,\partial}$ and its associated σ-field $\mathcal{B}_W^{0,\partial} = \sigma(\boldsymbol{\pi}_W^{0,\partial})$, and Example 2.13 for the definition of the kernel $\mathbf{G}_W^{\Lambda}$.

DEFINITION 3.4. Let $\mathbf{P} \in \mathscr{P}$ be a BRT and $W \in \mathbb{P}$. The *inner entropy of* $\mathbf{P}$ *in* W is then defined by

$$\mathcal{H}_W^{\mathrm{in}}(\mathbf{P}) := \mathcal{H}\big(\mathbf{P}_W; P_{W,0}\boldsymbol{\Pi}_W^{\Lambda}|\mathcal{B}_W^{0,\partial}\big) = \mathcal{H}\big(\mathbf{P}_W; \mathbf{P}_W^{0,\partial} \otimes \mathbf{G}_W^{\Lambda}\big). \tag{3.7}$$

(According to physical convention we should add a minus sign, but here we prefer to ignore this convention.)

So, the attribute "inner" means that this entropy compares the evolution of $\mathbf{P}$ with that of the STIT model only for those cells that are completely contained in W, while the evolution of all other cells hitting W is ignored. The idea of using a conditional, "inner" entropy without boundary effects has been exploited before by Föllmer and Snell [12] in the setup of Gibbs measures on general graphs.

Next, let $[n] := [-n/2, n/2]^d$ denote the closed centred cube of volume n^d. For a translation invariant BRT $\mathbf{P} \in \mathscr{P}_\Theta$, one expects that the limiting inner entropy per unit volume

$$\lim_{n\to\infty} n^{-d}\mathcal{H}_{[n]}^{\mathrm{in}}(\mathbf{P})$$

exists, which is then called the *inner entropy density* of $\mathbf{P}$ (relative to the reference STIT cutting rule Λ). Indeed, our result is the following; see (2.37) for the definition of the extended Palm measure $\overline{\mathbf{P}}^0$.

THEOREM 3.5. *For each* $\mathbf{P} \in \mathscr{P}_\Theta$, *there exists the possibly infinite limit*

$$h^{\mathrm{in}}(\mathbf{P}) := \lim_{n\to\infty} n^{-d}\mathcal{H}_{[n]}^{\mathrm{in}}(\mathbf{P}).$$

If this limit is finite, $\mathbf{P}$ *admits a translation covariant division kernel* Φ, *and*

$$\begin{aligned} h^{\mathrm{in}}(\mathbf{P}) &= \mathcal{H}\big(\overline{\mathbf{P}}^0 \otimes \Phi; \overline{\mathbf{P}}^0 \otimes \Lambda^*\big) \\ &= \int \mathrm{d}\overline{\mathbf{P}}^0(s,\mathbf{T}_s,c)\mathcal{H}\big(\Phi(s,\mathbf{T}_s,c,\cdot); \mathbb{1}_{\langle c\rangle}\Lambda\big). \end{aligned} \tag{3.8}$$

So, the inner entropy density $h^{\mathrm{in}}(\mathbf{P})$ is the conditional relative entropy of its division kernel Φ with respect to Λ^* when the branching tessellation and its cell are selected according to the extended Palm measure $\overline{\mathbf{P}}^0$. In particular, if $h^{\mathrm{in}}(\mathbf{P})$ is finite then the division kernel Φ of $\mathbf{P}$ admits a Radon–Nikodym density with respect to Λ^*.

It is natural to expect that the relative entropy density is affine and lower semi-continuous with compact level sets, at least under some natural caveats. We show this for a topology that is finer than the common weak topology, but is not metrisable. Namely, we define the *topology* τ_{loc} *of local convergence on* $\mathscr{P}$ as the coarsest topology for which the mapping $\mathbf{P} \mapsto \int f \,\mathrm{d}\mathbf{P}$ is continuous for every bounded local function f. It is then clear that $\mathscr{P}_\Theta$ is closed in $\mathscr{P}$. Recalling the definition (2.33) of the hitting intensity $i_1(\mathbf{P})$, we can then state the following.

THEOREM 3.6. *The inner entropy density h^{in} is affine and lower semi-continuous in τ_{loc}. Moreover, for any two constants $0 \le \beta, \gamma < \infty$ and every $P \in \mathscr{P}_\Theta(\mathbb{T})$, the restricted level set*

$$\mathscr{P}_{\Theta,P,\beta,\gamma} := \{\mathbf{P} \in \mathscr{P}_\Theta : \mathbf{P} \circ \boldsymbol{\pi}_0^{-1} = P, i_1(\mathbf{P}) \le \beta, h^{\mathrm{in}}(\mathbf{P}) \le \gamma\}$$

is compact and sequentially compact in τ_{loc}.

3.3. *Variational principle and existence.* Here, we change our perspective: rather than describing a given BRT in terms of its division kernel Φ, we will now suppose that a "nice" division kernel Ψ is given in advance. As we will see, Ψ gives rise to an "inner energy" functional on $\mathscr{P}_\Theta$, and further to an associated "inner free energy", which in turn leads to a variational principle and an existence proof for BRTs with division kernel Ψ. Here are the conditions on Ψ we need.

DEFINITION 3.7. Let us call a division kernel Ψ *moderate* if there exists a measurable density function ψ on the set

$$\{(s, \mathbf{T}_s, c, H) : 0 \le s \le 1, \mathbf{T}_s \in \mathbb{BT}_s, c \in T_s, H \in \langle c\rangle\}$$

satisfying

$$\Psi(s, \mathbf{T}_s, c, \mathrm{d}H) = \psi(s, \mathbf{T}_s, c, H)\mathbb{1}_{\langle c\rangle}(H)\Lambda(\mathrm{d}H)$$

such that the following holds for all arguments:

(M1) ψ is *covariant* under translations, in that

$$\psi(s, \mathbf{T}_s, c, H) = \psi(s, \mathbf{T}_s - x, c - x, H - x)$$

for all $x \in \mathbb{R}^d$.

(M2) ψ has *bounded range*, meaning that there exists a constant $0 \le r = r_\Psi < \infty$ such that $\psi(s, \mathbf{T}_s, c, H) = \psi(s, \mathbf{T}'_s, c, H)$ whenever $\mathbf{T}_{c+B_r,s} = \mathbf{T}'_{c+B_r,s}$. Here, B_r stands for the closed centred ball with radius r.

(M3) ψ is *bounded and bounded away from zero*, that is, there exists a constant $\kappa_\Psi < \infty$ such that $|\log \psi| \le \kappa_\Psi$.

(M4) Ψ is *approximately STIT for large cells*, which is to say that there exists a constant $\kappa'_\Psi < \infty$ such that

$$\int_{\langle c\rangle} \Lambda(\mathrm{d}H)\big|\psi(s, \mathbf{T}_s, c, H) - 1\big| \le \kappa'_\Psi.$$

[In view of the boundedness assumption (M3), this condition involves only the cells c for which $\Lambda(\langle c\rangle)$ is large.]

To give some understanding of these assumptions, we set up an analogy with the unbounded spin systems of classical statistical mechanics. A branching tessellation $\mathbf{T}_s$ at some time s may be viewed as a collection of (unbounded) "spins" that consist of cells together with their prospective cutting hyperplanes and are located at the sites $m(c)$, $c \in T_s$, of $\mathbb{R}^d$. The interaction energy of a "spin" (c, H) at time s with its surrounding tessellation $\mathbf{T}_s$ is given by $-\log \psi(s, \mathbf{T}_s, c, H)$. Assumption (M1) then expresses a natural spatial homogeneity, and (M3) the uniform boundedness of the local energies. Assumption (M2) stipulates that the range of interaction is bounded—in the units of real space, not in the units of the graph of sites which is random and difficult to handle. In particular, $\psi(s, \mathbf{T}_s, c, H)$ may depend at least on the evolution of all cells completely inside the r-neighbourhood $c + B_r$ of c (which typically contains most adjacent cells if r is chosen large enough). It may also depend on the colours of all cells that hit but are not contained in $c + B_r$; this is because the colour remains unchanged if a cell is intersected with a region. Finally, (M4) means that the interacting system is close to the noninteracting reference system unless the "spins" are suitably confined. This type of assumption is quite common for interacting systems of unbounded spins; we need it also here, although it excludes the possibility that Ψ is scale-invariant.

Obviously, the STIT kernel $\Psi = \Lambda^*$ of Example 2.9 is moderate. More generally, assumptions (M1)–(M3) hold for the cell-driven division kernels in (2.21) whenever the density φ there is uniformly bounded from above and away from zero; (M4) can be achieved by setting $\Psi = \Lambda^*$ for cells with large radius. In Example 2.10, (M1) trivially holds, (M2) holds for each $r > 0$, and (M3) follows from the assumptions on β stated there. Example 2.11 violates the bounded-range property (M2) in the most extreme way conceivable.

A moderate division kernel induces a functional on $\mathscr{P}_\Theta$ which, in analogy to the standard Gibbs theory, may be called the (negative) *inner energy in W for Ψ*, and is defined by

$$\mathcal{U}_W^{\mathrm{in}}(\mathbf{P}; \Psi) = \int \mathbf{P}(\mathrm{d}\mathbf{T}) \sum_{(s,c,H)\in\mathscr{D}(\mathbf{T})\,:\, c\subset W} \log \psi(s, \mathbf{T}_{s-}, c, H); \tag{3.9}$$

$\mathbf{P} \in \mathscr{P}_\Theta$. (Note that in statistical mechanics the energy always appears negatively in the exponent; so it should not be surprising that $\log \psi$ shows up here. But we suppress the minus sign.) Likewise, there is a term which comes from a normalisation (in our case of the distribution of jump times), and thus may be considered as an analog of the pressure in statistical mechanics. In the present setup, however, this quantity is not only a functional of Ψ, but also of the BRTs $\mathbf{P} \in \mathscr{P}_\Theta$, namely,

$$\mathcal{V}_W^{\text{in}}(\mathbf{P};\Psi) = \int d\overline{\mathbf{P}}(s,\mathbf{T}_s) \sum_{c\in T_s\,:\,c\subset W} \int_{\langle c\rangle} \Lambda(dH)\big(\psi(s,\mathbf{T}_s,c,H)-1\big). \tag{3.10}$$

THEOREM 3.8. *For every moderate division kernel Ψ and every $\mathbf{P} \in \mathscr{P}_\Theta$ admitting a covariant division kernel Φ, the following finite limits exist and can be identified*:

$$u^{\text{in}}(\mathbf{P};\Psi) := \lim_{n\to\infty} n^{-d}\,\mathcal{U}_{[n]}^{\text{in}}(\mathbf{P}_{[n]};\Psi) = \int \log\psi\, d\overline{\mathbf{P}}^0 \otimes \Phi,$$

$$v^{\text{in}}(\mathbf{P};\Psi) := \lim_{n\to\infty} n^{-d}\,\mathcal{V}_{[n]}^{\text{in}}(\mathbf{P}_{[n]};\Psi) = \int (\psi-1)\, d\overline{\mathbf{P}}^0 \otimes \Lambda^*.$$

In particular, $|u^{\text{in}}(\mathbf{P};\Psi)| \le \kappa_\Psi i_1(\mathbf{P})$ *and* $|v^{\text{in}}(\mathbf{P};\Psi)| \le \kappa'_\Psi i_1(\mathbf{P})$.

The energy terms above can be combined with the inner entropy density to define the *inner excess free energy density of* $\mathbf{P}$ *for* Ψ, namely,

$$h^{\text{in}}(\mathbf{P};\Psi) := h^{\text{in}}(\mathbf{P}) - u^{\text{in}}(\mathbf{P};\Psi) + v^{\text{in}}(\mathbf{P};\Psi), \tag{3.11}$$

where the right-hand side is set equal to $+\infty$ if $h^{\text{in}}(\mathbf{P}) = +\infty$. In fact, in the finite case it will turn out that

$$h^{\text{in}}(\mathbf{P};\Psi) = \int d\overline{\mathbf{P}}^0(s,\mathbf{T}_s,c)\mathcal{H}\big(\Phi(s,\mathbf{T}_s,c,\cdot);\Psi(s,\mathbf{T}_s,c,\cdot)\big), \tag{3.12}$$

where Φ is a division kernel for $\mathbf{P}$. The following variational principle for BRTs is then immediate.

THEOREM 3.9. *Let Ψ be any moderate division kernel. A BRT $\mathbf{P} \in \mathscr{P}_\Theta$ then admits Ψ as its division kernel if and only if $h^{\text{in}}(\mathbf{P};\Psi) = 0$.*

In particular, this can be used to prove the following result.

THEOREM 3.10. *For any moderate division kernel Ψ and every $P \in \mathscr{P}_\Theta(\mathbb{T})$, there exists a translation invariant BRT $\mathbf{P} \in \mathscr{P}_\Theta$ with initial distribution P and division kernel Ψ.*

There is a large variety of initial random tessellations P to which this existence theorem applies. The most common examples are the Poisson–Voronoi tessellation, the Poisson–Delaunay tessellation, and the Poisson hyperplane tessellation, which are well known to satisfy the moment condition (2.40). Further examples are the Delaunay tessellations that are constructed from tempered Gibbsian point processes with tile interaction, as studied in [9]. Unfortunately, we cannot allow a start in a degenerate tessellation with the full space $\mathbb{R}^d$ as its only cell (of any colour), which would be of major interest; cf. the discussion in Example 2.9 on the STIT measure $\mathbf{\Pi}^{\Lambda,\infty}$. However and as already indicated above, we can choose the initial distribution P to be the time-δ distribution of $\mathbf{\Pi}^{\Lambda,\infty}$ for some small $\delta > 0$. Up to a time shift, this means that there exists a BRT with degenerate start for any moderate division kernel Ψ with an initial cutoff of the form $\psi(s,\cdot,\cdot,\cdot) = 1$ for $0 \le s < \delta$ and some small δ. In the special case of shape-driven tessellations as in (2.21), the existence of a BRT with degenerate initial tessellation has been proved in [26] under regularity assumptions.

Since $h^{\text{in}}(\cdot;\Psi)$ is affine, the last two theorems imply the following.

COROLLARY 3.11. *For any moderate division kernel Ψ, the convex set $\mathscr{G}_\Theta(\Psi)$ of all translation invariant BRTs admitting Ψ is a face of $\mathscr{P}_\Theta$. That is, the extremal elements of $\mathscr{G}_\Theta(\Psi)$ are in fact extremal in $\mathscr{P}_\Theta$, and thereby ergodic under translations.*

It is clear that for each ergodic $\mathbf{P} \in \mathscr{G}_\Theta(\Psi)$ its initial distribution $P = \mathbf{P} \circ \boldsymbol{\pi}_0^{-1}$ is also ergodic. The converse holds whenever the correspondence between an initial distribution $P \in \mathscr{P}_\Theta(\mathbb{T})$ and its associated $\mathbf{P} \in \mathscr{G}_\Theta(\Psi)$ is one-to-one. This, however, does not hold in general, as our concluding remark shows.

REMARK 3.12. *Uniqueness and phase transition*. It is natural to ask whether or not the convex set $\mathscr{G}(P,\Phi)$ of all BRTs with initial distribution $P \in \mathscr{P}(\mathbb{T})$ and division kernel Φ is a singleton. In general, this is not the case. To provide an example, let $d = 2$ and consider the division kernel Φ defined in equation (2.22) of Example 2.11. Let $T_{\text{reg}} = \{[1] + i : i \in \mathbb{Z}^2\}$ be the regular tessellation of $\mathbb{R}^2$ into unit squares and $P \in \mathscr{P}_\Theta(\mathbb{T})$ be given by $P = \int_{[1]} \mathrm{d}x\, \delta_{T_{\text{reg}}-x}$. Further, let $\mathbf{P}^{\text{hor}}$ be the STIT tessellation with initial distribution P and driving measure Λ_{hor} as introduced in Example 2.11, and define $\mathbf{P}^{\text{vert}}$ analogously. It is then clear that these BRTs live on the spaces $\mathbb{T}_{\text{hor}}$, respectively, $\mathbb{T}_{\text{vert}}$ for all positive times. As a consequence, $\mathbf{P}^{\text{hor}}$ and $\mathbf{P}^{\text{vert}}$ are two distinct BRTs which both belong to $\mathscr{G}_\Theta(\Phi)$ and have the same initial distribution P.

Although the infinite-range interaction of this example is somewhat artificial, we learn that uniqueness does not hold automatically. Instead, the phenomenon of nonuniqueness, or phase transition, which is a central issue of statistical mechanics, shows up also in the present setting. In analogy to standard results on Gibbs measures (cf. [13], Section 8.3), we will show in Proposition 4.17 below that uniqueness does hold for suitable division kernels of bounded range in one spatial dimension. Uniqueness is also known in the noninteracting case (2.21) when the initial tessellation is degenerate and the density φ exhibits some regularity properties [26]. We leave it to the future to find sufficient conditions for uniqueness in higher dimensions as well as examples of bounded-range division kernels exhibiting phase transition. In fact, Figure 2 (right) suggests that a phase transition might already occur for the (moderate) model of Example 2.10.

4. Proofs.

4.1. *Some properties of local BRTs.* Before entering into the proofs of our results, we will establish some auxiliary properties of local BRTs. First we will express the local evolution of a BRT in a more explicit form. Throughout this section, we let $W \in \mathbb{P}_\cup$ be an arbitrary window. For any division kernel Φ_W in W, we introduce the abbreviation

$$\hat{\phi}_W(a, b; \mathbf{T}) = \int_a^b \hat{\phi}_W(s, \mathbf{T}_s)\, \mathrm{d}s, \tag{4.1}$$

where $0 \le a < b \le 1$, $\mathbf{T} \in \mathbb{BT}_W$ and $\hat{\phi}_W(s, \mathbf{T}_s) = \widehat{\Phi}_W(s, \mathbf{T}_s, T_s \times \langle W \rangle)$ is as in Remark 2.7. For every $T_0 \in \mathbb{T}_W$, we define a measure on $\mathbb{BT}_W$ by

$$\begin{aligned} \mathbf{\Pi}_W^{\Phi_W}(T_0, \cdot) &= \sum_{n \ge 0} \int \cdots \int_{\{0 \le s_1 < \cdots < s_n \le 1\}} \mathrm{d}s_1 \cdots \mathrm{d}s_n \prod_{i=1}^n \int \widehat{\Phi}_W\bigl(s_i, \mathbf{T}_{s_{i-1}}, \mathrm{d}(c_i, H_i)\bigr) \\ &\quad \times \exp\bigl[-\hat{\phi}_W(0, 1; \mathbf{T})\bigr] \mathbb{1}_{\{\boldsymbol{\pi}_0(\mathbf{T}) = T_0, \mathscr{D}(\mathbf{T}) = \{(s_i, c_i, H_i) : 1 \le i \le n\}\}} \delta_{\mathbf{T}}, \end{aligned} \tag{4.2}$$

where $s_0 := 0$ and the last indicator function simply means that $\mathbf{T}$ is the unique branching tessellation which starts from T_0 and is successively defined by the division events (s_i, c_i, H_i); recall (2.5).

LEMMA 4.1. *Let Φ_W be a division kernel and $\mathbf{P}_W \in \mathscr{P}(\mathbb{BT}_W)$ a BRT in W. Then the following statements are equivalent.*

(a) $\mathbf{P}_W$ *admits the division kernel* Φ_W.

(b) $\mathbf{\Pi}_W^{\Phi_W}(T_0, \cdot)$ *is the conditional distribution of* $\mathbf{P}_W$ *given* $\boldsymbol{\pi}_0(\mathbf{T}) = T_0$.

(c) *For every nonnegative measurable function f on $\overline{\mathbb{BT}}_W \times \mathbb{C}_W \times \langle W \rangle$,*

$$\begin{aligned} &\int \mathbf{P}_W(\mathrm{d}\mathbf{T}) \sum_{(s,c,H) \in \mathscr{D}(\mathbf{T})} f(s, \mathbf{T}_{s-}, c, H) \\ &\quad = \int_0^1 \mathrm{d}s \int \mathbf{P}_{W,s}(\mathrm{d}\mathbf{T}_s) \int \widehat{\Phi}_W\bigl(s, \mathbf{T}_s, \mathrm{d}(c, H)\bigr) f(s, \mathbf{T}_s, c, H). \end{aligned} \tag{4.3}$$

(d) $\mathbf{P}_W$ *has the infinitesimal generators $\mathbb{L}_{W,s}^{\Phi_W}$ of* (2.15), *in that the forward equation*

$$\int g\, \mathrm{d}\mathbf{P}_{W,t} - \int g\, \mathrm{d}\mathbf{P}_{W,0} = \int_0^t \mathrm{d}s \int \mathrm{d}\mathbf{P}_{W,s}\, \mathbb{L}_{W,s}^{\Phi_W} g \tag{4.4}$$

holds for all $t \in [0, 1]$ and bounded measurable functions g on $\mathbb{BT}_W$.

PROOF. (a) *implies* (b). Recall the recursion steps from Remark 2.7. For $i = 0$, $T_0 = \mathbf{T}_0$ is chosen according to the distribution $\mathbf{P}_{W,0}$. For each $i \geq 1$, conditionally on the first $(i-1)$ division events, the ith division event (s_i, c_i, H_i) for a random tessellation $\mathbf{T}$ with division rule Φ_W is chosen according to the distribution

$$\exp\big[-\hat{\phi}_W(s_{i-1}, s_i; \mathbf{T})\big]\,\mathrm{d}s_i\,\widehat{\Phi}_W\big(s_i, \mathbf{T}_{s_{i-1}}, \mathrm{d}(c_i, H_i)\big);$$

here we have used that $\mathbf{T}_s = \mathbf{T}_{s_{i-1}}$ for $s \in [s_{i-1}, s_i)$. So, on the event $\{|\mathscr{D}(\mathbf{T})| = n\}$, the joint distribution of the elements of $\mathscr{D}(\mathbf{T})$ is the product of these conditional measures for $i = 1, \ldots, n$, times the probability that s_n is the last division time before 1, which is $\exp[-\hat{\phi}_W(s_n, 1; \mathbf{T})]$.

(b) *implies* (c). Fix any initial tessellation T_0. On the set of all $\mathbf{T}$ with fixed number $n := |\mathscr{D}(\mathbf{T})| \geq 1$ of division events, the measure $\mathbf{\Pi}_W^{\Phi_W}(T_0, \cdot)$ has a product structure. The elements of $\mathscr{D}(\mathbf{T})$ can be labeled with $i \in \{1, \ldots, n\}$ according to their temporal order. For each i, we extract the ith term from the product, omit its index i, and separate the terms concerning the division events before and after time $s = s_i$. That is, we write $\mathscr{D}(\mathbf{T}) = \mathscr{D}' \cup \{(s, c, H)\} \cup \mathscr{D}''$ and separate the respective conditional measures. By Fubini's theorem, s can be considered as fixed. For given s, $\mathscr{D}'$ has the same distribution as $\mathscr{D}(\mathbf{T}_s)$, but we still have the condition that the extracted division event (s, c, H) has rank i in $\mathscr{D}(\mathbf{T})$. This condition disappears by summing over i and $n \geq i$. Finally, the division events in $\mathscr{D}''$ can be integrated out because these do not enter into $f(s, \mathbf{T}_{s-}, c, H)$, and an integration over T_0 gives (c).

(c) *implies* (d). Applying equation (4.3) to the function

$$f(s, \mathbf{T}, c, H) = \mathbb{1}_{[0,t]}(s)\big[g\big(\oslash_{s,c,H}(\mathbf{T}_s)\big) - g(\mathbf{T}_s)\big],$$

we find that for each $t \in [0, 1]$

$$\begin{aligned}
&\int_0^t \mathrm{d}s \int \mathrm{d}\mathbf{P}_{W,s}\,\mathbb{L}_{W,s}^{\Phi_W} g \\
&\qquad = \int \mathbf{P}_W(\mathrm{d}\mathbf{T}_W) \sum_{(s,c,H)\in\mathscr{D}(\mathbf{T}_W)\,:\, s\leq t} \big[g(\mathbf{T}_{W,s}) - g(\mathbf{T}_{W,s-})\big] \\
&\qquad = \int \mathbf{P}_W(\mathrm{d}\mathbf{T}_W)\big[g(\mathbf{T}_{W,t}) - g(\mathbf{T}_{W,0})\big] \\
&\qquad = \int g\,\mathrm{d}\mathbf{P}_{W,t} - \int g\,\mathrm{d}\mathbf{P}_{W,0}.
\end{aligned}$$

(d) *implies* (a). In principle, this follows from [11] which, however, makes use of a time-continuity condition on Φ_W. We thus indicate a direct argument. For brevity, we omit most indices referring to W. Let $0 \leq s < t \leq 1$ and $\mathbf{P}_{t|\mathbf{T}_s}$ be a regular version of the conditional probability of $\mathbf{P}_{W,t}$ given $\boldsymbol{\pi}_{W,s} = \mathbf{T}_s$. Using (4.4) for a function g of the form $g = \mathbb{1}_A \mathbb{1}_B$ with $A \in \mathcal{B}_{W,s}$ and $B \in \sigma(\mathbf{T} \mapsto (T_{W,u})_{s<u\leq 1})$ and varying A, one readily finds that

$$\mathbf{P}_{t|\mathbf{T}_s}(B) - \delta_{\mathbf{T}_s}(B) = \int_s^t \mathrm{d}u \int \mathrm{d}\mathbf{P}_{u|\mathbf{T}_s}\,\mathbb{L}_u^{\Phi}\mathbb{1}_B \tag{4.5}$$

for almost all $\mathbf{T}_s$. We now fix $\mathbf{T}_s$ and think of each $\mathbf{T}_u$ as an element of $\mathbb{BT}_W$ which is constant on $[u, 1]$. Also, for $\mathbf{T} \in \mathbb{BT}_W$ we let $\tau(\mathbf{T})$ be the time of the first jump of $\mathbf{T}$ after time s, which is set equal to ∞ when there is no jump during $[s, 1]$. Setting $B = \{\mathbf{T} \in \mathbb{BT}_W : \tau(\mathbf{T}) > 1\}$, we then find from (4.5) that

$$\mathbf{P}_{t|\mathbf{T}_s}(\tau > t) = 1 - \int_s^t \mathrm{d}u\, \hat{\phi}(u, \mathbf{T}_s)\mathbf{P}_{u|\mathbf{T}_s}(\tau > u)$$

and, therefore, $\mathbf{P}_{t|\mathbf{T}_s}(\tau > t) = \exp[-\hat{\phi}(s, t; \mathbf{T}_s)]$. In other words, τ has the conditional distribution used in Remark 2.7.

Next, let $\Gamma \subset \mathbb{C} \times \mathbb{H}$ be measurable and

$$\oslash_\Gamma(T_s) = \{\oslash_{c,H}(T_s) : c \in T_s, (c, H) \in \Gamma\}.$$

Consider the set $B = \{\mathbf{T} \in \mathbb{BT}_W : T_1 \in \oslash_\Gamma(T_s)\}$ and let $\tau_2(\mathbf{T})$ the time of the second jump of $\mathbf{T}$ after s (which again is set equal to ∞ if no second jump exists). Then $\mathbb{1}_B(\mathbf{T}_u) = \mathbb{1}_{\{\tau \le u < \tau_2, T_\tau \in \oslash_\Gamma(T_s)\}}(\mathbf{T})$ for $u > s$, and (4.5) implies that

$$\begin{aligned}
&\mathbf{P}_{t|\mathbf{T}_s}(\tau \le t < \tau_2, T_\tau \in \oslash_\Gamma(T_s)) \\
&\quad = \int_s^t \mathrm{d}u\, \widehat{\Phi}(u, \mathbf{T}_s, \Gamma)\mathbf{P}_{u|\mathbf{T}_s}(\tau > u) \\
&\qquad - \int_{(s,t]\times\oslash_\Gamma(T_s)} \mathbf{P}_{t|\mathbf{T}_s}((\tau, \mathbf{T}_\tau) \in \mathrm{d}(v, \mathbf{T}_v)) \int_v^t \mathrm{d}u\, \hat{\phi}(u, \mathbf{T}_v)\mathbf{P}_{t|\mathbf{T}_v}(\tau > u).
\end{aligned}$$

Using the explicit conditional distribution of τ derived above, we thus find that the first term on the right-hand side of the above equation is equal to

$$\int_s^t \mathrm{d}u\, \widehat{\Phi}(u, \mathbf{T}_s, \Gamma) \exp[-\hat{\phi}(s, u; \mathbf{T}_s)],$$

whereas the second term equals

$$\begin{aligned}
&\int_{(s,t]\times\oslash_\Gamma(T_s)} \mathbf{P}_{t|\mathbf{T}_s}((\tau, \mathbf{T}_\tau) \in \mathrm{d}(v, \mathbf{T}_v))\mathbf{P}_{t|\mathbf{T}_v}(v < \tau \le t) \\
&\qquad = \mathbf{P}_{t|\mathbf{T}_s}(\tau_2 \le t, T_\tau \in \oslash_\Gamma(T_s)).
\end{aligned}$$

We thus arrive at the equation

$$\mathbf{P}_{t|\mathbf{T}_s}(\tau \le t, T_\tau \in \oslash_\Gamma(T_s)) = \int_s^t \mathrm{d}u\, \exp[-\hat{\phi}(s, u; \mathbf{T}_s)]\widehat{\Phi}(u, \mathbf{T}_s, \Gamma),$$

which assures that $(\tau, \mathbf{T}_\tau)$ has the correct conditional distribution of Remark 2.7. □

Note that the joint integrating measure on the right-hand side of (4.3) can be written in the concise form $\overline{\mathbf{P}}_W \otimes \widehat{\Phi}_W$ or, equivalently, $\overline{\mathbf{C}}^{\mathbf{P}_W} \otimes \Phi$, where $\overline{\mathbf{P}}_W = \overline{\mathbf{P}} \circ \bar{\boldsymbol{\pi}}_W^{-1}$, $\overline{\mathbf{P}}$ is given by (2.11) and $\overline{\mathbf{C}}^{\mathbf{P}_W}$ by (2.38). We will switch between both representations according to convenience.

COROLLARY 4.2. *Let $\mathbf{P}_W, \mathbf{Q}_W \in \mathscr{P}(\mathbb{BT}_W)$ be two BRTs in W. Suppose $\mathbf{Q}_W$ admits a division kernel Ψ_W, and $\mathbf{P}_W \ll \mathbf{Q}_W$. Then there exists a measurable function $\varphi_W(s, \mathbf{T}_s, c, H) \geq 0$ such that the measure kernel*

$$\Phi_W(s, \mathbf{T}_s, c, \mathrm{d}H) := \varphi_W(s, \mathbf{T}_s, c, H)\Psi_W(s, \mathbf{T}_s, c, \mathrm{d}H)$$

is a division kernel for $\mathbf{P}_W$.

PROOF. For brevity, we introduce the measure kernel

$$\mathbf{D}_W(\mathbf{T}, \cdot) = \sum_{(s,c,H)\in\mathscr{D}(\mathbf{T})} \delta_{(s,\mathbf{T}_{s-},c,H)} \tag{4.6}$$

for $\mathbf{T} \in \mathbb{BT}_W$. The integration on the left-hand side of (4.3) is then with respect to the measure $\mathbf{P}_W\mathbf{D}_W$. Since $\mathbf{P}_W \ll \mathbf{Q}_W$ by assumption, it follows that $\mathbf{P}_W\mathbf{D}_W \ll \mathbf{Q}_W\mathbf{D}_W$ with a Radon–Nikodym density f, say. It also follows that $\overline{\mathbf{P}}_W \ll \overline{\mathbf{Q}}_W$ with a density g. Define

$$\varphi_W(s, \mathbf{T}_s, c, H) = f(s, \mathbf{T}_s, c, H)/g(s, \mathbf{T}_s)$$

if the denominator is positive, and zero otherwise. Then we obtain, using equation (4.3) for $(\mathbf{Q}_W, \Psi_W)$ in place of $(\mathbf{P}_W, \Phi_W)$,

$$\mathbf{P}_W\mathbf{D}_W = f(\mathbf{Q}_W\mathbf{D}_W) = f(\overline{\mathbf{Q}}_W \otimes \widehat{\Psi}) = \overline{\mathbf{P}}_W \otimes (\varphi_W\widehat{\Psi}_W).$$

In view of Lemma 4.1, this means that $\mathbf{P}_W$ admits the division kernel $\Phi_W := \varphi_W\Psi_W$. $\square$

Finally, we look at the first-moment condition (2.10).

LEMMA 4.3. *Let Φ_W be a division kernel for W, and suppose its total mass satisfies the uniform bound (2.14). Then*

$$\int \mathbf{\Pi}_W^{\Phi_W}(T_0, \mathrm{d}\mathbf{T})|T_1| \leq e^{\phi}|T_0|$$

for all initial tessellations $T_0 \in \mathbb{T}_W$. Moreover, for every $\varepsilon > 0$, one can find a number $\tau < \infty$ such that

$$\int \mathbf{\Pi}_W^{\Phi_W}(T_0, \mathrm{d}\mathbf{T})\big(|T_1| - \tau|T_0|\big)_+ \leq \varepsilon|T_0|$$

for all $T_0 \in \mathbb{T}_W$.

PROOF. Recall the description of $\mathbf{\Pi}_W^{\Phi_W}(T_0, \cdot)$ in Remark 2.7. The algorithm there implies that, for each i with $s_i \leq 1$, the holding time $s_i - s_{i-1}$ dominates an exponential time with parameter $|T_{s_{i-1}}|\phi$, independently of the previous recursion steps. Hence, the process $|T_s|$ is stochastically dominated by the Furry–Yule process $Z_s \in \mathbb{N}$ with birth rate ϕ, namely the pure birth Markov process which starts in $k = |T_0|$ and jumps from any $j \geq 1$ to $j + 1$ with rate $j\phi$. Equivalently, Z_s can be described as the branching process in which each individual, independently of all others, lives for an exponential time with parameter ϕ and then splits into two offspring. In particular, the descendance trees of each of the k ancestors are independent, and it is sufficient to look at the number of descendants at time s in each of these trees. This number is known to have the geometric distribution with mean $e^{\phi s}$. A proof of this can be found, for example, in [21], Examples 6.4, 6.8 or Exercise 6.11.

As for the second assertion, we conclude from the convexity of the function $a \mapsto (a - \tau)_+$ that

$$\big(|T_1| - \tau|T_0|\big)_+ \leq \sum_{c \in T_0} \big(|T_{c,1}| - \tau\big)_+.$$

Here, $|T_{c,1}|$ is the number of descendants of the initial cell c at time 1, which is stochastically dominated by the geometric random variable Z_1. As $\mathbb{E}(Z_1 - \tau)_+ \to 0$ as $\tau \to \infty$, the result follows immediately. □

4.2. *Significance and construction of global division kernels.* Here, we prove Theorems 3.1 and 3.3. We begin with the equivalence theorem (Theorem 3.1). Most work will be necessary for deriving the Gibbs property (b) from statement (c), the characterisation of the jump intensity measure. To this end, we need to introduce a modification of the outer projection for a given window $W \in \mathbb{P}$, which refers to a larger but bounded window $W' \in \mathbb{P}$ rather than the full space $\mathbb{R}^d$. Namely, for $W \subset W' \in \mathbb{P}$ and any $\mathbf{T}_{W'} \in \mathbb{BT}_{W'}$ let

$$\mathbf{T}_W^{W',\mathrm{out}} = \big(\{c \in T_{W',s} : c \not\subset \mathrm{int}(W)\}\big)_{0 \leq s \leq 1} \tag{4.7}$$

be the evolution of the cells hitting $W' \setminus \mathrm{int}(W)$. In particular, if $W' = W$ then $\mathbf{T}_W^{W,\mathrm{out}} = \mathbf{T}_W^{\partial}$. We also set $\mathbf{T}_W^{W',0,\mathrm{out}} = (T_{W,0}^{\mathrm{in}}, \mathbf{T}_W^{W',\mathrm{out}})$ and let $\mathcal{B}_W^{W',0,\mathrm{out}}$ denote the σ-field on $\mathbb{BT}_{W'}$ generated by the mapping $\mathbf{T}_{W'} \mapsto \mathbf{T}_W^{W',0,\mathrm{out}}$.

PROOF OF THEOREM 3.1. We establish the circle (a) ⇒ (c) ⇒ (b) ⇒ (a).

(a) *implies* (c). Let f be a bounded nonnegative measurable function on $\overline{\mathbb{BT}} \times \mathbb{C} \times \mathbb{H}$, and suppose there is some $W \in \mathbb{P}$ such that (i) $f(s, \mathbf{T}_s, c, H)$ is $\mathcal{B}_W$-measurable as a function of $\mathbf{T}$, and (ii) $f(s, \mathbf{T}_s, c, H) = 0$ unless $c \subset W$ and $|\mathscr{D}(\mathbf{T}_{W,s})| \leq K$ for some $K < \infty$. Define

$$g(\mathbf{T}) = \sum_{(s,c,H) \in \mathscr{D}(\mathbf{T})} f(s, \mathbf{T}_{s-}, c, H).$$

By assumption, g is bounded and local, and $g(\mathbf{T}_0) = 0$ for every $\mathbf{T}$. Moreover, if s is not a jump time of $\mathbf{T}_{W,s}$ then

$$g\big(\oslash_{s,c,H}(\mathbf{T}_s)\big) - g(\mathbf{T}_s) = f(s, \mathbf{T}_s, c, H)$$

and, therefore, $\mathbb{L}_s^\Phi g(\mathbf{T}_s) = \int \mathrm{d}\widehat{\Phi}(s, \mathbf{T}_s, \cdot, \cdot) f(s, \mathbf{T}_s, \cdot, \cdot)$. The forward equation thus shows that

$$\int g \,\mathrm{d}\mathbf{P} = \int_0^1 \mathrm{d}s \int \mathrm{d}\mathbf{P}_s \,\mathbb{L}_s^\Phi g = \int f \,\mathrm{d}(\overline{\mathbf{P}} \otimes \widehat{\Phi}),$$

which is (c) for our particular f. The case of general f now follows by letting $K \to \infty$ and using a monotone class argument.

(c) *implies* (b). Fix a window $W \in \mathbb{P}$ and let $\mathbf{P}_W^{\mathrm{in}}(\cdot|\mathbf{T}_W^{0,\mathrm{out}})$ be a regular version of the conditional distribution of $\boldsymbol{\pi}_W^{\mathrm{in}}$ under the condition $\boldsymbol{\pi}_W^{0,\mathrm{out}} = \mathbf{T}_W^{0,\mathrm{out}}$. We need to show that this probability kernel almost surely coincides with $\mathbf{G}_W^\Phi(\cdot|\mathbf{T}_W^{0,\mathrm{out}})$. (In particular, this will imply that the latter is almost surely well defined.) Pick any two nonnegative measurable functions $g(\mathbf{T}_W^{0,\mathrm{out}})$ and $h(s, \mathbf{T}_{W,s}^{\mathrm{in}}, c, H)$ of the indicated arguments. We suppose g is local, in that $g(\mathbf{T}_W^{0,\mathrm{out}}) = g(\mathbf{T}_W^{W',0,\mathrm{out}})$ for some $W' \in \mathbb{P}$ containing W. Consider the integral

$$\int \mathbf{P}(\mathrm{d}\mathbf{T}) g\big(\mathbf{T}_W^{0,\mathrm{out}}\big) \int \mathbf{P}_W^{\mathrm{in}}\big(\mathrm{d}\mathbf{S}|\mathbf{T}_W^{0,\mathrm{out}}\big) \sum_{(s,c,H)\in\mathscr{D}(\mathbf{S})} h(s, \mathbf{S}_{s-}, c, H). \tag{4.8}$$

By the definition of conditional distribution, this is equal to

$$\int \mathbf{P}(\mathrm{d}\mathbf{T}) g\big(\mathbf{T}_W^{0,\mathrm{out}}\big) \sum_{(s,c,H)\in\mathscr{D}(\mathbf{T}_W^{\mathrm{in}})} h\big(s, \mathbf{T}_{W,s-}^{\mathrm{in}}, c, H\big).$$

In view of the locality assumption on g, the integrand actually only depends on $\mathbf{T}_{W'}$, which is a pure jump process and therefore strongly Markov. Writing (s_i, c_i, H_i) for the ith division event of $\mathbf{T}_{W'}$ in temporal order, we can rewrite the last expression in the form

$$\sum_{i\geq 1} \int \mathbf{P}_{W'}(\mathrm{d}\mathbf{T}_{W'}) g\big(\mathbf{T}_W^{W',0,\mathrm{out}}\big) \mathbb{1}_{\{s_i<1, c_i\subset W\}}(\mathbf{T}_{W'}) h\big(s_i, \mathbf{T}_{W,s_i-}^{\mathrm{in}}, c_i, H_i\big). \tag{4.9}$$

Now, both h and the indicator function in the integrand are measurable with respect to the σ-field $\mathcal{B}_{W',s_i}$ of all events $A \in \mathcal{B}_{W'}$ with $A \cap \{s_i \leq t\} \in \mathcal{B}_{W',t}$ for all t. By the strong Markov property, we can therefore replace the function $g(\mathbf{T}_W^{W',0,\mathrm{out}})$ by its conditional expectation $g(s_i, \mathbf{T}_{W',s_i})$ relative to $\mathcal{B}_{W',s_i}$. Furthermore, the process $\mathbf{T}_W^{W',0,\mathrm{out}}$ is itself a Markov jump process. [In fact, it can be considered as the BRT in W' for the division kernel which equals $\Phi_{W'}(s, \mathbf{T}_{W'}, c, \cdot)$ if $c \not\subset W$ and is identically zero otherwise.] This means that

$$g(s_i, \mathbf{T}_{W',s_i}) = g\big(s_i, \mathbf{T}_{W,s_i}^{W',0,\mathrm{out}}\big) = g(s_i, \mathbf{T}_{W',s_i-})$$

when $c_i \subset W$. Altogether, we find that the expression (4.9) is equal to

$$\int \mathbf{P}(\mathrm{d}\mathbf{T}) \sum_{(s,c,H)\in\mathscr{D}(\mathbf{T}) : c\subset W} g(s, \mathbf{T}_{W',s-}) h\big(s, \mathbf{T}_{W,s-}^{\mathrm{in}}, c, H\big).$$

By statement (c), this in turn coincides with

$$\int_0^1 \mathrm{d}s \int \mathbf{P}_s(\mathrm{d}\mathbf{T}_s) \sum_{c\in T_s\,:\, c\subset W} \int \Phi(s, \mathbf{T}_s, c, \mathrm{d}H) g(s, \mathbf{T}_{W',s}) h\big(s, \mathbf{T}^{\mathrm{in}}_{W,s}, c, H\big),$$

which by the Markov property is equal to

$$\int_0^1 \mathrm{d}s \int \mathbf{P}(\mathrm{d}\mathbf{T}) g\big(\mathbf{T}^{0,\mathrm{out}}_W\big) \sum_{c\in T^{\mathrm{in}}_{W,s}} \int \Phi(s, \mathbf{T}_s, c, \mathrm{d}H) h\big(s, \mathbf{T}^{\mathrm{in}}_{W,s}, c, H\big).$$

Taking conditional expectation with respect to $\mathbf{T}^{0,\mathrm{out}}_W$ and using the conditional division kernel $\Phi^{\mathrm{in}}_W(\cdot|\mathbf{T}^{\mathrm{out}}_W)$ from Definition 2.12, we can rewrite this as

$$\int \mathbf{P}(\mathrm{d}\mathbf{T}) g\big(\mathbf{T}^{0,\mathrm{out}}_W\big) \int_0^1 \mathrm{d}s \int \mathbf{P}^{\mathrm{in}}_{W,s}\big(\mathrm{d}\mathbf{S}_s|\mathbf{T}^{0,\mathrm{out}}_W\big) \\ \times \int \widehat{\Phi}^{\mathrm{in}}_W\big(s, \mathbf{S}_s, c, \mathrm{d}H|\mathbf{T}^{\mathrm{out}}_W\big) h(s, \mathbf{S}_s, c, H).$$

Since the underlying spaces are Borel, a comparison of (4.8) with the preceding expression shows that, for almost all $\mathbf{T}^{0,\mathrm{out}}_W$,

$$\overline{\mathbf{P}}^{\mathrm{in}}_W\big(\cdot|\mathbf{T}^{0,\mathrm{out}}_W\big) \otimes \widehat{\Phi}^{\mathrm{in}}_W\big(\cdot|\mathbf{T}^{\mathrm{out}}_W\big) = \int \mathbf{P}^{\mathrm{in}}_W\big(\mathrm{d}\mathbf{S}|\mathbf{T}^{0,\mathrm{out}}_W\big) \sum_{(s,c,H)\in\mathscr{D}(\mathbf{S})} \delta_{(s,\mathbf{S}^{\mathrm{in}}_{s-},c,H)},$$

which corresponds to (4.3). Lemma 4.1 therefore implies that $\mathbf{P}^{\mathrm{in}}_W(\cdot|\mathbf{T}^{0,\mathrm{out}}_W)$ coincides with $\mathbf{G}^{\Phi}_W(\cdot|\mathbf{T}^{0,\mathrm{out}}_W)$. This completes the proof of the Gibbs property (b).

(b) *implies* (a). Let g be a bounded function which is $\mathcal{B}_W$-measurable for some $W \in \mathbb{P}$. For any n with $W \subset [n]$ let

$$A_n = \big\{\mathbf{T} \in \mathbb{BT} : c \cap W = \varnothing \text{ for all } c \in T_0 \text{ with } c \cap \partial[n] \neq \varnothing\big\}. \tag{4.10}$$

Obviously, $A_n \in \mathcal{B}^{\partial}_{[n],0}$. Also, since for each $T \in \mathbb{T}$ the union of all cells hitting W is contained in some $[n]$, we have $A_n \uparrow \mathbb{BT}$ as $n \to \infty$. Furthermore, if $\mathbf{T} \in A_n$ then the inner window $\mathrm{in}_{[n]}(s, \mathbf{T}^{\partial}_{[n]})$ [defined in (2.31)] contains W for all s. Using Definition 2.12 and Lemma 4.1, we thus obtain that

$$\int \mathbf{G}^{\Phi}_W\big(\mathrm{d}\mathbf{S}|\mathbf{T}^{0,\mathrm{out}}_W\big)\Big[g(\mathbf{S}_t) - g(\mathbf{S}_0) - \int_0^t \mathrm{d}s\, \mathbb{L}_s^{\Phi^{\mathrm{in}}_W(\cdot|\mathbf{T}^{\mathrm{out}}_W)} g(\mathbf{S}_s)\Big] = 0$$

for all $0 < t \le 1$ and $\mathbf{T} \in A_n$. Integrating this over $\int_{A_n} \mathbf{P}(\mathrm{d}\mathbf{T})$, applying the Gibbs property (b) and letting $n \to \infty$ we arrive at (a). $\square$

Before turning to the proof of Corollary 3.2 it is worthwhile to introduce a condensed notation for property (c) of Theorem 3.1. So, we introduce the measure kernel

$$\mathbf{D}(\mathbf{T}, \cdot) = \sum_{(s,c,H)\in\mathscr{D}(\mathbf{T})} \delta_{(s,\mathbf{T}_{s-},c,H)}$$

from $\mathbb{BT}$ to $\overline{\mathbb{BT}} \times \mathbb{C} \times \mathbb{H}$, which catches the behaviour of $\mathbf{T}$ at all cell division events; it is analogous to the kernel $\mathbf{D}_W$ within a window W, which was defined at (4.6). Statement (c) of Theorem 3.1 can then be written in the concise form

$$\mathbf{PD} := \int \mathbf{P}(\mathrm{d}\mathbf{T})\mathbf{D}(\mathbf{T}, \cdot) = \overline{\mathbf{P}} \otimes \widehat{\Phi}. \tag{4.11}$$

PROOF OF COROLLARY 3.2. Fix some $W \in \mathbb{P}$, recall the definitions of $\tilde{\pi}_W$ and Δ_W at (3.1), and note that $\mathbf{D}_W$ is supported on $\overline{\mathbb{BT}}_W \times \Delta_W$. Since

$$\mathscr{D}(\mathbf{T}_W) = \{(s, c \cap W, H) : (s, c, H) \in \mathscr{D}(\mathbf{T}), (c, H) \in \tilde{\pi}_W^{-1}\Delta_W\},$$

we have $\mathbf{D}_W(\mathbf{T}_W, A \times B) = \mathbf{D}(\mathbf{T}, \bar{\boldsymbol{\pi}}_W^{-1}A \times \tilde{\pi}_W^{-1}B)$ for all $\mathbf{T} \in \mathbb{BT}$ and all events $A \subset \overline{\mathbb{BT}}_W$ and $B \subset \Delta_W$ and, therefore, by (4.11)

$$\begin{aligned}\mathbf{P}_W\mathbf{D}_W(A \times B) &= \mathbf{PD}(\bar{\boldsymbol{\pi}}_W^{-1}A \times \tilde{\pi}_W^{-1}B) = \overline{\mathbf{P}} \otimes \widehat{\Phi}(\bar{\boldsymbol{\pi}}_W^{-1}A \times \tilde{\pi}_W^{-1}B) \\ &= \int_{\bar{\boldsymbol{\pi}}_W^{-1}A} \mathrm{d}\overline{\mathbf{P}}(s, \mathbf{T}_s)\widehat{\Phi}(s, \mathbf{T}_s, \tilde{\pi}_W^{-1}B).\end{aligned}$$

So, if $\widehat{\Phi}_W$ is defined as in the corollary then $\mathbf{P}_W\mathbf{D}_W = \overline{\mathbf{P}}_W \otimes \widehat{\Phi}_W$. Lemma 4.1 thus shows that $\mathbf{P}_W$ admits the kernel Φ_W. $\square$

Finally, we turn to the construction of division kernels for BRTs satisfying (LAC).

PROOF OF THEOREM 3.3. *Part* 1: *Extension of local division kernels*. By condition (LAC), Corollary 4.2 implies that for each $W \in \mathbb{P}$ there exists a cell division kernel Φ_W in W such that $\mathbf{P}_W$ is a BRT for Φ_W. (In fact, Φ_W is absolutely continuous with respect to Λ^*, but we do not need this here.) So, it merely remains to construct a global common extension Φ of these kernels Φ_W.

By Lemma 4.1 and the preceding proof of Corollary 3.2, we know that

$$\overline{\mathbf{P}}_W \otimes \widehat{\Phi}_W = \mathbf{P}_W\mathbf{D}_W = \mathbf{PD} \circ (\bar{\boldsymbol{\pi}}_W \otimes \tilde{\pi}_W)^{-1} \tag{4.12}$$

on $\overline{\mathbb{BT}}_W \times \Delta_W$. As a consequence, the measures $\overline{\mathbf{P}}_W \otimes \widehat{\Phi}_W$ with $W \in \mathbb{P}$ are consistent in the sense that

$$(\overline{\mathbf{P}}_{W'} \otimes \widehat{\Phi}_{W'}) \circ (\bar{\boldsymbol{\pi}}_W \otimes \tilde{\pi}_W)^{-1} = \overline{\mathbf{P}}_W \otimes \widehat{\Phi}_W \qquad \text{on } \overline{\mathbb{BT}}_W \times \Delta_W \tag{4.13}$$

for $W \subset W' \in \mathbb{P}$. To see that these measures admit a common extension, we first localise to a fixed window $V \in \mathbb{P}$. For $W \supset V$, we write $\overline{\mathbf{P}}_W \otimes 1_V \widehat{\Phi}_W$ for the restriction of $\overline{\mathbf{P}}_W \otimes \widehat{\Phi}_W$ to the set $\overline{\mathbb{BT}}_W \times \tilde{\pi}_V^{-1}\Delta_V$. These measures have a finite total mass that does not depend on W. Indeed, (4.12) and the first-moment condition (2.10) imply that

$$\begin{aligned}\|\overline{\mathbf{P}}_W \otimes 1_V \widehat{\Phi}_W\| &= \int \mathbf{P}(\mathrm{d}\mathbf{T})\big|\{(s, c, H) \in \mathscr{D}(\mathbf{T}) : (c, H) \in \tilde{\pi}_V^{-1}\Delta_V\}\big| \\ &\le \int \mathbf{P}(\mathrm{d}\mathbf{T})|T_{V,1}| < \infty.\end{aligned} \tag{4.14}$$

Since all spaces under consideration are Borel spaces, we can thus apply an abstract version of the Kolmogorov extension theorem [16], Corollary 6.15, to obtain a finite measure on $\overline{\mathbb{BT}} \times \tilde{\pi}_V^{-1}\Delta_V$, to be denoted by $\overline{\mathbf{P}} \otimes \mathbb{1}_V \widehat{\Phi}$, which satisfies

$$(\overline{\mathbf{P}} \otimes \mathbb{1}_V \widehat{\Phi}) \circ (\bar{\boldsymbol{\pi}}_W \otimes \tilde{\pi}_W)^{-1} = \overline{\mathbf{P}}_W \otimes \mathbb{1}_V \widehat{\Phi}_W$$

for all $W \in \mathbb{P}$ with $W \supset V$. Since V is arbitrary and

$$\bigcup_{V \in \mathbb{P}} \tilde{\pi}_V^{-1}\Delta_V = \Delta := \{(c, H) \in \mathbb{C} \times \mathbb{H} : H \in \langle c \rangle\},$$

the measures $\overline{\mathbf{P}} \otimes \mathbb{1}_V \widehat{\Phi}$ can be glued together to a locally finite measure $\overline{\mathbf{P}} \otimes \widehat{\Phi}$ on $\overline{\mathbb{BT}} \times \Delta$ satisfying

$$(\overline{\mathbf{P}} \otimes \widehat{\Phi}) \circ (\bar{\boldsymbol{\pi}}_W \otimes \tilde{\pi}_W)^{-1} = \overline{\mathbf{P}}_W \otimes \widehat{\Phi}_W \qquad \text{on } \overline{\mathbb{BT}}_W \times \Delta_W \tag{4.15}$$

for all $W \in \mathbb{P}$. As we have indicated by the notation, disintegration shows that this measure is indeed the product of $\overline{\mathbf{P}}$ with a locally finite measure kernel $\widehat{\Phi}$.

We next need to show that this $\widehat{\Phi}$ is really a (cumulative) division kernel. By construction, each $\widehat{\Phi}(s, \mathbf{T}_s, \cdot)$ is supported on Δ, which means that each $\Phi(s, \mathbf{T}_s, c, \cdot)$ is supported on $\langle c \rangle$. In fact, considering the set $\Delta(T_s) = \{(c, H) : c \in T_s, H \in \langle c \rangle\}$ and its complement $\neg\Delta(T_s)$, we can write

$$\begin{aligned} &\int \mathrm{d}\overline{\mathbf{P}}(s, \mathbf{T}_s)\widehat{\Phi}(s, \mathbf{T}_s, \neg\Delta(T_s)) \\ &\qquad = \int \mathrm{d}(\overline{\mathbf{P}} \otimes \widehat{\Phi})(s, \mathbf{T}_s, c, H) \lim_{W \uparrow \mathbb{R}^d} \bigl(1 - \mathbb{1}_{\tilde{\pi}_W^{-1}\Delta(T_{W,s})}(c, H)\bigr), \end{aligned}$$

and the last term vanishes by (4.15) and Fatou's lemma. So, we can conclude that, for $\overline{\mathbf{P}}$-almost all $(s, \mathbf{T}_s)$, $\widehat{\Phi}(s, \mathbf{T}_s, \cdot)$ is indeed supported on $\Delta(T_s)$, as required.

Finally, combining (4.12) and (4.15) we find that statement (c) of Theorem 3.1 holds for all f of the form

$$f(s, \mathbf{T}, c, H) = \mathbb{1}_{\langle c \cap W \rangle}(H) f_W(s, \mathbf{T}_W, c \cap W, H)$$

with some $W \in \mathbb{P}$ and a measurable function f_W. As this can be extended to general f by a monotone class argument, it follows that Φ is a division kernel for $\mathbf{P}$.

Part 2: *Averaging over translations*. Suppose now that $\mathbf{P}$ is invariant under translations, and let Φ be a global division kernel for $\mathbf{P}$, which exists by part 1 of the proof. For $x \in \mathbb{R}^d$, let ϑ_x be the spatial translation by $-x$, which acts on $\overline{\mathbb{BT}}$ via $\vartheta_x : (s, \mathbf{T}_s) \mapsto (s, \mathbf{T}_s - x)$ and, by hypothesis, leaves $\overline{\mathbf{P}}$ invariant. As before, we use the same symbol for the translation $\vartheta_x : (c, H) \mapsto (c - x, H - x)$ acting on $\mathbb{C} \times \mathbb{H}$. For $x \in \mathbb{R}^d$ let $\widehat{\Phi}^x(s, \mathbf{T}_s, \cdot) := \widehat{\Phi}(s, \mathbf{T}_s + x, \vartheta_x^{-1} \cdot)$. We first claim that

$$\overline{\mathbf{P}} \otimes \widehat{\Phi} = \overline{\mathbf{P}} \otimes \widehat{\Phi}^x \tag{4.16}$$

for all x. Indeed, let $A \in \mathcal{B}(\overline{\mathbb{BT}})$ and $B \in \mathcal{B}(\mathbb{C} \times \mathbb{H})$ be arbitrarily given. Then we can write, using the ϑ_x-invariance of $\mathbf{P}$ in the first step,

$$\begin{aligned}
(\overline{\mathbf{P}} \otimes \widehat{\Phi}^x)(A \times B) &= (\overline{\mathbf{P}} \otimes \widehat{\Phi})(\vartheta_x^{-1} A \times \vartheta_x^{-1} B) \\
&= \int \mathbf{P}(\mathrm{d}\mathbf{T}) \sum_{(s,c,H) \in \mathscr{D}(\mathbf{T}) \,:\, (c-x, H-x) \in B} \mathbb{1}_A(s, \mathbf{T}_{s-}) \\
&= \int \mathbf{P}(\mathrm{d}\mathbf{T}) \sum_{(s,c,H) \in \mathscr{D}(\mathbf{T}) \,:\, (c, H) \in B} \mathbb{1}_A(s, \mathbf{T}_{s-}) \\
&= (\overline{\mathbf{P}} \otimes \widehat{\Phi})(A \times B),
\end{aligned}$$

proving (4.16). The second and the fourth step come from Theorem 3.1(c), and in the third step we observed that $\mathscr{D}(\mathbf{T})$ consists of the shifted elements of $\mathscr{D}(\mathbf{T} - x)$ and then used again the translation invariance of $\mathbf{P}$. Equation (4.16) shows that $\widehat{\Phi}^x(\cdot, \cdot, B) = \widehat{\Phi}(\cdot, \cdot, B)$ $\overline{\mathbf{P}}$-almost surely for each B.

To obtain an everywhere covariant version of Φ, we pick a countable generator $\mathcal{G}$ of $\mathcal{B}(\mathbb{C} \times \mathbb{H})$ which is stable under intersections. We also let Γ be the set of all $(s, \mathbf{T}_s) \in \overline{\mathbb{BT}}$ which are such that $\widehat{\Phi}(s, \mathbf{T}_s, B) = \widehat{\Phi}^y(s, \mathbf{T}_s, B)$ for all $B \in \mathcal{G}$ [and thus all $B \in \mathcal{B}(\mathbb{C} \times \mathbb{H})$] and the countably many lattice elements $y \in \mathbb{Z}^d$. Then $\overline{\mathbf{P}}(\Gamma) = 1$ by (4.16). We further define the kernel

$$\widetilde{\Phi}(s, \mathbf{T}_s, \cdot) = \begin{cases} \widehat{\Phi}(s, \mathbf{T}_s, \cdot), & \text{if } (s, \mathbf{T}_s) \in \Gamma, \\ \widehat{\Lambda}^*(s, \mathbf{T}_s, \cdot), & \text{otherwise,} \end{cases}$$

where $\widehat{\Lambda}^*$ is the cumulative STIT kernel of Example 2.9. It is then clear that $\widetilde{\Phi}$ is a version of $\widehat{\Phi}$ which satisfies $\widetilde{\Phi}^y = \widetilde{\Phi}$ for all $y \in \mathbb{Z}^d$.

To achieve the covariance under the full translation group, we finally define

$$\overline{\Phi} = \int_{[1]} \widetilde{\Phi}^x \, \mathrm{d}x,$$

where $[1]$ is the centred unit cube in $\mathbb{R}^d$. Then for each $y \in \mathbb{R}^d$, we have

$$\overline{\Phi}^y = \int_{[1]+y} \widetilde{\Phi}^x \, \mathrm{d}x = \overline{\Phi}$$

because $[1] + y$ can be decomposed into finitely many pieces which are lattice translations of corresponding pieces of $[1]$. On the other hand, since

$$\overline{\mathbf{P}} \otimes \overline{\Phi} = \int_{[1]} \overline{\mathbf{P}} \otimes \widetilde{\Phi}^x \, \mathrm{d}x = \int_{[1]} \overline{\mathbf{P}} \otimes \widehat{\Phi}^x \, \mathrm{d}x = \overline{\mathbf{P}} \otimes \widehat{\Phi}$$

by (4.16), $\overline{\Phi}$ is also a version of $\widehat{\Phi}$. $\square$

We conclude this subsection with two supplements to the preceding proofs. The first deals with the consistency properties (4.13), respectively, (4.15), and the second with a localised version of the Gibbs property.

REMARK 4.4 (Consistency of kernel densities). Consider two windows $W, W' \in \mathbb{P}$ with $W \subset W'$ and a BRT $\mathbf{P} \in \mathscr{P}$ satisfying $\mathbf{P}_{W'} \ll \mathbf{P}_{W',0}\mathbf{\Pi}^{\Lambda}_{W'}$. Let φ_W and $\varphi_{W'}$ be the Λ-densities of the division kernels Φ_W and $\Phi_{W'}$ of $\mathbf{P}_W$ and $\mathbf{P}_{W'}$, which exist by Corollary 4.2. The consistency equation (4.13) then means that

$$\varphi_W(s, \mathbf{T}_{W,s}, c, H) = \int \mathbf{P}_{W',s|\mathbf{T}_{W,s}}(\mathrm{d}\mathbf{T}_{W',s})\varphi_{W'}\big(s, \mathbf{T}_{W',s}, \pi_W^{-1}(c, T_{W',s}), H\big)$$

for $\overline{\mathbf{P}}_W \otimes \Lambda^*_W$ almost all arguments. Here, $\mathbf{P}_{W',s|\mathbf{T}_{W,s}}$ stands for a regular version of the conditional distribution of $\mathbf{T}_{W',s}$ under $\mathbf{P}_s$ given $\mathbf{T}_{W,s}$, and $c' = \pi_W^{-1}(c, T_{W',s})$ is the unique element of $T_{W',s}$ with $c' \cap W = c$. An analogous statement holds for $W' = \mathbb{R}^d$ when $\mathbf{P}$ admits a global division kernel with a Λ-density.

REMARK 4.5 (Conditional BRTs with finite horizon). Fix two windows $W, W' \in \mathbb{P}$ with $W \subset W'$ and let $\mathbf{P}_{W'}$ be a BRT in W' for a division kernel $\Phi_{W'}$. Furthermore, replace $\mathbb{R}^d$ by W' in Definition 2.12 and use the conditional division kernel

$$\Phi_W^{\mathrm{in}}(s, \mathbf{S}_s, c, \cdot|\mathbf{T}_{W'}) := \Phi_{W'}\big(s, \mathbf{S}_s \cup \mathbf{T}_{W,s}^{W',\mathrm{out}}, c, \cdot\big),$$

to obtain a conditional BRT $\mathbf{G}_W^{\Phi_{W'}}(\cdot|\mathbf{T}_W^{W',0,\mathrm{out}})$ in W; here we use the notation introduced in and after (4.7). The arguments in the proof of Theorem 3.1, (c) $\Rightarrow$ (b), then show that the kernel $\mathbf{G}_W^{\Phi_{W'}}$ is a regular version of the conditional distribution of $\boldsymbol{\pi}_W^{\mathrm{in}}$ for $\mathbf{P}_{W'}$ under the condition $\mathcal{B}_W^{W',0,\mathrm{out}}$.

4.3. *On the inner entropy density.* We first recall some standard properties of relative entropy. A basic fact is the variational formula, which states that

$$\mathcal{H}(\mu, \nu) = \sup_g \left[\int g \,\mathrm{d}\mu - \log \int e^g \,\mathrm{d}\nu\right], \tag{4.17}$$

for any two probability measures μ, ν on a common measurable space. Here, the supremum extends over all bounded measurable functions on this space; see [31], Theorem 4.1. On the one hand, the variational formula implies the useful estimate

$$\int g \,\mathrm{d}\mu \le \mathcal{H}(\mu, \nu) + \log \int e^g \,\mathrm{d}\nu \tag{4.18}$$

for any nonnegative measurable g. On the other hand, using Jensen's inequality it follows immediately that $\mathcal{H}(\mu, \nu)$ is jointly measure convex in both arguments simultaneously. Also, it is jointly lower semi-continuous in (μ, ν) in the topology generated by the integrals of bounded measurable functions. Finally, if μ and ν are restricted to a sub-σ-field $\mathcal{A}$ then relative entropy is increasing in $\mathcal{A}$. Alternative proofs of these facts can be found in [13], Section 15.1, for example. Since $\mathcal{H}(a\mu; b\nu) = a\mathcal{H}(\mu, \nu) + b\varrho(a/b)$ for $a, b > 0$ and normalised μ, ν [recall (3.5)], the last facts extend directly to the case of finite measures, except that the convexity then holds in the first argument only.

Now, turning to the proof of Theorem 3.5 we proceed with a series of lemmas. Let $\mathbf{P} \in \mathscr{P}_\Theta$ be arbitrarily given and $P = \mathbf{P} \circ \boldsymbol{\pi}_0^{-1}$ its initial distribution. We can clearly assume that

$$\liminf_{n\to\infty} n^{-d} \mathcal{H}_{[n]}^{\text{in}}(\mathbf{P}) < \infty \tag{4.19}$$

because otherwise there is nothing to show.

LEMMA 4.6. *Condition* (4.19) *implies condition* (LAC).

PROOF. We fix a window $W \in \mathbb{P}$ and consider the sets A_n defined in (4.10). Recall that for $\mathbf{T} \in A_n$ the inner window $\text{in}_{[n]}(s, \mathbf{T}_{[n]}^{\partial})$ contains W for all s, so that

$$\mathbf{G}_{[n]}^{\Lambda}\big(B|\mathbf{T}_{[n]}^{0,\partial}\big) = \boldsymbol{\Pi}_W^{\Lambda}(T_{W,0}, B)$$

for all $B \in \mathcal{B}_W$. Suppose now that $\mathcal{H}_{[n]}^{\text{in}}(\mathbf{P}) < \infty$ and recall the notation of Definition 3.4. Writing $\mathbf{P}_{[n]}^{\text{in}}(\cdot|\mathbf{T}_{[n]}^{0,\partial})$ for a regular conditional distribution of $\boldsymbol{\pi}_{[n]}^{\text{in}}$ under the condition $\boldsymbol{\pi}_{[n]}^{0,\partial} = \mathbf{T}_{[n]}^{0,\partial}$, we then have

$$\mathbf{P}_{[n]}^{\text{in}}\big(\cdot|\mathbf{T}_{[n]}^{0,\partial}\big) \ll \mathbf{G}_{[n]}^{\Lambda}\big(\cdot|\mathbf{T}_{[n]}^{0,\partial}\big) \qquad \text{for } \mathbf{P}\text{-almost all } \mathbf{T}.$$

Therefore, if $B \in \mathcal{B}_W$ is such that $P\boldsymbol{\Pi}^{\Lambda}(B) = 0$ then $\boldsymbol{\Pi}_W^{\Lambda}(T_{W,0}, B) = 0$ for almost all $\mathbf{T}$, and thus $\mathbf{P}_{[n]}^{\text{in}}(B|\mathbf{T}_{[n]}^{0,\partial}) = 0$ for almost all $\mathbf{T} \in A_n$. Hence,

$$\mathbf{P}(B \cap A_n) = \int_{A_n} \mathbf{P}(\mathrm{d}\mathbf{T})\mathbf{P}_{[n]}^{\text{in}}\big(B|\mathbf{T}_{[n]}^{0,\partial}\big) = 0.$$

Letting $n \to \infty$ through the integers n with $\mathcal{H}_{[n]}^{\text{in}}(\mathbf{P}) < \infty$, we thus obtain that $\mathbf{P}(B) = 0$. So, we have shown that $\mathbf{P} \ll P\boldsymbol{\Pi}^{\Lambda}$ on $\mathcal{B}_W$, and the proof is complete. □

Combining the preceding lemma with Theorem 3.3, we can conclude that $\mathbf{P}$ admits a global division kernel Φ. Hence, for each window $W \in \mathbb{P}$, the conditional distribution of $\boldsymbol{\pi}_W^{\text{in}}$ given $\mathcal{B}_W^{0,\partial}$ under $\mathbf{P}$, respectively, $P\boldsymbol{\Pi}^{\Lambda}$ are equal to the localised conditional BRTs $\mathbf{G}_W^{\Phi_W}(\cdot|\mathbf{T}_W^{0,\partial})$, respectively, $\mathbf{G}_W^{\Lambda}(\cdot|\mathbf{T}_W^{0,\partial})$ introduced in Remark 4.5, respectively, Example 2.13. It follows that

$$\mathcal{H}_W^{\text{in}}(\mathbf{P}) = \int \mathbf{P}(\mathrm{d}\mathbf{T})\mathcal{H}\big(\mathbf{G}_W^{\Phi_W}(\cdot|\mathbf{T}_W^{0,\partial}); \mathbf{G}_W^{\Lambda}(\cdot|\mathbf{T}_W^{0,\partial})\big). \tag{4.20}$$

This expression can be specified as follows.

LEMMA 4.7. *Under* (4.19), *we have for each* $W \in \mathbb{P}$

$$\mathcal{H}_W^{\mathrm{in}}(\mathbf{P}) = \int \mathrm{d}\overline{\mathbf{P}}_W(s, \mathbf{T}_W) \sum_{c \in T_{W,s}^{\mathrm{in}}} \mathcal{H}\big(\Phi_W(s, \mathbf{T}_{W,s}, c, \cdot); \mathbb{1}_{\langle c \rangle}\Lambda\big). \tag{4.21}$$

PROOF. By Lemma 4.6 and Corollary 4.2, Φ_W is absolutely continuous with respect to Λ; we write φ_W for the associated Radon–Nikodym density. Using the equivalence of Lemma 4.1(a) and (b) separately for the intervals between the "immigration times" (2.32), we obtain the following identity for the Radon–Nikodym density of $\mathbf{G}_W^{\Phi_W}(\cdot|\mathbf{T}_W^{0,\partial})$ relative to $\mathbf{G}_W^{\Lambda}(\cdot|\mathbf{T}_W^{0,\partial})$:

$$\begin{aligned}\log \frac{\mathrm{d}\mathbf{G}_W^{\Phi_W}(\cdot|\mathbf{T}_W^{0,\partial})}{\mathrm{d}\mathbf{G}_W^{\Lambda}(\cdot|\mathbf{T}_W^{0,\partial})}(\mathbf{T}_W^{\mathrm{in}}) &= [\hat{\lambda}_W^{\mathrm{in}}(0,1;\mathbf{T}_W^{\mathrm{in}}) - \hat{\phi}_W^{\mathrm{in}}(0,1;\mathbf{T}_W)] \\ &\quad + \sum_{(s,c,H) \in \mathscr{D}(\mathbf{T}_W^{\mathrm{in}})} \log \varphi_W(s, \mathbf{T}_{W,s-}, c, H),\end{aligned} \tag{4.22}$$

where $\mathbf{T}_W \in \mathbb{BT}_W$, $\mathscr{D}(\mathbf{T}_W^{\mathrm{in}})$ is the associated set of division events with $c \subset W$, and similarly to (4.1),

$$\hat{\phi}_W^{\mathrm{in}}(0,1;\mathbf{T}_W) = \int_0^1 \mathrm{d}s \sum_{c \in T_{W,s}^{\mathrm{in}}} \Phi_W\big(s, \mathbf{T}_{W,s}, c, \langle c \rangle\big)$$

and $\hat{\lambda}_W^{\mathrm{in}}(0,1;\mathbf{T}_W^{\mathrm{in}}) = \int_0^1 \mathrm{d}s \sum_{c \in T_{W,s}^{\mathrm{in}}} \Lambda(\langle c \rangle)$.

Now, $\mathcal{H}_W^{\mathrm{in}}(\mathbf{P})$ is simply the integral of (4.22) over

$$\mathbf{P}_W^{0,\partial}\big(\mathrm{d}\mathbf{T}_W^{0,\partial}\big)\mathbf{G}_W^{\Phi_W}\big(\mathrm{d}\mathbf{T}_W^{\mathrm{in}}|\mathbf{T}_W^{0,\partial}\big) = \mathbf{P}_W(\mathrm{d}\mathbf{T}_W).$$

By the equation in Lemma 4.1(c), integration of the last term in (4.22) gives the contribution

$$\int \mathrm{d}\overline{\mathbf{P}}_W(s, \mathbf{T}_{W,s}) \sum_{c \in T_{W,s}^{\mathrm{in}}} \int_{\langle c \rangle} \Phi_W(s, \mathbf{T}_{W,s}, c, \mathrm{d}H) \log \varphi_W(s, \mathbf{T}_{W,s}, c, H).$$

In terms of the measure $\overline{\mathbf{C}}^{\mathbf{P}_W,\mathrm{in}}$ which is defined by restricting the sum in (2.38) to the cells $c \in T_{W,s}^{\mathrm{in}}$, this can be rewritten in the concise form $\int \mathrm{d}\overline{\mathbf{C}}^{\mathbf{P}_W,\mathrm{in}} \otimes \Lambda^* \varphi_W \log \varphi_W$. Likewise, we have

$$\begin{aligned}\int \mathbf{P}_W(\mathrm{d}\mathbf{T}_W)&[\hat{\lambda}_W^{\mathrm{in}}(0,1;\mathbf{T}_W^{\mathrm{in}}) - \hat{\phi}_W^{\mathrm{in}}(0,1;\mathbf{T}_W)] \\ &= \int \mathrm{d}\overline{\mathbf{C}}^{\mathbf{P}_W,\mathrm{in}} \otimes \Lambda^*[1 - \varphi_W].\end{aligned}$$

Consequently, the $\mathbf{P}_W$-integral of (4.22) is equal to $\int \mathrm{d}\overline{\mathbf{C}}^{\mathbf{P}_W,\mathrm{in}} \otimes \Lambda^* \varrho(\varphi_W)$, and (4.21) follows by recalling (3.4). □

The final step in the proof of Theorem 3.5 is the following. Let $h^{\text{in}}(\mathbf{P})$ be defined by (3.8). For brevity, we write $h_n^{\text{in}}(\mathbf{P}) = n^{-d}\mathcal{H}_{[n]}^{\text{in}}(\mathbf{P})$.

LEMMA 4.8. *Under* (4.19),

$$\lim_{n\to\infty} h_n^{\text{in}}(\mathbf{P}) = \sup_{n\geq 1} h_n^{\text{in}}(\mathbf{P}) = h^{\text{in}}(\mathbf{P}).$$

PROOF. We claim first that $\liminf_{n\to\infty} h_n^{\text{in}}(\mathbf{P}) \geq h^{\text{in}}(\mathbf{P})$. We pick any $0 < \eta < 1$ and $\ell < \infty$ and restrict the sum in (4.21) for $W = [n]$ to the cells of T_s with midpoint in $[n\eta]$ and radius at most ℓ. More precisely, we let $L_n \subset \mathbb{R}^d$ be such that the set $\{[\eta] + x : x \in L_n\}$ is a tessellation of the cube $[n\eta]$. (Note that $|L_n| = n^d$.) Also, we take any $0 < \varepsilon < 1 - \eta$ and let n be so large that $\eta + \ell < \varepsilon n$. Then we can write

$$h_n^{\text{in}}(\mathbf{P}) \geq n^{-d} \sum_{x\in L_n} h_{n,\eta,\ell}(x)$$

with

$$h_{n,\eta,\ell}(x) := \int \mathrm{d}\overline{\mathbf{P}}(s, \mathbf{T}_s) \sum_{c\in T_s\cap\Gamma_{\eta,\ell}(x)} \mathcal{H}\big(\Phi_{[n]}(s, \mathbf{T}_{[n],s}, c, \cdot); \mathbb{1}_{\langle c\rangle}\Lambda\big),$$

where $\Gamma_{\eta,\ell}(x)$ is the set of all cells c satisfying $m(c) \in [\eta] + x$ and $r(c) \leq \ell$. Now, whether or not a cell $c \in T_{[n],s}$ belongs to $\Gamma_{\eta,\ell}(x)$ can be decided by looking at the restriction $T_{[\varepsilon n]+x,s}$. So, using Remark 4.4 and Jensen's inequality together with the translation invariance of $\mathbf{P}$ and the covariance equation (3.3) we find for each $x \in L_n$,

$$\begin{aligned}
h_{n,\eta,\ell}(x) &= \int \mathrm{d}\overline{\mathbf{P}}_{[\varepsilon n]+x}(s, \mathbf{T}_{[\varepsilon n]+x,s}) \\
&\qquad \times \sum_{c\in T_s\cap\Gamma_{\eta,\ell}(x)} \int \mathbf{P}_{[n],s|\mathbf{T}_{[\varepsilon n]+x,s}}(\mathrm{d}\mathbf{T}_{[n],s}) \\
&\qquad\qquad \times \int_{\langle c\rangle} \Lambda(\mathrm{d}H)\varrho\big(\varphi_{[n]}(s, \mathbf{T}_{[n],s}, c, H)\big) \\
&\geq \int \mathrm{d}\overline{\mathbf{P}}_{[\varepsilon n]}(s, \mathbf{T}_{[\varepsilon n],s}) \sum_{c\in T_s\cap\Gamma_{\eta,\ell}} \int_{\langle c\rangle} \Lambda(\mathrm{d}H)\varrho\big(\varphi_{[\varepsilon n]}(s, \mathbf{T}_{[\varepsilon n],s}, c, H)\big) \\
&= \mathcal{H}\big(\mathbb{1}_{\Gamma_{\eta,\ell}}\overline{\mathbf{C}}^{\mathbf{P}} \otimes \Phi|_{\mathcal{B}_{[\varepsilon n]}}; \mathbb{1}_{\Gamma_{\eta,\ell}}\overline{\mathbf{C}}^{\mathbf{P}} \otimes \Lambda^*|_{\mathcal{B}_{[\varepsilon n]}}\big).
\end{aligned}$$

In the last expression, $\mathcal{B}_{[\varepsilon n]}$ is identified with the σ-field that is generated by the projection $\bar{\boldsymbol{\pi}}_{[\varepsilon n]} \otimes \text{id}$, and $\Gamma_{\eta,\ell} := \Gamma_{\eta,\ell}(0)$ is viewed as a set in the product space $\overline{\mathbb{BT}} \times \mathbb{C} \times \mathbb{H}$. In the limit as $n \to \infty$, Perez' continuity theorem for relative entropies (cf. [13], Proposition 15.6) implies that the last relative entropy converges to

$$\begin{aligned}&\mathcal{H}\big(\mathbb{1}_{\Gamma_{\eta,\ell}}\overline{\mathbf{C}}^{\mathbf{P}}\otimes\Phi;\mathbb{1}_{\Gamma_{\eta,\ell}}\overline{\mathbf{C}}^{\mathbf{P}}\otimes\Lambda^*\big)\\&\qquad=\int \mathrm{d}\overline{\mathbf{P}}(s,\mathbf{T}_s)\sum_{c\in T_s\cap\Gamma_{\eta,\ell}}\mathcal{H}\big(\Phi(s,\mathbf{T}_s,c,\cdot);\mathbb{1}_{\langle c\rangle}\Lambda\big).\end{aligned}$$

By the time-integrated version of the Palm formula (2.35) and the shift covariance of Φ and Λ, the last integral is equal to

$$h_{\eta,\ell}:=\eta^d\int_{\{r(c)\le\ell\}}\mathrm{d}\overline{\mathbf{P}}^0(s,\mathbf{T}_s,c)\mathcal{H}\big(\Phi(s,\mathbf{T}_s,c,\cdot);\mathbb{1}_{\langle c\rangle}\Lambda\big).$$

Altogether, we find that $\liminf_{n\to\infty} h_n^{\mathrm{in}}(\mathbf{P})\ge h_{\eta,\ell}$, and the claim follows by letting $\eta\to 1$ and $\ell\to\infty$.

It remains to show that $h_n^{\mathrm{in}}(\mathbf{P})\le h^{\mathrm{in}}(\mathbf{P})$. By (4.19) and the above, $h^{\mathrm{in}}(\mathbf{P})<\infty$. This implies that the kernel Φ admits a Radon–Nikodym density relative to Λ. Applying Remark 4.4 and Jensen's inequality as above, we conclude from (4.21) that

$$h_n^{\mathrm{in}}(\mathbf{P})\le n^{-d}\int \mathrm{d}\overline{\mathbf{P}}(s,\mathbf{T}_s)\sum_{c\in T_s:\,c\subset[n]}\mathcal{H}\big(\Phi(s,\mathbf{T}_s,c,\cdot);\mathbb{1}_{\langle c\rangle}\Lambda\big).$$

The condition under the sum above implies that $m(c)\in[n]$. Using again (2.35) in its time-integrated version, we thus find that the last expression is not larger than $h^{\mathrm{in}}(\mathbf{P})$. The proof is thus complete. □

REMARK 4.9. The inner entropy of a BRT $\mathbf{P}$ in a window $W\in\mathbb{P}$ can be defined by considering the tessellations not only in W but also in some neighborhood of W. Namely, if Φ is a division kernel for $\mathbf{P}$ and $r>0$, one can introduce the quantity

$$\mathcal{H}_W^{r,\mathrm{in}}(\mathbf{P})=\int \mathbf{P}(\mathrm{d}\mathbf{T})\mathcal{H}\big(\mathbf{G}_W^{\Phi_{W+B_r}}(\cdot|\mathbf{T}_W^{W+B_r,0,\mathrm{out}});\mathbf{G}_W^{\Lambda}(\cdot|\mathbf{T}_W^{0,\partial})\big),$$

which is called the *inner entropy of* $\mathbf{P}$ *in* W *with horizon* r. Here, the first of the conditional BRTs $\mathbf{G}$ is as in Remark 4.5. A glance at the preceding proof then shows that Lemma 4.8 can be extended to yield

$$\lim_{n\to\infty} n^{-d}\mathcal{H}_{[n]}^{r,\mathrm{in}}(\mathbf{P})=h^{\mathrm{in}}(\mathbf{P}).$$

Next, we turn to the proof of Theorem 3.6, which is split into two lemmas.

LEMMA 4.10. *The inner entropy density* h^{in} *is affine.*

PROOF. As noticed after (4.18), relative entropy is a jointly convex function of probability measures. This shows that the inner entropies $\mathcal{H}_W^{\mathrm{in}}(\mathbf{P})=$

$\mathcal{H}(\mathbf{P}_W; \mathbf{P}_W^{0,\partial} \otimes \mathbf{G}_W^{\Lambda})$ are convex in $\mathbf{P}$, and so is their limit $h^{\text{in}}(\mathbf{P})$. The proof is therefore completed by showing that this limit is also concave. So, let $\mathbf{P}, \mathbf{P}' \in \mathscr{P}_\Theta$, $0 < a < 1$, $\widehat{\mathbf{P}} = a\mathbf{P} + (1-a)\mathbf{P}'$, and assume without loss of generality that $h^{\text{in}}(\widehat{\mathbf{P}}) < \infty$. By Lemma 4.8, it follows that $h_n^{\text{in}}(\widehat{\mathbf{P}}) < \infty$ for all n. In particular, $\widehat{\mathbf{P}}_{[n]} \ll \widehat{\mathbf{P}}_{[n]}^{0,\partial} \otimes \mathbf{G}_{[n]}^{\Lambda}$ with a Radon–Nikodym density g_n. The Radon–Nikodym theorem further implies that $\mathbf{P}_{[n]} \ll \widehat{\mathbf{P}}_{[n]}$ and $\mathbf{P}'_{[n]} \ll \widehat{\mathbf{P}}_{[n]}$ with densities f_n and f_n', respectively. It is clear that $af_n + (1-a)f_n' = 1$ almost surely for $\widehat{\mathbf{P}}_{[n]}$. Moreover, it follows that $\mathbf{P}_{[n]}^{0,\partial} = f_n^{0,\partial}\widehat{\mathbf{P}}_{[n]}^{0,\partial}$ for a suitable Radon–Nikodym density $f_n^{0,\partial}$. We conclude that $\mathbf{P}_{[n]} = (f_n g_n / f_n^{0,\partial})\mathbf{P}_{[n]}^{0,\partial} \otimes \mathbf{G}_{[n]}^{\Lambda}$. Since $f_n \le 1/a$ and $\int d\mathbf{P}_{[n]} \log f_n^{0,\partial} = \mathcal{H}(\mathbf{P}_{[n]}^{0,\partial}; \widehat{\mathbf{P}}_{[n]}^{0,\partial}) \ge 0$, this gives

$$n^d h_n^{\text{in}}(\mathbf{P}) = \int d\mathbf{P}_{[n]} \log \frac{f_n g_n}{f_n^{\partial}} \le \int d\mathbf{P}_{[n]} \log g_n + \log \frac{1}{a}.$$

Together with the analogous inequality for $\mathbf{P}'$, we finally end up with the estimate

$$a h_n^{\text{in}}(\mathbf{P}) + (1-a) h_n^{\text{in}}(\mathbf{P}') \le n^{-d} \int d\widehat{\mathbf{P}}_{[n]} \log g_n + o(1) = h_n^{\text{in}}(\widehat{\mathbf{P}}) + o(1).$$

The result thus follows from Lemma 4.8 by letting $n \to \infty$. $\square$

As for the topological properties of h^{in}, we note first that its lower semi-continuity is a direct consequence of Lemma 4.8 and the lower semi-continuity of relative entropy; recall the discussion below (4.18). Since $\mathbf{T} \mapsto |T_{[1],1}|$ is the supremum of bounded local functions, it is also evident that the hitting intensity $i_1(\cdot)$ is lower semi-continuous. It follows that the restricted level sets $\mathscr{P}_{\Theta,P,\beta,\gamma}$ (as introduced in Theorem 3.6) are closed. The following lemma, which can be viewed as a refinement of Lemma 4.6, will imply that they are in fact compact; as the intensity bound is not needed here, we put $\beta = \infty$.

LEMMA 4.11. *The restricted level sets $\mathscr{P}_{\Theta,P,\infty,\gamma}$ are locally equi-continuous in the following sense: for each $W \in \mathbb{P}$ and $0 \le \gamma < \infty$ and every sequence $B_k \in \mathcal{B}_W$ with $B_k \downarrow \varnothing$ as $k \to \infty$, one has*

$$\lim_{k\to\infty} \sup_{\mathbf{P}\in\mathscr{P}_{\Theta,P,\infty,\gamma}} \mathbf{P}(B_k) = 0.$$

PROOF. Let $W \in \mathbb{P}$ and a sequence $B_k \in \mathcal{B}_W$ with $B_k \downarrow \varnothing$ be given. Pick some $\varepsilon > 0$ and consider the events A_n defined in (4.10). Recall that $A_n \uparrow \mathbb{BT}$. Since $A_n \in \mathcal{B}_{[n],0}^{\partial}$, $\mathbf{P}(A_n)$ depends only on the initial distribution of $\mathbf{P}$, which is P for all $\mathbf{P} \in \mathscr{P}_{\Theta,P,\infty,\gamma}$. So, there is an n with $\mathbf{P}(A_n) \ge 1-\varepsilon$ for all $\mathbf{P} \in \mathscr{P}_{\Theta,P,\infty,\gamma}$.

Next, each $\mathbf{P} \in \mathscr{P}_{\Theta,P,\infty,\gamma}$ admits some division kernel Φ, and $\mathcal{H}_{[n]}^{\text{in}}(\mathbf{P}) \le \eta := n^d\gamma$ by Lemma 4.8. Since

$$\mathcal{H}_{[n]}^{\text{in}}(\mathbf{P}) = \int \mathbf{P}(d\mathbf{T}) \mathcal{H}\big(\mathbf{G}_{[n]}^{\Phi_{[n]}}(\cdot|\mathbf{T}_{[n]}^{0,\partial}); \mathbf{G}_{[n]}^{\Lambda}(\cdot|\mathbf{T}_{[n]}^{0,\partial})\big)$$

by definition and Remark 4.5, we can conclude that the set

$$H_n := \{\mathbf{T} \in \mathbb{BT} : \mathcal{H}(\mathbf{G}_{[n]}^{\Phi_{[n]}}(\cdot|\mathbf{T}_{[n]}^{0,\partial}); \mathbf{G}_{[n]}^{\Lambda}(\cdot|\mathbf{T}_{[n]}^{0,\partial})) \leq \eta/\varepsilon\}$$

in $\mathcal{B}_{[n]}^{0,\partial}$ has measure at least $1-\varepsilon$ for $\mathbf{P}$. It follows that

$$\mathbf{P}(B_k) \leq 2\varepsilon + \mathbf{P}(B_k \cap A_n \cap H_n) = 2\varepsilon + \int_{A_n \cap H_n} \mathbf{P}(\mathrm{d}\mathbf{T})\mathbf{G}_{[n]}^{\Phi_{[n]}}(B_k|\mathbf{T}_{[n]}^{0,\partial})$$

because $\mathrm{in}_{[n]}(s, \mathbf{T}_{[n]}^{\partial}) \supset W$ for all $\mathbf{T} \in A_n$ and all s; recall (2.31). The next step is to use the inequality (4.18). For $\mathbf{T} \in A_n \cap H_n$, this inequality shows that

$$(\eta/\varepsilon^2)\mathbf{G}_{[n]}^{\Phi_{[n]}}(B_k|\mathbf{T}_{[n]}^{\partial}) \leq (\eta/\varepsilon) + \log \int \mathrm{d}\mathbf{\Pi}_W^{\Lambda}(T_{W,0}, \cdot) \exp[(\eta/\varepsilon^2)\mathbb{1}_{B_k}]$$

since $\mathbf{G}_{[n]}^{\Lambda}(\cdot|\mathbf{T}_{[n]}^{\partial}) = \mathbf{\Pi}_W^{\Lambda}(T_{W,0}, \cdot)$ on $\mathcal{B}_W$ when $\mathbf{T} \in A_n$. Inserting this into the previous inequality, we find

$$\begin{aligned}\sup_{\mathbf{P} \in \mathscr{P}_{\Theta,P,\infty,\gamma}} & \mathbf{P}(B_k) \\ & \leq 3\varepsilon + (\varepsilon^2/\eta) \int P_W(\mathrm{d}T_W) \log \int \mathrm{d}\mathbf{\Pi}_W^{\Lambda}(T_W, \cdot) \exp[(\eta/\varepsilon^2)\mathbb{1}_{B_k}].\end{aligned}$$

Letting $k \to \infty$, using the dominated convergence theorem, and noting that ε was chosen arbitrarily, we arrive at the lemma. □

The preceding lemma verifies the conditions of Propositions 4.9 and 4.15 of [13], which imply that $\mathscr{P}_{\Theta,P,\infty,\gamma}$ is relatively compact and relatively sequentially compact within the class of all translation invariant BRTs. However, this does not yet imply that each limit of a net in $\mathscr{P}_{\Theta,P,\infty,\gamma}$ also satisfies the first-moment condition. (This is because h^{in} is the limit of *conditional* entropies which do not allow to control the number of cells that hit the boundary. But this number enters into the hitting intensity i_1.) The simplest way to deal with this problem is to add the bound $i_1 \leq \beta$ which trivially implies (2.33) also for all limiting BRTs. The proof of Theorem 3.6 is therefore complete.

4.4. *Free energy density*, *variational principle*, *existence*. Throughout this section, we fix a moderate division kernel Ψ. Our first item is the existence of the energy density.

PROOF OF THEOREM 3.8. Let $\mathbf{P} \in \mathscr{P}_\Theta$ be a BRT with a covariant division kernel Φ. By property (c) of Theorem 3.1, the inner energy of $\mathbf{P}$ in a window $W \in \mathbb{P}$ can be written in the form

$$\begin{aligned}&\mathcal{U}_W^{\mathrm{in}}(\mathbf{P}; \Psi) \\ &\quad = \int \mathrm{d}\overline{\mathbf{P}}(s, \mathbf{T}_s) \sum_{c \in T_s : c \subset W} \int \Phi(s, \mathbf{T}_s, c, \mathrm{d}H) \log \psi(s, \mathbf{T}_s, c, H).\end{aligned} \tag{4.23}$$

Since Φ and ψ are covariant, the time-integrated version of the Palm formula (2.35) shows that the last term can be written in the form

$$\int \mathrm{d}\overline{\mathbf{P}}^0(s, \mathbf{T}_s, c)\,\mathrm{vol}(x : c + x \subset W) \int \Phi(s, \mathbf{T}_s, c, \mathrm{d}H) \log \psi(s, \mathbf{T}_s, c, H).$$

Hence,

$$n^{-d}\mathcal{U}^{\mathrm{in}}_{[n]}(\mathbf{P}; \Psi) = u^{\mathrm{in}}(\mathbf{P}; \Psi) + \delta_n(\mathbf{P}; \Psi)$$

with

$$\left|\delta_n(\mathbf{P}; \Psi)\right| \le \kappa_\Psi \int \mathrm{d}\overline{\mathbf{P}}^0(s, \mathbf{T}_s, c)\,\mathrm{vol}\bigl(x \in [1] : c/n + x \not\subset [1]\bigr)\Phi\bigl(s, \mathbf{T}_s, c, \langle c\rangle\bigr)$$

by (M3). The volume term above is bounded by 1 and tends to 0 as $n \to \infty$. To apply the dominated convergence theorem, we thus need to show that the total mass of $\overline{\mathbf{P}}^0 \otimes \Phi$ is finite. But the Palm formula and (4.14) show that this mass is at most $i_1(\mathbf{P})$. This completes the proof of the first part of Theorem 3.8 and implies the bound on $|u^{\mathrm{in}}(\mathbf{P}; \Psi)|$.

The proof of the second part is similar: the Palm formula gives

$$\begin{aligned}&\mathcal{V}^{\mathrm{in}}_W(\mathbf{P}; \Psi)\\ &\quad = \int \mathrm{d}\overline{\mathbf{P}}^0(s, \mathbf{T}_s, c)\,\mathrm{vol}(x : c + x \subset W) \int_{\langle c\rangle} \Lambda(\mathrm{d}H)\bigl(\psi(s, \mathbf{T}_s, c, H) - 1\bigr)\end{aligned}$$

and thus $n^{-d}\mathcal{V}^{\mathrm{in}}_{[n]}(\mathbf{P}; \Psi) = v^{\mathrm{in}}(\mathbf{P}; \Psi) + \delta'_n(\mathbf{P}; \Psi)$ with a remainder term δ'_n which, by assumption (M4), is bounded in modulus by κ'_Ψ times

$$\int \mathrm{d}\overline{\mathbf{P}}^0(s, \mathbf{T}_s, c)\,\mathrm{vol}\bigl(x \in [1] : c/n + x \not\subset [1]\bigr).$$

By (2.39) and the dominated convergence theorem, this bound vanishes in the limit $n \to \infty$. The proof of Theorem 3.8 is therefore complete. □

REMARK 4.12. Exploiting Theorem 3.1(c) and the Palm formula (2.35) in the same way as in the first part of the preceding proof, one finds that the energy density can be written in the alternative form

$$u^{\mathrm{in}}(\mathbf{P}; \Psi) = \int \mathbf{P}(\mathrm{d}\mathbf{T}) \sum_{(s,c,H)\in\mathscr{D}(\mathbf{T}) : m(c)\in[1]} \log \psi(s, \mathbf{T}_{s-}, c, H),$$

in which the division kernel of $\mathbf{P}$ does not appear. In particular, it follows that $u^{\mathrm{in}}(\cdot; \Psi)$ is affine.

Turning to the proof of the variational principle, Theorem 3.9, we introduce an *inner relative entropy of a BRT* $\mathbf{P}$ *in a window* $W \in \mathbb{P}$ *with horizon* $r = r_\Psi$ *relative* *to* Ψ as follows: if $\mathbf{P}$ admits a division kernel Φ we set, using the notation of Remark 4.5,

$$\begin{aligned}\mathcal{H}_W^{r,\mathrm{in}}(\mathbf{P};\Psi) \\ = \int \mathbf{P}(\mathrm{d}\mathbf{T})\mathcal{H}\big(\mathbf{G}_W^{\Phi_{W+B_r}}(\cdot|\mathbf{T}_W^{W+B_r,0,\mathrm{out}});\mathbf{G}_W^{\Psi_{W+B_r}}(\cdot|\mathbf{T}_W^{W+B_r,0,\mathrm{out}})\big);\end{aligned} \tag{4.24}$$

otherwise we set $\mathcal{H}_W^{r,\mathrm{in}}(\mathbf{P};\Psi)=\infty$. (Compare this definition with Remark 4.9, where $\Psi=\Lambda^*$.) By the bounded-range property (M2) of Ψ and Corollary 3.2, the conditional BRT $\mathbf{G}_W^{\Psi_{W+B_r}}$ in (4.24) actually coincides with $\mathbf{G}_W^{\Psi}$. We then have the following convergence to the quantity $h^{\mathrm{in}}(\mathbf{P};\Psi)$ in (3.11).

COROLLARY 4.13. *Let* Ψ *be a moderate division kernel and* $r=r_\Psi$ *its range. Then*

$$h^{\mathrm{in}}(\mathbf{P};\Psi)=\lim_{n\to\infty} n^{-d}\mathcal{H}_{[n]}^{r,\mathrm{in}}(\mathbf{P};\Psi)$$

for all $\mathbf{P}\in\mathscr{P}_\Theta$. *The limit is finite if and only if* $h^{\mathrm{in}}(\mathbf{P})<\infty$, *and then equation* (3.12) *holds.*

PROOF. An analog of equation (4.22) gives for each n the identity

$$\mathcal{H}_{[n]}^{r,\mathrm{in}}(\mathbf{P};\Psi)=\mathcal{H}_{[n]}^{r,\mathrm{in}}(\mathbf{P})-\mathcal{U}_{[n]}^{\mathrm{in}}(\mathbf{P};\Psi)+\mathcal{V}_{[n]}^{\mathrm{in}}(\mathbf{P};\Psi), \tag{4.25}$$

which is a counterpart to (3.11). Also, the estimates in the proof of Theorem 3.8 show that the second and third term on the right-hand side are bounded in modulus by a finite constant times n^d. The convergence result thus follows directly from Remark 4.9 and Theorem 3.8 (together with Lemma 4.6 and Theorem 3.3).

Next, suppose that $h^{\mathrm{in}}(\mathbf{P})<\infty$ and let φ and ψ be the Radon–Nikodym densities of Φ and Ψ with respect to Λ^*. Inserting the explicit expressions for all quantities, we then obtain

$$\begin{aligned}&h^{\mathrm{in}}(\mathbf{P})-u^{\mathrm{in}}(\mathbf{P};\Psi)+v^{\mathrm{in}}(\mathbf{P};\Psi)\\ &\quad=\int \mathrm{d}\overline{\mathbf{P}}^0\otimes\Lambda^*[-\varphi+\varphi\log\varphi-\varphi\log\psi+\psi]\\ &\quad=\int \mathrm{d}\overline{\mathbf{P}}^0\otimes\Lambda^*\psi\varrho(\varphi/\psi)=\int \mathrm{d}\overline{\mathbf{P}}^0\,\mathcal{H}(\Phi;\Psi),\end{aligned}$$

which is (3.12). □

The variational principle, Theorem 3.9, follows directly from equation (3.12) and thus from the preceding corollary.

Next we address the existence problem for BRTs with given division kernel, as stated in Theorem 3.10. We still keep a moderate Ψ fixed and let $r=r_\Psi$ be its range. We also fix an initial distribution $P\in\mathscr{P}_\Theta(\mathbb{T})$. We will construct a BRT

$\mathbf{P}$ with initial distribution $\boldsymbol{\pi}_0(\mathbf{P}) = P$ and division kernel Ψ as a cluster point of some approximating measures $\mathbf{P}^{n,\mathrm{av}}$.

Specifically, for any n we let $\bar{n} = n + r$ and consider the shifted cubes $[n]_i = [n] + \bar{n}i$, $i \in \mathbb{Z}^d$, which are separated by a grid of corridors of width r. Let $[n]_\bullet = \bigcup_{i \in \mathbb{Z}^d} [n]_i$ be their union. We introduce a BRT $\mathbf{P}^n$ for which the cells that hit the corridors between the boxes $[n]_i$ evolve according to Λ^* and, conditioned on this STIT evolution, the cells inside these boxes evolve independently according to Ψ. (This is inspired by the familiar construction of independent repetitions in disjoint blocks, which is often used in large deviation theory; see [13], (15.52), for example. Using the STIT process in the corridors between the blocks, we avoid an artificial cutting of cells at the block boundaries.) Formally, we introduce the projection

$$\boldsymbol{\pi}^{0,\mathrm{out}}_{[n]_\bullet} : \mathbf{T} \mapsto \mathbf{T}^{0,\mathrm{out}}_{[n]_\bullet} := \Bigl(\bigcup_{i \in \mathbb{Z}^d} T^{\mathrm{in}}_{[n]_i,0}, \bigcap_{i \in \mathbb{Z}^d} \mathbf{T}^{\mathrm{out}}_{[n]_i} \Bigr),$$

and define

$$\mathbf{P}^n = \bigl(P\boldsymbol{\Pi}^\Lambda\bigr)^{0,\mathrm{out}}_{[n]_\bullet} \otimes \bigotimes_{i \in \mathbb{Z}^d} \mathbf{G}^\Psi_{[n]_i}. \tag{4.26}$$

More explicitly, $\mathbf{P}^n$ is defined by its integrals

$$\int f \, \mathrm{d}\mathbf{P}^n = \int P\boldsymbol{\Pi}^\Lambda(\mathrm{d}\mathbf{T}) \prod_{i \in \mathbb{Z}^d} \int \mathbf{G}^\Psi_{[n]_i}\bigl(\mathrm{d}\mathbf{S}_i \,|\, \mathbf{T}^{0,\mathrm{out}}_{[n]_i}\bigr) f\Bigl(\mathbf{T}^{0,\mathrm{out}}_{[n]_\bullet} \cup \bigcup_i \mathbf{S}_i\Bigr)$$

for measurable functions $f \geq 0$ on $\mathbb{BT}$. By the bounded-range property (M2), the conditional BRTs $\mathbf{G}^\Psi_{[n]_i}(\cdot|\mathbf{T}^{0,\mathrm{out}}_{[n]_i})$ depend only on $\mathbf{T}^{0,\mathrm{out}}_{[n]_\bullet}$, so that $\mathbf{P}^n$ is well defined. It is easily seen that $\mathbf{P}^n$ is a BRT with initial distribution P and division kernel

$$\Psi^n(s, \mathbf{T}, c, \cdot) = \begin{cases} \Psi(s, \mathbf{T}, c, \cdot), & \text{if } c \subset [n]_i \text{ for some } i \in \mathbb{Z}^d, \\ \Lambda\bigl(\langle c\rangle \cap \cdot\bigr), & \text{otherwise.} \end{cases} \tag{4.27}$$

To achieve translation invariance, we introduce the average

$$\mathbf{P}^{n,\mathrm{av}} = \bar{n}^{-d} \int_{[\bar{n}]} \mathrm{d}x \, \mathbf{P}^n \circ \vartheta_x^{-1}. \tag{4.28}$$

The next two lemmas show that the BRTs $\mathbf{P}^{n,\mathrm{av}}$ belong to a restricted level set of the inner entropy density, and thus have a cluster point.

LEMMA 4.14. (a) *There exists a constant* $\beta < \infty$ *such that* $i_1(\mathbf{P}^{n,\mathrm{av}}) \leq \beta$ *for all* n.

(b) *For every* $\varepsilon > 0$, *there exists some* $\tau < \infty$ *such that*

$$\int \mathbf{P}^{n,\mathrm{av}}(\mathrm{d}\mathbf{T}) |T_{[1],1}| \mathbb{1}_{\{|T_{[1],1}| \geq \tau\}} \leq \varepsilon \qquad \textit{for all } n.$$

PROOF. Let $\kappa = \kappa_\Psi$. Since $\psi \le e^\kappa$ by (M3), it follows that each kernel Ψ^n also has a Λ-density ψ^n satisfying $\psi^n \le e^\kappa$ for all n. With the help of Remark 4.4, we can further conclude that this bound remains true after localisation to a window $W \in \mathbb{P}$ (relative to $\mathbf{P}^n$); that is, the localised kernel Ψ^n_W has a Λ-density ψ^n_W with $\psi^n_W \le e^\kappa$. In particular, if $W = [1] + x$ is a translate of the unit cube, then

$$\Psi^n_W\big(s, \mathbf{T}_{W,s}, c, \langle c\rangle\big) \le e^\kappa \Lambda\big(\langle\!\langle [1] \rangle\!\rangle\big) =: \alpha < \infty \tag{4.29}$$

for all possible arguments. In view of Lemma 4.3, it follows that

$$\int \mathbf{P}^n(\mathrm{d}\mathbf{T})|T_{[1]+x,1}| \le \beta := i_0(P)e^\alpha, \tag{4.30}$$

and statement (a) follows by averaging over x.

To prove (b), we still let $W = [1] + x$ and define $\varepsilon_1 = \varepsilon/4i_0(P)$. By Lemma 4.3, there exists a number τ_1 with

$$\sup_n \int \mathbf{P}^n(\mathrm{d}\mathbf{T})\big(|T_{W,1}| - \tau_1|T_{W,0}|\big)_+ \le \varepsilon_1 i_0(P) = \varepsilon/4.$$

For any τ_2 we then find (by distinguishing whether or not $\tau_1|T_{W,0}| \le \tau_2$) that

$$\int \mathbf{P}^n(\mathrm{d}\mathbf{T})\big(|T_{W,1}| - \tau_2\big)_+ \le \varepsilon/4 + e^\alpha \int P(\mathrm{d}T)|T_{[1]}|\mathbb{1}_{\{|T_{[1]}| > \tau_2/\tau_1\}},$$

which is at most $\varepsilon/2$ for suitable choice of τ_2. Setting $\tau = 2\tau_2$ and using that

$$|T_{W,1}| \le 2\big(|T_{W,1}| - \tau_2\big)_+ \qquad \text{on } \big\{|T_{W,1}| \ge \tau\big\},$$

we then see that

$$\int \mathbf{P}^n(\mathrm{d}\mathbf{T})|T_{[1]+x,1}|\mathbb{1}_{\{|T_{[1]+x,1}| \ge \tau\}} \le \varepsilon$$

for all n. Statement (b) thus follows by taking the average over $x \in [\bar n]$. □

LEMMA 4.15. *$h^{\mathrm{in}}(\mathbf{P}^{n,\mathrm{av}}; \Psi) \to 0$ as $n \to \infty$.*

PROOF. Fix any n and let $\mathbf{P}^n$ and Ψ^n be given by (4.26) and (4.27). Consider the inner relative entropy of $\mathbf{P}^{n,\mathrm{av}}$ relative to Ψ in a large cube $W = [k]$ and with horizon $r = r_\Psi$, as defined in (4.24). In concise notation, (4.24) reads

$$\mathcal{H}^{r,\mathrm{in}}_W(\mathbf{P}; \Psi) = \mathcal{H}\big(\mathbf{P}_{W+B_r}; \mathbf{P}^{W+B_r,0,\mathrm{out}}_W \otimes \mathbf{G}^\Psi_W\big).$$

As relative entropy is jointly measure convex, we have

$$\mathcal{H}^{r,\mathrm{in}}_W\big(\mathbf{P}^{n,\mathrm{av}}; \Psi\big) \le \bar n^{-d} \int_{[\bar n]} \mathrm{d}x\, \mathcal{H}^{r,\mathrm{in}}_{W+x}\big(\mathbf{P}^n; \Psi\big).$$

To estimate this further, we note that $\mathbf{P}^n$ has the division kernel Ψ^n. A combination of (4.25), (4.23), (3.10) and an analog of (4.21) thus gives the formula

$$\mathcal{H}^{r,\mathrm{in}}_{W+x}\big(\mathbf{P}^n; \Psi\big) = \int \mathrm{d}\overline{\mathbf{P}}^n(s, \mathbf{T}_s) \sum_{c \in T_s : c \subset W+x} \mathcal{H}\big(\Psi^n; \Psi | s, \mathbf{T}_s, c\big),$$

where $\mathcal{H}(\Psi^n; \Psi|s, \mathbf{T}_s, c) = \mathcal{H}(\Psi^n(s, \mathbf{T}_s, c, \cdot); \Psi(s, \mathbf{T}_s, c, \cdot))$ for brevity. We can further use that $W + x \subset [k + \bar{n}]$ when $x \in [\bar{n}]$. Altogether, we obtain

$$\mathcal{H}_W^{r,\text{in}}(\mathbf{P}^{n,\text{av}}; \Psi) \le \int d\overline{\mathbf{P}}^n(s, \mathbf{T}_s) \sum_{c \in T_s : c \subset [k+\bar{n}]} \mathcal{H}(\Psi^n; \Psi|s, \mathbf{T}_s, c).$$

Next, it is clear from (4.27) that $\mathcal{H}(\Psi^n; \Psi|\cdot, \cdot, c) = 0$ when $c \subset [n]_i$ for some i. On the other hand, for any cell c hitting the corridors between the boxes $[n]_i$ we have $\mathcal{H}(\Psi^n; \Psi|\cdot, \cdot, c) = \mathcal{H}(\Lambda^*; \Psi|\cdot, \cdot, c)$, which is bounded by a constant. Indeed, the function $\varrho(a)$ defined in (3.5) is bounded by a multiple of $|a - 1|$ as long as $a \le e^{\kappa_\Psi}$. Assumptions (M3) and (M4) therefore imply that

$$\mathcal{H}(\Lambda^*; \Psi|\cdot, \cdot, c) = \int_{\langle c \rangle} \Lambda(dH) \psi(\cdot, \cdot, c, H) \varrho\big(1/\psi(\cdot, \cdot, c, H)\big) \le \tilde{\kappa}_\Psi$$

for some constant $\tilde{\kappa}_\Psi < \infty$ and all c hitting the corridors.

Now let $k = (\ell - 1)\bar{n}$ for some integer ℓ and L_ℓ be such that $\{[1] + x : x \in L_\ell\}$ is a tessellation of $[\ell\bar{n}] \setminus [n]_\bullet$. The preceding estimates then show that

$$\begin{aligned}\mathcal{H}_{[k]}^{r,\text{in}}(\mathbf{P}^{n,\text{av}}; \Psi) &\le \sum_{x \in L_\ell} \int d\overline{\mathbf{P}}^n(s, \mathbf{T}_s) \sum_{c \in T_s : c \cap ([1]+x) \neq \varnothing} \mathcal{H}(\Psi^n; \Psi|s, \mathbf{T}_s, c) \\ &\le \tilde{\kappa}_\Psi \sum_{x \in L_\ell} \int d\overline{\mathbf{P}}^n(s, \mathbf{T}_s) |T_{[1]+x,1}| \le \tilde{\kappa}_\Psi \beta |L_\ell|.\end{aligned}$$

The last inequality comes from (4.30). Letting $\ell \to \infty$ and applying Corollary 4.13, we finally see that

$$h^{\text{in}}(\mathbf{P}^{n,\text{av}}; \Psi) \le \tilde{\kappa}_\Psi \beta \lim_{\ell \to \infty} \frac{(\ell\bar{n})^d - \ell^d n^d}{((\ell - 1)\bar{n})^d} = \tilde{\kappa}_\Psi \beta \big(1 - (n/\bar{n})^d\big).$$

This proves the lemma. □

Combining equation (3.11) with the last lemma and the bounds in Theorem 3.8 and Lemma 4.14(a), one finds that

$$h^{\text{in}}(\mathbf{P}^{n,\text{av}}) \le (\kappa_\Psi + \kappa'_\Psi)\beta + 1 =: \gamma < \infty$$

when n is large enough. That is, the measures $\mathbf{P}^{n,\text{av}}$ eventually belong to the sequentially compact level set $\mathscr{P}_{\Theta, P, \beta, \gamma}$ of Theorem 3.6. This means that a subsequence converges in τ_{loc} to some $\mathbf{P}$ in this set. We need to show that $\mathbf{P}$ has the division kernel Ψ. In view of Theorem 3.9, this will follow once we have shown that $h^{\text{in}}(\mathbf{P}; \Psi) = 0$. By the last lemma, it is therefore sufficient to verify that $h^{\text{in}}(\cdot; \Psi)$ is lower semi-continuous on the closure of the sequence $\{\mathbf{P}^{n,\text{av}} : n \ge 1\}$. In view of equation (3.11) and Theorem 3.6, this follows from the next lemma, which completes the proof of Theorem 3.10.

LEMMA 4.16. *The functionals $u^{\text{in}}(\cdot;\Psi)$ and $v^{\text{in}}(\cdot;\Psi)$ are continuous on the closure $\mathscr{C}$ of the sequence $\{\mathbf{P}^{n,\text{av}}: n\geq 1\}$.*

PROOF. First, we observe that the estimate in Lemma 4.14(b) holds not only for all $\mathbf{P}^{n,\text{av}}$, but even for all $\mathbf{P}\in\mathscr{C}$. This is because the integral there is a lower semi-continuous function of the integrating measure. We further know from Lemma 4.6 that each $\mathbf{P}\in\mathscr{C}$ satisfies (LAC). Hence, Theorem 3.8 and Remark 4.12 can be applied.

It follows that $u^{\text{in}}(\mathbf{P};\Psi)=\int u\,\mathrm{d}\mathbf{P}$ for the function

$$u(\mathbf{T})=\sum_{(s,c,H)\in\mathscr{D}(\mathbf{T}):\,m(c)\in[1]}\log\psi(s,\mathbf{T}_{s-},c,H)$$

on $\mathbb{BT}$, which in general is neither bounded nor local. We will therefore replace u by a truncated version

$$u_{\tau,\ell}(\mathbf{T})=\mathbb{1}_{\{|\mathbf{T}_{[1],1}|\leq\tau\}}\sum_{(s,c,H)\in\mathscr{D}(\mathbf{T}):\,m(c)\in[1],r(c)\leq\ell}\log\psi(s,\mathbf{T}_{s-},c,H),$$

for suitable numbers τ and ℓ; $r(c)$ is again the radius of c. The function $u_{\tau,\ell}$ is bounded in modulus by $\kappa_\Psi\tau$ and also local because of (M2). It differs from u by at most $\kappa_\Psi(\delta_\tau+\delta_\ell)$ with the error functions

$$\delta_\tau(\mathbf{T})=\mathbb{1}_{\{|\mathbf{T}_{[1],1}|>\tau\}}|\mathbf{T}_{[1],1}|,\qquad \delta_\ell(\mathbf{T})=\sum_{(s,c,H)\in\mathscr{D}(\mathbf{T}):\,m(c)\in[1]}\mathbb{1}_{\{r(c)>\ell\}}.$$

As noticed at the beginning of this proof, we have $\sup_{\mathbf{P}\in\mathscr{C}}\int\delta_\tau\,\mathrm{d}\mathbf{P}\to 0$ as $\tau\to\infty$. On the other hand, the function δ_ℓ is not larger than

$$\delta'_\ell(\mathbf{T})=\sum_{c_0\in T_0:\,c_0\cap[1]\neq\varnothing,r(c_0)>\ell}|T_{c_0\cap[1],1}|,$$

and Lemma 4.3 gives the estimate

$$\sup_{\mathbf{P}\in\mathscr{C}}\int\delta'_\ell\,\mathrm{d}\mathbf{P}\leq e^{\alpha}\int P(\mathrm{d}T)\sum_{c_0\in T_0:\,c_0\cap[1]\neq\varnothing}\mathbb{1}_{\{r(c_0)>\ell\}}$$

for the constant α in (4.29) because each $\mathbf{P}\in\mathscr{C}$ has initial distribution P. This bound does not depend on n and tends to 0 as $\ell\to\infty$ because $i_0(P)<\infty$. We have thus shown that the restriction of $u^{\text{in}}(\cdot;\Psi)$ to $\mathscr{C}$ is the uniform limit of the functions $\mathbf{P}\mapsto\int u_{\tau,\ell}\,\mathrm{d}\mathbf{P}$, which are continuous in τ_{loc}.

The analogous result for $v^{\text{in}}(\mathbf{P}^{n,\text{av}};\Psi)$ is achieved in a similar way by truncating the function

$$v(\mathbf{T})=\int_0^1\mathrm{d}s\sum_{c\in T_s:\,m(c)\in[1]}\int_{\langle c\rangle}\Lambda(\mathrm{d}H)\big(\psi(s,\mathbf{T}_s,c,H)-1\big)$$

and using (M4). $\square$

As the proof of Theorem 3.10 is now complete, we turn to its corollary.

PROOF OF COROLLARY 3.11. Suppose $\mathbf{P} \in \mathscr{G}_\Theta(\Psi)$ is not extremal in $\mathscr{P}_\Theta$. Then $\mathbf{P} = a\mathbf{P}^1 + (1-a)\mathbf{P}^2$ for some $0 < a < 1$ and two distinct BRTs $\mathbf{P}^1, \mathbf{P}^2 \in \mathscr{P}_\Theta$. By Theorem 3.6, Remark 4.12 and Theorem 3.9, it follows that

$$0 = h^{\text{in}}(\mathbf{P}; \Psi) = ah^{\text{in}}(\mathbf{P}^1; \Psi) + (1-a)h^{\text{in}}(\mathbf{P}^2; \Psi),$$

so that $\mathbf{P}^1, \mathbf{P}^2$ both belong to $\mathscr{G}_\Theta(\Psi)$. Hence, $\mathbf{P}$ is not extremal in $\mathscr{G}_\Theta(\Psi)$. $\square$

Our final observation concerns the uniqueness problem discussed in Remark 3.12. We will exploit the fact that, in one space dimension, we always have that $\sum_{c \in T_W} \Lambda(\langle c\rangle) = \Lambda(\langle W\rangle)$ when $W \in \mathbb{P}$ and $T_W \in \mathbb{T}_W$. Consider the following variants of conditions (M3) and (M4):

(M3′) Ψ is *STIT-bounded*, in that $\Psi \le K_\Psi \Lambda^*$ for some constant $K_\Psi < \infty$.

(M4′) Ψ is *STIT for large cells*, in that $\Psi(\cdot, \cdot, c, \cdot) = \Lambda(\langle c\rangle \cap \cdot)$ whenever $\operatorname{diam}(c) \ge r'_\Psi$ for some constant $r'_\Psi < \infty$.

PROPOSITION 4.17. *Suppose that the space dimension is $d = 1$. Let $P \in \mathscr{P}(\mathbb{T})$ and Ψ be a division kernel satisfying* (M2), (M3′) *and* (M4′). *Then there exists at most one BRT for Ψ with initial distribution P.*

PROOF. Suppose there exist two distinct BRTs $\mathbf{P}$, $\mathbf{P}'$ for Ψ with the same initial distribution P. Consider the difference measure $\mathbf{P}^\delta = \mathbf{P} - \mathbf{P}'$ and fix an interval $[k] \in \mathbb{P}$. Let g be $\mathcal{B}_{[k]}$-measurable with $|g| \le 1$. Using property (a) of Theorem 3.1, we obtain for each $0 < t \le 1$ the identity

$$\int g\,\mathrm{d}\mathbf{P}_t^\delta = \int_0^t \mathrm{d}s \int \mathbf{P}_s^\delta(\mathrm{d}\mathbf{T}_s)\mathbb{L}_s^\Psi g(\mathbf{T}_s)$$

with

$$\mathbb{L}_s^\Psi g(\mathbf{T}_s) = \sum_{c \in T_s : c \cap [k] \ne \varnothing} \int_{\langle c \cap [k]\rangle} \Psi(s, \mathbf{T}_s, c, \mathrm{d}H)\big[g\big(\oslash_{s,c,H}(\mathbf{T}_s)\big) - g(\mathbf{T}_s)\big].$$

Now, (M2) and (M4′) imply that $\mathbb{L}_s^\Psi g(\mathbf{T}_s)$ depends only on $\mathbf{T}_{[k+r],s}$ with $r = 2(r_\Psi + r'_\Psi)$. On the other hand, using (M3′) and the additivity of $c \mapsto \Lambda(\langle c\rangle)$ we find that

$$\big|\mathbb{L}_s^\Psi g(\mathbf{T}_s)\big| \le 2K_\Psi \sum_{c \ni T_s : c \cap [k] \ne \varnothing} \Lambda\big(\langle c \cap [k]\rangle\big) = 2K_\Psi \Lambda\big(\langle [k]\rangle\big) =: \alpha k.$$

The total variation norm $\delta_k(t) := \|\mathbf{P}_{[k],t}^\delta\|$ thus satisfies the inequality

$$\delta_k(t) \le \alpha k \int_0^t \delta_{k+r}(s)\,\mathrm{d}s \tag{4.31}$$

of Gronwall type. (Note that δ_k is increasing and, therefore, measurable.) Since $\delta_{k+nr}(s) \le 2$, we obtain by n-fold iteration

$$\delta_k(t) \le 2\alpha^n (k+nr)^n t^n/n! \le 2e^k \big(\alpha t e^r\big)^n$$

and thus, in the limit as $n \to \infty$, $\delta_k(t) = 0$ for all $t < \varepsilon := 1/(\alpha e^r)$ and all k. Inserting this into (4.31) and repeating the estimate, we obtain that $\delta_k(t) = 0$ for all $t < 2\varepsilon$ and all k. Continuing in this way, we finally find that $\delta_k(1) = 0$ for all k, which means that $\mathbf{P} = \mathbf{P}'$. $\square$

Acknowledgements. We would like to thank Claudia Redenbach for providing the simulation pictures of Figure 2, and an anonymous referee for helpful hints and suggestions.

REFERENCES

[1] ARAK, T. and SURGAILIS, D. (1989). Markov fields with polygonal realizations. *Probab. Theory Related Fields* **80** 543–579. MR0980687

[2] ARAK, T. and SURGAILIS, D. (1991). Consistent polygonal fields. *Probab. Theory Related Fields* **89** 319–346. MR1113222

[3] BERTIN, E., BILLIOT, J.-M. and DROUILHET, R. (1999). Existence of Delaunay pairwise Gibbs point process with superstable component. *J. Stat. Phys.* **95** 719–744. MR1700922

[4] BERTOIN, J. (2006). *Random Fragmentation and Coagulation Processes*. Cambridge Univ. Press, Cambridge. MR2253162

[5] CATTIAUX, P., ROELLY, S. and ZESSIN, H. (1996). Une approche gibbsienne des diffusions browniennes infini-dimensionnelles. *Probab. Theory Related Fields* **104** 147–179. MR1373374

[6] DAI PRA, P. (1993). Large deviations and stationary measures for interacting particle systems. *Stochastic Process. Appl.* **48** 9–30. MR1237166

[7] DAI PRA, P., ROELLY, S. and ZESSIN, H. (2002). A Gibbs variational principle in space-time for infinite-dimensional diffusions. *Probab. Theory Related Fields* **122** 289–315. MR1894070

[8] DEREUDRE, D., DROUILHET, R. and GEORGII, H.-O. (2012). Existence of Gibbsian point processes with geometry-dependent interactions. *Probab. Theory Related Fields* **153** 643–670. MR2948688

[9] DEREUDRE, D. and GEORGII, H.-O. (2009). Variational characterisation of Gibbs measures with Delaunay triangle interaction. *Electron. J. Probab.* **14** 2438–2462. MR2563247

[10] DEUSCHEL, J.-D. (1986). Nonlinear smoothing of infinite-dimensional diffusion processes. *Stochastics* **19** 237–261. MR0872463

[11] FELLER, W. (1940). On the integro-differential equations of purely discontinuous Markoff processes. *Trans. Amer. Math. Soc.* **48** 488–515. MR0002697

[12] FÖLLMER, H. and SNELL, J. L. (1977). An "inner" variational principle for Markov fields on a graph. *Z. Wahrsch. Verw. Gebiete* **39** 187–195. MR0445642

[13] GEORGII, H.-O. (2011). *Gibbs Measures and Phase Transitions*, 2nd ed. de Gruyter, Berlin. MR2807681

[14] ISRAEL, R. B. (1979). *Convexity in the Theory of Lattice Gases*. Princeton Univ. Press, Princeton, NJ. MR0517873

[15] KALLENBERG, O. (1983). *Random Measures*, 3rd ed. Akademie Verlag, Berlin. MR0818219

[16] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. Springer, New York. MR1876169

[17] MECKE, J., NAGEL, W. and WEISS, V. (2008). A global construction of homogeneous random planar tessellations that are stable under iteration. *Stochastics* **80** 51–67. MR2384821

[18] NAGEL, W. and WEISS, V. (2005). Crack STIT tessellations: Characterization of stationary random tessellations stable with respect to iteration. *Adv. in Appl. Probab.* **37** 859–883. MR2193987

[19] OKABE, A., BOOTS, B., SUGIHARA, K. and CHIU, S. N. (2000). *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley, Chichester. MR1770006

[20] REDENBACH, C. and THÄLE, C. (2013). On the arrangement of cells in planar STIT and Poisson line tessellations. *Methodol. Comput. Appl. Probab.* **15** 643–654. MR3085884

[21] ROSS, S. M. (2003). *Introduction to Probability Models*, 8th ed. Academic Press, Amsterdam.

[22] SCHNEIDER, R. and WEIL, W. (2008). *Stochastic and Integral Geometry*. Springer, Berlin. MR2455326

[23] SCHREIBER, T. and THÄLE, C. (2010). Second-order properties and central limit theory for the vertex process of iteration infinitely divisible and iteration stable random tessellations in the plane. *Adv. in Appl. Probab.* **42** 913–935. MR2796670

[24] SCHREIBER, T. and THÄLE, C. (2011). Intrinsic volumes of the maximal polytope process in higher dimensional STIT tessellations. *Stochastic Process. Appl.* **121** 989–1012. MR2775104

[25] SCHREIBER, T. and THÄLE, C. (2012). Second-order theory for iteration stable tessellations. *Probab. Math. Statist.* **32** 281–300. MR3021459

[26] SCHREIBER, T. and THÄLE, C. (2013). Shape-driven nested Markov tessellations. *Stochastics* **85** 510–531. MR3176472

[27] SCHREIBER, T. and THÄLE, C. (2013). Limit theorems for iteration stable tessellations. *Ann. Probab.* **41** 2261–2278. MR3098072

[28] SCHREIBER, T. and THÄLE, C. (2013). Geometry of iteration stable tessellations: Connection with Poisson hyperplanes. *Bernoulli* **19** 1637–1654. MR3129028

[29] STOYAN, D., KENDALL, D. G. and MECKE, J. (1995). *Stochastic Geometry*, 2nd ed. Wiley, Chichester.

[30] THÄLE, C., WEISS, V. and NAGEL, W. (2012). Spatial STIT tessellations: Distributional results for I-segments. *Adv. in Appl. Probab.* **44** 635–654. MR3024603

[31] VARADHAN, S. R. S. (1988). Large deviations and applications. In *École D'Été de Probabilités de Saint-Flour XV–XVII*, 1985–87. *Lecture Notes in Math.* **1362** 1–49. Springer, Berlin. MR0983371

H.-O. GEORGII
MATHEMATISCHES INSTITUT
LUDWIG-MAXIMILIANS-UNIVERSITÄT
THERESIENSTRASSE 39
80333 MÜNCHEN
GERMANY
E-MAIL: georgii@math.lmu.de

T. SCHREIBER
FACULTY OF MATHEMATICS
AND COMPUTER SCIENCE
NICOLAUS COPERNICUS UNIVERSITY
UL. CHOPINA 12/18
87-100 TORUŃ
POLAND

C. THÄLE
FAKULTÄT FÜR MATHEMATIK
RUHR-UNIVERSITÄT BOCHUM
UNIVERSITÄTSSTRASSE 150
44801 BOCHUM
GERMANY
E-MAIL: christoph.thaele@rub.de