Adv. Appl. Prob. (SGSA) **46,** 643–660 (2014)
Printed in Northern Ireland

STIT TESSELLATIONS HAVE TRIVIAL TAIL σ-ALGEBRA

SERVET MARTÍNEZ,* *Universidad de Chile*
WERNER NAGEL,** *Friedrich-Schiller-Universität Jena*

Abstract

We consider homogeneous STIT tessellations Y in the ℓ-dimensional Euclidean space $\mathbb{R}^\ell$ and show the triviality of the tail σ-algebra. This is a sharpening of the mixing result by Lachièze-Rey (2001).

Keywords: Stochastic geometry; random process of tessellations; ergodic theory; tail σ-algebra

2010 Mathematics Subject Classification: Primary 60D05
Secondary 60J25; 60J75; 37A25

1. Introduction

Let $Y = (Y_t : t > 0)$ be the STIT tessellation process, which is a Markov process taking values in the space of tessellations of the ℓ-dimensional Euclidean space $\mathbb{R}^\ell$. The process Y is spatially stationary (that is, its law is invariant under translations of the space) and on every polytope with nonempty interior W (called a window) the induced tessellation process, denoted by $Y \wedge W = (Y_t \wedge W : t > 0)$, is a pure jump process. The process Y was first constructed in [16], and in Section 4 we give a brief description of it and recall some of its main properties.

In stochastic geometry, ergodic and mixing properties as well as weak dependencies in space are studied. For example, Heinrich *et al.* considered mixing properties for Voronoi and some other tessellations, Poisson cluster processes, and germ–grain models, and derived laws of large numbers and central limit theorems; see [6], [9], and [8]. In the monograph by Daley and Vere-Jones [4, pp. 207–209] the concept of the tail σ-algebra and of mixing conditions for random measures and point processes is developed.

For STIT tessellations, Lachièze-Rey [12] showed that they are mixing in space. We pursue the problem of whether sharper mixing properties also hold. In the present paper we prove tail triviality for STIT tessellations. This is a step towards stronger mixing conditions like α- or β-mixing; cf. [2, Chapter 2].

In Section 2 we introduce a definition for the tail σ-algebra $\mathcal{B}_\infty(\mathbb{T})$ on the space $\mathbb{T}$ of tessellations in the ℓ-dimensional Euclidean space $\mathbb{R}^\ell$. This definition relies essentially on the definition of the Borel σ-algebra with respect to the Fell topology on the set of closed subsets of $\mathbb{R}^\ell$ (cf. [18]) as well as on the mentioned definition of the tail σ-algebra for random measures and point processes, as given in [4].

Our main result is formulated in Section 3, Theorem 2: for the distribution of a STIT tessellation, the tail σ-algebra $\mathcal{B}_\infty(\mathbb{T})$ is trivial, i.e. all its elements (the terminal events) have

Received 15 October 2012; revision received 26 August 2013.
* Postal address: Departamento Ingeniería Matemática and Centro Modelamiento Matemático, Universidad de Chile, UMI 2807 CNRS, Casilla 170-3, Correo 3, Santiago, Chile. Email address: smartine@dim.uchile.cl
** Postal address: Institut für Stochastik, Friedrich-Schiller-Universität Jena, Ernst-Abbe-Platz 2, D-07743 Jena, Germany. Email address: werner.nagel@uni-jena.de

either probability 1 or 0. A detailed proof is given in Section 6. A key concept is the apparently new definition in Section 5 of the encapsulation time of a window inside another window.

Finally, we compare STIT tessellations with Poisson hyperplane tessellations, and we show that the tail σ-algebra $\mathcal{B}_\infty(\mathbb{T})$ is not trivial with respect to the distribution of Poisson hyperplane tessellations. This shows that STIT tessellations are intermediary in their long-range regime between Poisson hyperplane tessellations (with strong dependencies in very distant locations) and the Poisson–Voronoi tessellation (where the dependencies are local only).

Schreiber and Thäle [19], [20], [21] proved some limit theorems. They provided central limit theorems for the number of vertices and the total edge length in the two-dimensional case ($\ell = 2$). Furthermore, they proved that in dimensions $\ell \geq 3$ there appear nonnormal limits, e.g. for the total surface area of the cells of STIT tessellations. Because it is known (see [7]) that certain mixing properties—which are stronger than 'tail triviality'—provide sufficient conditions for central or other limit theorems, it remains as an open problem to characterize the strong mixing properties of STIT tessellations in a more precise way. The results by Schreiber and Thäle indicate that there will be some distinctions for different dimensions.

Let us mention that the triviality of the tail σ-algebra has revealed to be a powerful tool in some areas. In particular, in ergodic theory with the introduction of K-systems (see [17]) and in statistical physics where it characterizes when a Gibbs state is extremal for some specification (see Theorem 7.7 of [5]). The Lévy processes have trivial tail σ-algebra (Kolmogorov 0–1 law), and also around the origin the initial σ-algebra is trivial, so it satisfies the Blumenthal 0–1 law; see [1].

2. The space of tessellations and the tail σ-algebra

The concepts of ergodic theory, mixing, and triviality of invariant and of tail σ-algebras are well established for random processes on the real axis; see [10]. For random measures and point processes on a Euclidean space $\mathbb{R}^\ell$, the corresponding definitions can be found in [4]. In the abovementioned papers [6], [9], [8] models of stochastic geometry were considered.

In the present paper we adapt these concepts for random tessellations. An analogous result to Lemma 1 has already been shown in [10] and [4], and Lemma 2 is closely related to Theorem 2.14 of [2]. For completeness, we give a proof of these two lemmas.

Let $\mathbb{R}^\ell$ be the ℓ-dimensional Euclidean space, and denote by $\mathbb{T}$ the space of tessellations of this space as defined in [18, Chapter 10]. A tessellation can be considered as a set T of polytopes (the cells) with disjoint interiors and covering the Euclidean space, as well as a closed subset ∂T which is the union of the cell boundaries. There is an obvious one-to-one relation between both descriptions of a tessellation, and their measurable structures can be related appropriately; see [13] and [18].

Let $\mathcal{C}$ be the set of all compact subsets of $\mathbb{R}^\ell$. We endow $\mathbb{T}$ with the Borel σ-algebra $\mathcal{B}(\mathbb{T})$ of the Fell topology (also known as the topology of closed convergence), namely,

$$\mathcal{B}(\mathbb{T}) = \sigma(\{\{T \in \mathbb{T} \colon \partial T \cap C = \varnothing\} \colon C \in \mathcal{C}\}).$$

(As usual, for a class of sets $\mathcal{J}$, we denote by $\sigma(\mathcal{J})$ the smallest σ-algebra containing $\mathcal{J}$.)

A compact convex polytope W with nonempty interior in $\mathbb{R}^\ell$ is called a *window*. For a window W, we introduce the σ-algebra of sets of tessellations which are determined by restraints inside W:

$$\mathcal{B}_W(\mathbb{T}) = \sigma(\{\{T \in \mathbb{T} \colon \partial T \cap C = \varnothing\} \colon C \subseteq W,\ C \in \mathcal{C}\}).$$

By definition, $\mathcal{B}_W(\mathbb{T}) \subset \mathcal{B}(\mathbb{T})$ is a sub-σ-algebra. We note that if $W' \subseteq W$ then $\mathcal{B}_{W'}(\mathbb{T}) \subseteq \mathcal{B}_W(\mathbb{T})$.

Denote by $\mathbb{N} = \{1, 2, \ldots\}$ the set of positive integers. Let $(W_n : n \in \mathbb{N})$ be an increasing sequence of windows such that

$$\mathbb{R}^\ell = \bigcup_{n\in\mathbb{N}} W_n \quad \text{and} \quad W_n \subset \operatorname{int} W_{n+1} \quad \text{for all } n \in \mathbb{N}. \tag{1}$$

We have

$$\mathcal{B}_{W_n}(\mathbb{T}) \nearrow \mathcal{B}(\mathbb{T}) \quad \text{as } n \nearrow \infty,$$

which means that $\sigma(\bigcup_{n\in\mathbb{N}} \mathcal{B}_{W_n}(\mathbb{T})) = \mathcal{B}(\mathbb{T})$.

In order to define the tail σ-algebra, we will also consider sets of tessellations which are determined by their behavior outside a window W, i.e. in its complement W^c. We define the σ-algebra

$$\mathcal{B}_{W^c}(\mathbb{T}) = \sigma(\{\{T \in \mathbb{T} \colon \partial T \cap C = \varnothing\} \colon C \subset W^c,\ C \in \mathcal{C}\}).$$

We have $\mathcal{B}_{W^c}(\mathbb{T}) \subset \mathcal{B}(\mathbb{T})$. On the other hand, if $W' \subseteq W$ then $\mathcal{B}_{W^c}(\mathbb{T}) \subseteq \mathcal{B}_{W'^c}(\mathbb{T})$.

Let $(W_n : n \in \mathbb{N})$ be an increasing sequence of windows satisfying the conditions in (1). Then, for every window W, there exists an m such that $W \subseteq W_m$ and so $\mathcal{B}_{W_m^c}(\mathbb{T}) \subseteq \mathcal{B}_{W^c}(\mathbb{T})$. Hence, $\bigcap_{n=1}^{\infty} \mathcal{B}_{W_n^c}(\mathbb{T}) \subseteq \mathcal{B}_{W^c}(\mathbb{T})$. Since this happens for every window W, we have proven that $\bigcap_{n=1}^{\infty} \mathcal{B}_{W_n^c}(\mathbb{T}) \subseteq \bigcap_{W \text{ window}} \mathcal{B}_{W^c}(\mathbb{T})$. But the converse inclusion obviously holds and so we have $\bigcap_{n=1}^{\infty} \mathcal{B}_{W_n^c}(\mathbb{T}) = \bigcap_{W \text{ window}} \mathcal{B}_{W^c}(\mathbb{T})$. This equality allows us to define, in analogy with the definition for point processes (see [4, Definition 12.3.IV]), the tail σ-algebra on the space of tessellations.

Definition 1. The *tail σ-algebra* is defined as

$$\mathcal{B}_\infty(\mathbb{T}) = \bigcap_{W \text{ window}} \mathcal{B}_{W^c}(\mathbb{T}) = \bigcap_{n=1}^{\infty} \mathcal{B}_{W_n^c}(\mathbb{T}),$$

where $(W_n : n \in \mathbb{N})$ is an increasing sequence of windows satisfying (1).

Let us fix P a probability measure on $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$. For $\mathcal{E}, \mathcal{D} \in \mathcal{B}(\mathbb{T})$, we write $\mathcal{E} = \mathcal{D} \bmod \mathrm{P}$ if $\mathrm{P}(\mathcal{E} \Delta \mathcal{D}) = 0$. Also, for a pair $\mathcal{B}', \mathcal{B}''$ of sub-σ-algebras of $\mathcal{B}(\mathbb{T})$, we write $\mathcal{B}' \subseteq \mathcal{B}'' \bmod \mathrm{P}$ if, for all $\mathcal{E} \in \mathcal{B}'$, there exists $\mathcal{D} \in \mathcal{B}''$ such that $\mathcal{E} = \mathcal{D} \bmod \mathrm{P}$. It is easily checked that every set of $\mathcal{B}(\mathbb{T})$ can be approximated by some set in $\bigcup_{n\in\mathbb{N}} \mathcal{B}_{W_n}(\mathbb{T})$, which means that

$$\mathcal{B}(\mathbb{T}) = \{\mathcal{E} \in \mathcal{B}(\mathbb{T}) \colon \text{for every } \epsilon > 0, \text{ there exist } n \in \mathbb{N} \text{ and } \mathcal{E}_n \in \mathcal{B}_{W_n}(\mathbb{T}) \colon \ \mathrm{P}(\mathcal{E} \Delta \mathcal{E}_n) < \epsilon\}. \tag{2}$$

Now we consider invariance (both of measurable sets and of distributions) under the group of translations of $\mathbb{R}^\ell$. Let $h \in \mathbb{R}^\ell$. For any set $D \subseteq \mathbb{R}^\ell$, put $D + h = \{x + h \colon x \in D\}$ and, for $T \in \mathbb{T}$, denote by $T + h$ the tessellation with the boundary $\partial(T + h) = \partial(T) + h$. For $\mathcal{E} \subseteq \mathbb{T}$, put $\mathcal{E} + h = \{T + h \colon T \in \mathcal{E}\}$. For all $\mathcal{E} \in \mathcal{B}(\mathbb{T})$, we have $\mathcal{E} + h \in \mathcal{B}(\mathbb{T})$ because $\{C \in \mathcal{C}\} = \{C + h \colon C \in \mathcal{C}\}$. The probability measure P is translation invariant if it satisfies $\mathrm{P}(\mathcal{E}) = \mathrm{P}(\mathcal{E} + h)$ for all $\mathcal{E} \in \mathcal{B}(\mathbb{T})$ and $h \in \mathbb{R}^\ell$.

Now assume that P is translation invariant. A set $\mathcal{E} \in \mathcal{B}(\mathbb{T})$ is said to be (P-)invariant if $\mathrm{P}(\mathcal{E} \Delta (\mathcal{E} + h)) = 0$ for all $h \in \mathbb{R}^\ell$. Let $\mathcal{I}(\mathbb{T})$ denote the set of all invariant subsets of $\mathbb{T}$. Note that if $\mathcal{E} \in \mathcal{I}(\mathbb{T})$, $\mathcal{D} \in \mathcal{B}(\mathbb{T})$, and $\mathcal{E} = \mathcal{D} \bmod \mathrm{P}$, then $\mathcal{D} \in \mathcal{I}(\mathbb{T})$. It is easily checked that $\mathcal{I}(\mathbb{T}) \subseteq \mathcal{B}(\mathbb{T})$ is a sub-σ-algebra, the *invariant σ-algebra* (see, e.g. [4]). We have the following inclusion relation which corresponds to Exercise 12.3.2. of [4, pp. 208–215] for random measures. For completeness, we give the proof.

https://doi.org/10.1239/aap/1409319553 Published online by Cambridge University Press

Lemma 1. *Suppose that* P *is translation invariant. Then*

$$\mathcal{I}(\mathbb{T}) \subseteq \mathcal{B}_\infty(\mathbb{T}) \bmod \mathrm{P}. \tag{3}$$

Proof. We define $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^\ell$, so $a\mathbf{1} = (a, \ldots, a)$ for $a \in \mathbb{R}$. Fix the sequence $(W_n = [-n, n]^\ell : n \in \mathbb{N})$. Note that if $m > 2n$ then $W_n - m\mathbf{1} \subset W_n^c$.

Let $\mathcal{E} \in \mathcal{I}(\mathbb{T})$. Then it follows from (2) that, for all $n \in \mathbb{N}$, there exists $k = k(n) > n$ and $\mathcal{E}_k \in \mathcal{B}_{W_k}(\mathbb{T})$ such that $\mathrm{P}(\mathcal{E} \Delta \mathcal{E}_k) < 2^{-n}$. Since $\mathcal{E}_k \in \sigma(\{\{T \in \mathbb{T} \colon \partial T \cap C = \varnothing\} \colon C \subseteq W_k,\ C \in \mathcal{C}\})$, we have, for $N > 2k$,

$$\begin{aligned}\mathcal{E}_k - N\mathbf{1} &\in \sigma(\{\{T \in \mathbb{T} \colon \partial T \cap C = \varnothing\} \colon C \subseteq W_k - N\mathbf{1},\ C \in \mathcal{C}\})\\ &\subseteq \sigma(\{\{T \in \mathbb{T} \colon \partial T \cap C = \varnothing\} \colon C \subset W_k^c,\ C \in \mathcal{C}\})\\ &= \mathcal{B}_{W_k^c}(\mathbb{T}).\end{aligned}$$

Since $\mathrm{P}(\mathcal{E} \Delta (\mathcal{E} - N\mathbf{1})) = 0$ and $\mathrm{P}((\mathcal{E} - N\mathbf{1}) \Delta (\mathcal{E}_k - N\mathbf{1})) = \mathrm{P}(\mathcal{E} \Delta \mathcal{E}_k) < 2^{-n}$, we obtain

$$\mathrm{P}(\mathcal{E} \Delta (\mathcal{E}_k - N\mathbf{1})) \le \mathrm{P}(\mathcal{E} \Delta (\mathcal{E} - N\mathbf{1})) + \mathrm{P}((\mathcal{E} - N\mathbf{1}) \Delta (\mathcal{E}_k - N\mathbf{1})) < 2^{-n}$$

for all $N > 2k$. Now choose $N = k^2$, $k > 2$, and use $k^2 > 2k$ to obtain

$$\mathrm{P}(\mathcal{E} \Delta (\mathcal{E}_k - k^2\mathbf{1})) \le 2^{-n} \quad \text{and} \quad \mathcal{E}_k - k^2\mathbf{1} \in \mathcal{B}_{W_k^c}(\mathbb{T}).$$

For $m \ge 1$, define

$$\mathcal{D}_m = \bigcup_{n > m} (\mathcal{E}_{k(n)} - k(n)^2\mathbf{1}).$$

This sequence of sets satisfies $\mathrm{P}(\mathcal{E} \Delta \mathcal{D}_m) \le 2^{-m}$, $\mathcal{D}_m \in \mathcal{B}_{W^c_{k(m)}}(\mathbb{T})$ and $\mathcal{D}_m$ decreases with m. We conclude that $\mathcal{D} = \bigcap_{m>1} \mathcal{D}_m$ satisfies $\mathrm{P}(\mathcal{E} \Delta \mathcal{D}) = 0$ and $\mathcal{D} \in \mathcal{B}_\infty(\mathbb{T})$. Hence, relation (3) holds, completing the proof.

So, if the tail σ-algebra $\mathcal{B}_\infty(\mathbb{T})$ is P-trivial then P is ergodic with respect to translations, because every invariant set $\mathcal{E} \in \mathcal{I}(\mathbb{T})$ also belongs to $\mathcal{B}_\infty(\mathbb{T})$ andso $\mathrm{P}(\mathcal{E}) = 0$ or $\mathrm{P}(\mathcal{E}) = 1$. Moreover, it implies mixing, an implication that is found in Proposition 12.3.V of [4] for the case of random measures. That is, if P is translation invariant and $\mathcal{B}_\infty(\mathbb{T})$ is P-trivial then the action of translations is mixing, i.e.

$$\begin{aligned}&\mathcal{B}_\infty(\mathbb{T}) \text{ is P-trivial}\\ &\implies \lim_{|h| \to \infty} \mathrm{P}(\mathcal{D} \cap (\mathcal{E} + h)) = \mathrm{P}(\mathcal{D})\,\mathrm{P}(\mathcal{E}) \quad \text{for all } \mathcal{D}, \mathcal{E} \in \mathcal{B}(\mathbb{T}).\end{aligned} \tag{4}$$

A sufficient condition ensuring that the tail σ-algebra is trivial is given by the following result.

Lemma 2. *Assume that, for every window W', all $\mathcal{D} \in \mathcal{B}_{W'}(\mathbb{T})$, and all $\epsilon > 0$, there exists a window W depending on $(W', \mathcal{D}, \epsilon)$ such that*

$$W' \subset \operatorname{int} W \quad \textit{and} \quad |\mathrm{P}(\mathcal{D} \cap \mathcal{E}) - \mathrm{P}(\mathcal{D})\,\mathrm{P}(\mathcal{E})| < \epsilon \quad \textit{for all } \mathcal{E} \in \mathcal{B}_{W^c}(\mathbb{T}). \tag{5}$$

Then the tail σ-algebra $\mathcal{B}_\infty(\mathbb{T})$ is P*-trivial, that is,*

$$\mathrm{P}(\mathcal{E}) = 0 \quad \textit{or} \quad \mathrm{P}(\mathcal{E}) = 1 \quad \textit{for all } \mathcal{E} \in \mathcal{B}_\infty(\mathbb{T}).$$

https://doi.org/10.1239/aap/1409319553 Published online by Cambridge University Press

Proof. Let $\mathcal{D} \in \mathcal{B}_\infty(\mathbb{T})$. Let $(W_n : n \in \mathbb{N})$ be an increasing sequence of windows satisfying (1). Let $\epsilon > 0$ be fixed. Since $\mathcal{D} \in \mathcal{B}(\mathbb{T})$, from (2), there exist $k \in \mathbb{N}$ and $\mathcal{D}_k \in \mathcal{B}_{W_k}(\mathbb{T})$ such that

$$\mathrm{P}(\mathcal{D} \Delta \mathcal{D}_k) < \epsilon. \tag{6}$$

Hypothesis (5) implies that there exists a window W, depending on $(W_k, \mathcal{D}_k, \epsilon)$, such that $W_k \subset \operatorname{int} W$ and, for all $\mathcal{E} \in \mathcal{B}_{W^c}(\mathbb{T})$, it holds that $|\,\mathrm{P}(\mathcal{D}_k \cap \mathcal{E}) - \mathrm{P}(\mathcal{D}_k)\,\mathrm{P}(\mathcal{E})| < \epsilon$. We know that there exists an $n \geq k$ such that $W \subseteq W_n$. So, $\mathcal{B}_{W_n^c}(\mathbb{T}) \subseteq \mathcal{B}_{W^c}(\mathbb{T})$. We then have

$$|\,\mathrm{P}(\mathcal{D}_k \cap \mathcal{E}) - \mathrm{P}(\mathcal{D}_k)\,\mathrm{P}(\mathcal{E})| < \epsilon \quad \text{for all } \mathcal{E} \in \mathcal{B}_{W_n^c}(\mathbb{T}).$$

Since $\mathcal{D} \in \mathcal{B}_\infty(\mathbb{T}) \subset \mathcal{B}_{W_n^c}(\mathbb{T})$, we obtain $|\,\mathrm{P}(\mathcal{D}_k \cap \mathcal{D}) - \mathrm{P}(\mathcal{D}_k)\,\mathrm{P}(\mathcal{D})| < \epsilon$. From (6) we deduce that $|\,\mathrm{P}(\mathcal{D} \cap \mathcal{D}) - \mathrm{P}(\mathcal{D})\,\mathrm{P}(\mathcal{D})| < 3\epsilon$. Since this occurs for all $\epsilon > 0$, we conclude that $\mathrm{P}(\mathcal{D}) = \mathrm{P}(\mathcal{D})\,\mathrm{P}(\mathcal{D})$, so $\mathrm{P}(\mathcal{D}) = 1$ or 0.

3. Main results

Let Λ be a (nonzero) measure on the space of hyperplanes $\mathcal{H}$ in $\mathbb{R}^\ell$. It is assumed that Λ is translation invariant and possesses the following locally finiteness property:

$$\Lambda([B]) < \infty \quad \text{for all bounded sets } B \subset \mathbb{R}^\ell,$$

where $[B] = \{H \in \mathcal{H} : H \cap B \neq \varnothing\}$. It is further assumed that the support set of Λ is such that there is no line in $\mathbb{R}^\ell$ with the property that all hyperplanes of the support are parallel to it (in order to obtain almost surely (a.s.) bounded cells in the constructed tessellation (cf. [18, Theorem 10.3.2]), which can also be applied to STIT tessellations).

Let $\mathbb{P}$ denote the law of the STIT process Y.

Theorem 1. *Let $Y = (Y_t : t > 0)$ denote a STIT tessellation process defined by the measure Λ on $\mathcal{H}$ in $\mathbb{R}^\ell$, and let W' be a window. Then, for all $t > 0$ and all $\varepsilon > 0$, there exists a window W such that $W' \subset \operatorname{int} W$, and, for all $\mathcal{D} \in \mathcal{B}_{W'}(\mathbb{T})$ and $\mathcal{E} \in \mathcal{B}_{W^c}(\mathbb{T})$,*

$$|\mathbb{P}(Y_t \in \mathcal{D} \cap \mathcal{E}) - \mathbb{P}(Y_t \in \mathcal{D})\mathbb{P}(Y_t \in \mathcal{E})| < \varepsilon.$$

Let $\mathrm{P}^t(\mathcal{E}) = \mathbb{P}(Y_t \in \mathcal{E})$ for $\mathcal{E} \in \mathcal{B}(\mathbb{T})$ be the marginal law of Y_t in $\mathbb{T}$. Theorem 1 implies that P^t satisfies the sufficient condition (5), and, thus, Lemma 2 implies the triviality of the tail σ-algebra $\mathcal{B}_\infty(\mathbb{T})$ for P^t. Hence, the following result holds.

Theorem 2. *For all $t > 0$, the tail σ-algebra is trivial for the STIT process Y, that is, for all $\mathcal{E} \in \mathcal{B}_\infty(\mathbb{T})$, we have $\mathbb{P}(Y_t \in \mathcal{E}) = 0$ or $\mathbb{P}(Y_t \in \mathcal{E}) = 1$.*

Relation (4) ensures that Theorem 2 is stronger than the mixing property shown by Lachièze-Rey [12].

4. The STIT model

4.1. A construction of STIT tessellations

For a window W, we denote the set of all tessellations of W by $\mathbb{T} \wedge W$. If $T \in \mathbb{T}$, we denote by $T \wedge W$ the induced tessellation on W. Its boundary is defined by $\partial(T \wedge W) = (\partial T \cap W) \cup \partial W$.

In the present paper we will refer to the construction given in [16] for all details (for an alternative but equivalent construction, see [13]). On every window W there exists a STIT

https://doi.org/10.1239/aap/1409319553 Published online by Cambridge University Press

tessellation process $Y \wedge W = (Y_t \wedge W : t > 0)$. It turns out that this is a pure jump Markov process and, hence, has the strong Markov property (see [3, Proposition 15.25]). Each marginal $Y_t \wedge W$ takes values in $\mathbb{T} \wedge W$. Furthermore, for any $t > 0$, the law of Y_t is consistent with respect to the windows, that is, if W' and W are windows such that $W' \subseteq W$, then $(Y_t \wedge W) \wedge W' \sim Y_t \wedge W'$, where '$\sim$' denotes the identity of distributions (for a proof, see [16]). This yields the existence of a STIT tessellation Y_t of $\mathbb{R}^\ell$ such that, for all windows W, the law of $Y_t \wedge W$ coincides with the law of the construction in the window. A global construction for a STIT process was provided in [15]. A STIT tessellation process $Y = (Y_t : t > 0)$ is a Markov process and each marginal Y_t takes values in $\mathbb{T}$.

Here, let us recall roughly the construction of $Y \wedge W$ carried out in [16].

The assumptions made on Λ at the beginning of Section 3 imply that $0 < \Lambda([W]) < \infty$ for every window W. Denote by $\Lambda_{[W]}$ the restriction of Λ to $[W]$ and by $\Lambda^W = \Lambda([W])^{-1}\Lambda_{[W]}$ the normalized probability measure. Let us take two independent families of independent random variables $D = (d_{n,m} : n, m \in \mathbb{N})$ and $\tau = (\tau_{n,m} : n, m \in \mathbb{N})$, where each $d_{n,m}$ has distribution Λ^W and each $\tau_{n,m}$ is exponentially distributed with parameter $\Lambda([W])$.

1. Even if, for $t = 0$, the STIT tessellation Y_0 is not defined in $\mathbb{R}^\ell$, we define $Y_0 \wedge W = \{W\}$, the trivial tessellation for the window W. Its unique cell is denoted by $C^1 = W$.

2. Any extant cell has a random lifetime, and at the end of its lifetime it is divided by a random hyperplane. The lifetime of $W = C^1$ is $\tau_{1,1}$, and at that time it is divided by $d_{1,1}$ into two cells denoted by C^2 and C^3.

3. Now, for any cell C^i which is generated in the course of the construction, the random sequences $(d_{i,m} : m \in \mathbb{N})$ and $\tau = (\tau_{i,m} : m \in \mathbb{N})$ are used, and the following rejection method is applied.

4. When the time $\tau_{i,1}$ is elapsed, the random hyperplane $d_{i,1}$ is thrown onto the window W. If it does not intersect C^i then this hyperplane is rejected, and we continue until a number z_i, which is the first index j for which a hyperplane $d_{i,j}$ intersects and, thus, divides C^i into two cells, $C^{l_1(i)}$ and $C^{l_2(i)}$, which are called the successors of C^i. Note that this random number z_i is finite a.s. Hence, the lifetime of C^i is a sum $\tau^*(C^i) := \sum_{m=1}^{z_i} \tau_{i,m}$. It is easy to check, see [16], that $\tau^*(C^i)$ is exponentially distributed with parameter $\Lambda([C^i])$. Note that $z_1 = 1$ and $\tau^*(C^1) = \tau_{1,1}$.

5. This procedure is performed for any extant cell independently. It starts in that moment when the cell is born by division of a larger (the predecessor) cell. In order to guarantee independence of the division processes for the individual cells, the successors of C^i get indexes $l_1(i), l_2(i)$ in $\mathbb{N}$ that are different, and that can be chosen as the smallest numbers which were not yet used before for other cells.

6. For each cell C^i, we denote by $\eta = \eta(i)$ the number of its ancestors and we denote by $(k_1(i), k_2(i), \dots, k_\eta(i))$ the sequence of indexes of the ancestors of C^i. So $W = C^{k_1(i)} \supset C^{k_2(i)} \supset \cdots \supset C^{k_\eta(i)} \supset C^i$. Hence, $k_1(i) = 1$. The cell C^i is born at time $\underline{\kappa}(C^i) = \sum_{l=1}^{\eta} \tau^*(C^{k_l(i)})$ and it dies at time $\bar{\kappa}(C^i) = \underline{\kappa}(C^i) + \tau^*(C^i)$; for C^1, this is $\underline{\kappa}(C^1) = 0$ and $\bar{\kappa}(C^1) = \tau^*(C^1)$. It is useful to put $k_{\eta+1}(i) = i$.

With this notation at each time $t > 0$, the tessellation $Y_t \wedge W$ is constituted by the cells C^i for which $\underline{\kappa}(C^i) \le t$ and $\bar{\kappa}(C^i) > t$. It is easy to see that at any time, a.s. at most one cell dies and a.s. at most only two cells are born.

Now we describe the generated cells as intersections of half-spaces. First note that, by translation invariance, it follows that $\Lambda([\{0\}]) = 0$. Hence, all the random hyperplanes $(d_{n,m}: n, m \in \mathbb{N})$ a.s. do not contain the point 0. Now, for a hyperplane $H \in \mathcal{H}$ such that $0 \notin H$, we denote by H^+ and H^- the closed half-spaces generated by H with the convention $0 \in \operatorname{int}(H^+)$. Hence, $C^{k_{l+1}(i)} = C^{k_l(i)} \cap d^{\pm}_{k_l(i),z_{k_l(i)}}$, where the sign in the upper index determines on which side of the dividing hyperplane the cell $C^{k_{l+1}(i)}$ is located. Then any cell can be represented as an intersection of W with half-spaces:

$$C^i = W \cap \bigcap_{l=1}^{\eta(i)} \bigcap_{m=1}^{z_{k_l(i)}} d^{s(k_l(i),m)}_{k_l(i),m} \cap \bigcap_{m=1}^{z_i-1} d^{s(i,m)}_{i,m}. \tag{7}$$

In the above relation we define the sign $s(j,m) \in \{+,-\}$ by the relation $C^j \subset d^{s(j,m)}_{j,m}$ for $j, m \in \mathbb{N}$. Note that the origin $0 \in C^i$ if and only if all signs in (7) satisfy $s(k_l(i),m) = +$ and $s(i,m) = +$.

Obviously, the set of cells can be organized as a dyadic tree by the relation 'C' is a successor of C'. This method of construction is used in [16]. For the following, it is important to observe that all the rejected hyperplanes $d_{k_l(i),m}$ are also included in this intersection because the intersection with the appropriate half-spaces does not alter the cell. Although the third set $\bigcap_{m=1}^{z_i-1} d^{s(i,m)}_{i,m}$ in intersection (7) does not modify the resulting set, we also include it because we will use this representation later.

In [16] it was shown that there is no explosion, so at each time $t > 0$ the number of cells of $Y_t \wedge W$, denoted by ξ_t, is finite a.s. Renumbering the cells, we write $\{C^i_t : i = 1, \dots, \xi_t\}$ for the set of cells of $Y_t \wedge W$.

4.2. Independent increments relation

The name STIT is an abbreviation for 'stochastic stability under the operation of iteration of tessellations'. Closely related to that stability is a certain independence of increments of the STIT process in time.

In order to explain the operation of iteration, we number the cells of a tessellation $T \in \mathbb{T}$ in the following way. Assign to each cell a reference point in its interior (e.g. the Steiner point, see [18, p. 613], or another point that is a.s. uniquely defined). Order the set of reference points of all cells of T by its distance from the origin. For random homogeneous tessellations, this order is a.s. unique. Then number the cells of T according to this order, starting with number 1 for the cell which contains the origin. Thus, we write $C(T)^1, C(T)^2, \dots$ for the cells of T.

For $T \in \mathbb{T}$ and $\vec{R} = (R^m : m \in \mathbb{N}) \in \mathbb{T}^{\mathbb{N}}$, we define the tessellation $T \boxplus \vec{R}$, referred to as the iteration of T and $\vec{R}$, by its set of cells

$$T \boxplus \vec{R} = \{C(T)^k \cap C(R^k)^l : k \in \mathbb{N}, l \in \mathbb{N},\ \operatorname{int}(C(T)^k \cap C(R^k)^l) \neq \varnothing\}.$$

So, we restrict R^k to the cell $C(T)^k$, and this is done for all $k = 1, \dots$. The same definition holds when the tessellation and the sequence of tessellations are restricted to some window.

To state the independence relation of the increments of the Markov process Y of STIT tessellations, we fix a copy of the random process Y and let $\vec{Y}' = (Y'^m : m \in \mathbb{N})$ be a sequence of independent copies of Y, all of them also being independent of Y. In particular, $Y'^m \sim Y$. For a fixed time $s > 0$, we set $\vec{Y}'_s = (Y'^m_s : m \in \mathbb{N})$. Then, from the construction and from the consistency property of Y, it is straightforward to see that the following property holds:

$$Y_{t+s} \sim Y_t \boxplus \vec{Y}'_s \quad \text{for all } t, s > 0. \tag{8}$$

https://doi.org/10.1239/aap/1409319553 Published online by Cambridge University Press

This relation was first stated in Lemma 2 of [16]. It implies that $Y_{2t} \sim Y_t \boxplus \vec{Y}_t'$. The STIT property means that

$$Y_t \sim 2(Y_t \boxplus \vec{Y}_t') \quad \text{for all } t > 0,$$

so $Y_t \sim 2Y_{2t}$. Here the multiplication with 2 stands for the transformation $x \mapsto 2x$, $x \in \mathbb{R}^\ell$.

5. The encapsulation time of an embedded window

In the present section we introduce the concept of the encapsulation time of a window W' which is embedded in a larger window W. This will provide an essential tool for the proof of the main result.

Let W' be a window, and let Λ be a measure on the set of hyperplanes, as introduced at the beginning of Section 3. Since Λ is supposed to be translation invariant, without loss of generality, it can be assumed that the origin $0 \in \operatorname{int}(W')$.

Let W be a window such that $W' \subset \operatorname{int}(W)$. A key idea is the investigation of the probability that the window W' and the complement W^{c} are separated by the edges of the STIT process Y (we say that W' is encapsulated within W) at some time, and, thus, after that time, the processes inside W' and outside W, are approximately independent.

For simplicity, denote by $C_t = C_t^1$ the (a.s. uniquely determined) cell of tessellation $Y_t \wedge W$ that contains the origin in its interior. Obviously, $C_0 = W$. Note that C_t decreases as time t increases. On the other hand, since $W' \subset W$, when we consider the STIT on W', we can take $(Y_t \wedge W) \wedge W'$.

Definition 2. Let W' and W be two windows with $0 \in \operatorname{int}(W')$ and $W' \subset \operatorname{int}(W)$, and let $t > 0$. We say that W' *is encapsulated inside W at time t* if the cell C_t that contains 0 in $Y_t \wedge W$ is such that

$$W' \subseteq C_t \subset \operatorname{int}(W).$$

We write $W'|_t W$ if W' is encapsulated inside W at time t.

We denote the *encapsulation time* by

$$\mathcal{S}(W', W) = \inf\{t > 0 \colon W'|_t W\},$$

where, as usual, we put $\mathcal{S}(W', W) = \infty$ if $\{t > 0 \colon W'|_t W\} = \varnothing$. Encapsulation of W' inside W means that $\mathcal{S}(W', W) < \infty$ or, equivalently, $W'|_t W$ for some $t > 0$. In other words, the boundaries $\partial W'$ and ∂W are completely separated by facets of the 0-cell before the smaller window W' is hit for the first time by a facet of the STIT tessellation.

We have $\{\mathcal{S}(W', W) \le t\} \in \sigma(Y_s \colon s \le t)$, so $\mathcal{S}(W', W)$ is a stopping time. Hence, $\mathcal{S}(W', W)$ is also a stopping time for the processes $Y \wedge W$. On the other hand, note that the distribution of $\mathcal{S}(W', W)$ does not depend on a particular method of construction of the STIT process Y. Furthermore, if the construction as described in Subsection 4.1 is performed in a window $\tilde{W} \supseteq W$ then the spatial consistency of STIT tessellations (see [16]) implies that $\mathcal{S}(W', W)$ does not depend on $\tilde{W}$. In some of our proofs we will assume that the starting process is Y, but in some others we will start from the STIT process $Y \wedge W$, as in the proof of Lemma 3.

Note that even if in the STIT construction we have 'independence (inside W' and outside W) after encapsulation', it has to be considered that the tessellation outside W also depends on the process until the encapsulation time.

For two Borel sets $A, B \subset \mathbb{R}^\ell$, we denote by

$$[A|B] = \{H \in \mathcal{H}: (A \subset \text{int}(H^+),\ B \subset \text{int}(H^-)) \quad \text{or} \quad (A \subset \text{int}(H^-),\ B \subset \text{int}(H^+))\},$$

the set of all hyperplanes that separate A and B. This set is a Borel set in $\mathcal{H}$.

For a window W (which is defined to be a convex polytope), we denote by f_a^W, $a = 1, \ldots, q$, the $(\ell - 1)$-dimensional facets of W. Let W' be another window such that $W' \subset \text{int}(W)$. We denote by

$$G_a = [W' | f_a^W], \qquad a = 1, \ldots, q,$$

the set of hyperplanes that separate W' from the facet f_a^W of W. Note that all these sets are nonempty, and they are not necessarily pairwise disjoint.

There exists a finite family $\{G'_a : a = 1, \ldots, q\}$ of pairwise disjoint nonempty measurable sets that satisfy

$$G'_a \subseteq G_a \quad \text{for all } a \in \{1, \ldots, q\}. \tag{9}$$

(For example, we can choose $G'_a = G_a \setminus \bigcup_{b<a} G_b$, or we can partition $\bigcup_{a=1}^q G_a$ alternatively.)

Lemma 3. *Let W' and W be two compact convex polytopes, with $W' \subset \text{int}(W)$. Let $\{G'_a : a = 1, \ldots, q\}$ be a finite class of nonempty disjoint measurable sets satisfying (9) and such that $\Lambda(G'_a) > 0$ for all $a = 1, \ldots, q$. Then,*

$$\mathbb{P}(\mathcal{S}(W', W) \le t) \ge \mathrm{e}^{-t\Lambda([W'])} \prod_{a=1}^{q} (1 - \mathrm{e}^{-t\Lambda(G'_a)}) + \int_0^t \Lambda([W'])\mathrm{e}^{-x\Lambda([W'])} \prod_{a=1}^{q} (1 - \mathrm{e}^{-x\Lambda(G'_a)})\, \mathrm{d}x. \tag{10}$$

Proof. Our starting point in the proof is the construction of the process $Y \wedge W$. Because we assume that $0 \in \text{int}(W')$, we focus on the genesis of the 0-cell $C_t = C_t^1$, $t \ge 0$, only. We use representation (7), and we emphasize that in this intersection the positive half-spaces of those hyperplanes $d_{i,m}$ which are rejected in the construction are also involved.

Now, we consider a Poisson point process

$$\Phi = \{(d_m, S_m) : m \in \mathbb{N}\}$$

on $[W] \times [0, \infty)$ with intensity measure $\Lambda^W \otimes \Lambda([W])\lambda_+$, where λ_+ denotes the Lebesgue measure on $\mathbb{R}_+ = [0, \infty)$. This point process can be considered as a marked hyperplane process where the marks are birth times (or as a 'rain of hyperplanes'). This choice of intensity measure corresponds to the families D and τ of random variables in Subsection 4.1: the interval between two sequential births of hyperplanes is exponentially distributed with parameter $\Lambda([W])$, and the law of the born hyperplanes is Λ^W. Thus, the S_m are sums of independent and identically distributed exponentially distributed random variables which are independent of the d_m. This corresponds to one of the standard methods to construct (marked) Poisson point processes (cf., e.g. [11]). Note that this Poisson process is used for the construction of the 0-cell exclusively.

Let η denote the number of ancestors of C_t with indexes $k_1, \ldots, k_\eta$, and let $Z_{k_l} = \sum_{i=1}^{l} z_{k_i}$, $l = 1, \ldots, \eta + 1$. Thus, we can write (7) for the 0-cell as

$$C_t = W \cap \bigcap_{l=1}^{\eta} \bigcap_{m=Z_{k_{l-1}}+1}^{Z_{k_l}} d_m^+ \cap \bigcap_{m=Z_{k_\eta}+1}^{Z_{k_{\eta+1}}-1} d_m^+ = W \cap \bigcap_{\{m:\, S_m < t\}} d_m^+.$$

https://doi.org/10.1239/aap/1409319553 Published online by Cambridge University Press

Define the random times

$$\sigma' = \min\{S\colon \text{there exist } (d, S) \in \Phi\colon d \in [W']\}$$
$$\text{and} \quad \sigma_a = \min\{S\colon \text{there exist } (d, S) \in \Phi\colon d \in G'_a\}, \qquad a = 1, \ldots q.$$

These are the first times that a hyperplane of Φ falls into the respective sets. Note that, a.s., these minima exist and are greater than 0 because all $\Lambda_{[W]}(G'_a) > 0$; furthermore, we are working on a bounded window W and Λ is assumed to be locally finite, so Λ^W is a probability measure. Let

$$\mathcal{M} = \max\{\sigma_a\colon a = 1, \ldots q\}.$$

By definition, for all $a = 1, \ldots, q$, there exists a $(d_{(a)}, S_{(a)}) \in \Phi$ with $d_{(a)} \in G'_a$ and $S_{(a)} \leq \mathcal{M}$. Then $f_a^W \subset d^-_{(a)}$ and $C_{S_{(a)}} \subseteq d^+_{(a)}$. Since $C_{\mathcal{M}} \subseteq C_{S_{(a)}}$, we deduce that

$$C_{\mathcal{M}} \subseteq \bigcap_{a=1}^{q} d^+_{(a)} \subset \operatorname{int}(W).$$

On the other hand, if $\sigma' \geq \mathcal{M}$ then W' is not intersected until the time $\mathcal{M} = \max\{\sigma_a\colon a = 1, \ldots, q\}$, so we have $W' \subseteq C_{\mathcal{M}}$. Then

$$W' \subseteq C_{\mathcal{M}} \subset \operatorname{int}(W).$$

We have shown that

$$\{\mathcal{M} \leq \sigma'\} \subseteq \{\mathcal{S}(W', W) \leq \mathcal{M}\}.$$

This relation straightforwardly implies the inclusion

$$\{\mathcal{M} \leq \min\{\sigma', t\}\} \subseteq \{\mathcal{S}(W', W) \leq t\}. \tag{11}$$

Indeed, from $\mathcal{M} \leq \sigma'$ we obtain $\mathcal{S}(W', W) \leq \mathcal{M}$, and we use $\mathcal{M} \leq t$ to obtain relation (11). We deduce that

$$\mathbb{P}(\mathcal{S}(W', W) \leq t) \geq \mathbb{P}(\mathcal{M} \leq \min\{\sigma', t\}).$$

Now, the sets $[W']$ and G'_a, $a = 1, \ldots, q$, are pairwise disjoint and, therefore, the restricted Poisson point processes $\Phi \cap ([W'] \times [0, \infty))$ and $\Phi \cap (G'_a \times [0, \infty))$, $a = 1, \ldots, q$, are independent. Hence, σ' and σ_a, $a = 1, \ldots, q$, are independent random variables. Then σ' and $\mathcal{M}$ are independent, and we obtain

$$\begin{aligned} \mathbb{P}(\mathcal{M} \leq \min\{\sigma', t\}) &= \mathbb{P}(\mathcal{M} \leq \sigma' \leq t) + \mathbb{P}(\mathcal{M} \leq t \leq \sigma') \\ &= \mathbb{P}(\mathcal{M} \leq \sigma' \leq t) + \mathbb{P}(\mathcal{M} \leq t)\mathbb{P}(t \leq \sigma'). \end{aligned}$$

Since σ' and σ_a, $a = 1, \ldots, q$, are exponentially distributed with the respective parameters $\Lambda([W'])$ and $\Lambda(G'_a) > 0$, $a = 1, \ldots, q$, we find that

$$\mathbb{P}(t \leq \sigma') = \mathrm{e}^{-t\Lambda([W'])} \quad \text{and} \quad \mathbb{P}(\mathcal{M} \leq t) = \prod_{a=1}^{q} (1 - \mathrm{e}^{-t\Lambda(G'_a)}).$$

Now, by denoting the density functions of σ' by p' and those of σ_i by p_i, we obtain

$$\begin{aligned} \mathbb{P}(\mathcal{M} \leq \sigma' \leq t) &= \int_0^t p'(x) \bigg(\int_0^x p_1(x_1)\,\mathrm{d}x_1 \cdots \int_0^x p_q(x_q)\,\mathrm{d}x_q \bigg)\,\mathrm{d}x \\ &= \int_0^t \Lambda([W'])\mathrm{e}^{-x\Lambda([W'])} \prod_{a=1}^{q} (1 - \mathrm{e}^{-x\Lambda(G'_a)})\,\mathrm{d}x. \end{aligned}$$

Therefore, (10) follows.

https://doi.org/10.1239/aap/1409319553 Published online by Cambridge University Press

Remark 1. (i) We stress that $\mathcal{M} \le \min\{\sigma', t\}$ is sufficient but not necessary for $\mathcal{S}(W', W) \le t$. There are other ways than separating the complete facets of W by single hyperplanes to encapsulate W' within W. Alternate geometric constructions are possible.

(ii) It is well known in convex geometry that $[W' | f_a^W] \ne \varnothing$ for all $a = 1, \ldots, q$. But, depending on the support of the measure Λ (in particular, if Λ is concentrated on a set of hyperplanes with only finitely many directions), there can be windows W such that $\Lambda([W' | f_a^W]) = 0$ for some a. In such cases the bound given in Lemma 3 is useless. Therefore, in the following, W will be adapted to Λ in order to have all $\Lambda(G'_a) > 0$. However, here we will not try to find an optimal W in the sense that the quantities $\Lambda(G'_a)$ could somehow be maximized.

(iii) As an example, consider the particular measure $\Lambda_\perp = \sum_{j=1}^{\ell} g_j \delta_j$, where $g_j > 0$ and δ_j is the translation invariant measure on the set of all hyperplanes that are orthogonal to the jth coordinate axis in $\mathbb{R}^\ell$, with the normalization $\delta_j([s_j]) = 1$, where s_j is a linear segment of length 1 and parallel to the jth coordinate axis. With each axis j we associate two indices $a \in \{1, \ldots, 2\ell\}$, where $a = j$ refers to the negative part of this axis and $a = j + \ell$ refers to its positive part. Let $W' = [-\alpha, \alpha]^\ell$ and $W = [-\beta, \beta]^\ell$ be two windows with $0 < \alpha < \beta$. If $(f_a^{W'}, f_a^W)$ is a pair of parallel facets (of W' and of W, respectively) which are orthogonal to the jth coordinate axis and hitting this axis at $-\alpha$ and $-\beta$, respectively, when $a = j$, or at α and β when $a = j+\ell,\ j = 1, \ldots, \ell$, we can choose G'_a as the set of all hyperplanes that are orthogonal to the jth axis and which separate f_a^W from $f_a^{W'}$. This implies that $\Lambda_\perp(G'_a) = g_j(\beta - \alpha)$. Simple geometric considerations yield

$$\{\mathcal{M} \le \min\{\sigma', t\}\} = \{\mathcal{S}(W', W) \le t\} \quad \text{a.s.},$$

and, hence, for $\Lambda_\perp$, we have the equality sign in (10).

We will use the following parameterization of hyperplanes. Let $\mathbb{S}^{\ell-1}$ be the unit hypersphere in $\mathbb{R}^\ell$. For $H \in \mathcal{H}$, $d(H) \in \mathbb{R}$ denotes its signed distance from the origin and $u(H) \in \mathbb{S}^{\ell-1}$ is its normal direction. We denote by $H(u, d)$ the hyperplane with the respective parameters $(u, d) \in \mathbb{R} \times \mathbb{S}^{\ell-1}$. The image of a nonzero, locally finite, and translation invariant measure Λ with respect to this parameterization can be written as the product measure

$$\gamma\lambda \otimes \theta, \tag{12}$$

where $\gamma > 0$ is a constant, λ is the Lebesgue measure on $\mathbb{R}$, and θ is an even probability measure on $\mathbb{S}^{\ell-1}$ (cf., e.g. [18, Theorem 4.4.1 and Theorem 13.2.12]). Here θ is even means that $\theta(A) = \theta(-A)$ for all Borel sets $A \subseteq \mathbb{S}^{\ell-1}$. The property that there is no line in $\mathbb{R}^\ell$ such that all hyperplanes of the support of Λ are parallel to it is equivalent to the property that θ is not concentrated on a great subsphere of $\mathbb{S}^{\ell-1}$, i.e. there is no one-dimensional subspace L_1 of $\mathbb{R}^\ell$ (with the orthogonal complement $L_1^\perp$) such that the support of θ equals $\mathbb{G} = L_1^\perp \cap \mathbb{S}^{\ell-1}$.

Lemma 4. *For all windows W' with $0 \in \operatorname{int}(W')$, there exists a compact convex polytope W with facets $f_1^W, \ldots, f_{2\ell}^W$ and $W' \subset \operatorname{int}(W)$, and pairwise disjoint sets $G'_a \subseteq [W' | f_a^W]$ such that $\Lambda(G'_a) > 0$ for all $a = 1, \ldots, 2\ell$.*

Proof. For $u \in \mathbb{S}^{\ell-1}$, we denote by $H_{W'}(u)$ the supporting (i.e. tangential) hyperplane to W' with normal direction $u \in \mathbb{S}^{\ell-1}$. By $h_{W'}(u)$ we denote the distance from the origin to $H_{W'}(u)$. This is the support function of W': $h_{W'}(u) = \max\{\langle x, u\rangle : x \in W'\}$ (see [18, p. 600]). With this notation we have $H_{W'}(u) = H(u, h_{W'}(u))$. Note that, for $d \in \mathbb{R}$, the hyperplane $H(u, d)$ is parallel to $H_{W'}(u)$ at a signed distance d from the origin.

The shape of the window W will depend on Λ. We use some ideas given in the proof of Theorem 10.3.2 of [18]. Under the given assumptions on the support of Λ, there exist points $u_1, \dots, u_{2\ell} \in \mathbb{S}^{\ell-1}$ which all belong to the support of θ and $0 \in \operatorname{int}(\operatorname{conv}\{u_1, \dots, u_{2\ell}\})$, i.e. the origin is in the interior of the convex hull. Now, the facets f_a^W of W are chosen to have normals u_a, and their distance from the origin is $|h_{W'}(u_a)| + 3$, $a = 1, \dots, 2\ell$. Formally,

$$W = \bigcap_{a=1}^{2\ell} H(u_a, h_{W'}(u_a) \pm 3)^+.$$

Note that the described condition on the choice of the directions $u_1, \dots, u_{2\ell}$ guarantees that W is bounded (see the proof of Theorem 10.3.2 of [18]).

From the definition of the support of a measure and some continuity arguments (applied to sets of hyperplanes), it follows that, for all u_a, there are pairwise disjoint neighborhoods $U_a \subset \mathbb{S}^{\ell-1}$ such that $\theta(U_a) > 0$, and, for the sets of hyperplanes,

$$G'_a = \{H \in \mathcal{H} \colon u(H) \in U_a,\ h_{W'}(u_a) \pm 1 < d(H) < h_{W'}(u_a) \pm 2\},$$

it holds that $G'_a \subset [W'|f_a^W]$. Hence,

$$\Lambda([W'|f_a^W]) \geq \Lambda(G'_a) = \gamma\theta(U_a) > 0 \quad \text{for all } a = 1, \dots, 2\ell.$$

Since the U_a are pairwise disjoint, the sets G'_a, $a = 1, \dots, 2\ell$, also have this property.

Remark 2. Because the directional distribution θ is assumed to be an even measure, in the construction above we can choose $u_{\ell+a} = -u_a$, $a = 1, \dots, \ell$. Then the facets f_a^W and $f_{\ell+a}^W$ are parallel.

Lemma 5. *The compact convex polytope W constructed in Lemma 4 also satisfies the following property. For all $\varepsilon > 0$, there exists a $t^*(\varepsilon) > 0$ such that the following encapsulation time relation holds. For all $s \in (0, t^*(\varepsilon)]$ and all $r \geq r(s)$, there exists $r(s) \geq 1$ such that*

$$\mathbb{P}(\mathcal{S}(W', rW) \leq s) > 1 - \varepsilon. \tag{13}$$

Proof. Let us use the notation introduced in Lemma 4 and its proof. For $r > 0$, we set $rG'_a = \{rh \colon h \in G'_a\}$. Then, it follows from elementary linear algebra that $G'_a \subset [W'|f_a^W]$ implies that $rG'_a \subset [W'|rf_a^W]$ for all $r > 1$. Furthermore, from (12) we find that

$$\Lambda([W'|rf_a^W]) \geq \Lambda(rG'_a) = \gamma r\theta(U_a) \quad \text{for all } a = 1, \dots, 2\ell.$$

Now define

$$L = \min\{\Lambda(G'_a) = \gamma\theta(U_a) \colon a = 1, \dots, 2\ell\}.$$

We have $L > 0$, and (10) yields

$$\begin{aligned}\mathbb{P}(&\mathcal{S}(W', rW) \leq s)\\ &\geq \mathrm{e}^{-s\Lambda([W'])}(1 - \mathrm{e}^{-srL})^{2\ell} + \int_0^s \Lambda([W'])\mathrm{e}^{-x\Lambda([W'])}(1 - \mathrm{e}^{-xrL})^{2\ell}\,\mathrm{d}x\\ &> \mathrm{e}^{-s\Lambda([W'])}(1 - \mathrm{e}^{-srL})^{2\ell}.\end{aligned}$$

Note that, for all $\varepsilon \in (0, 1)$, there exists $t^*(\varepsilon) > 0$ such that

$$\mathrm{e}^{-t^*(\varepsilon)\Lambda([W'])} > \sqrt{1 - \varepsilon}.$$

Then, for all $s \in (0, t^*(\varepsilon)]$, we have $\mathrm{e}^{-s\Lambda([W'])} > \sqrt{1 - \varepsilon}$. Furthermore, for any such $s \in (0, t^*(\varepsilon)]$, there is an $r(s) \geq 1$ with $(1 - \mathrm{e}^{-sr(s)L})^{2\ell} > \sqrt{1 - \varepsilon}$. This completes the proof.

It is clear that (13) holds for all $s > 0$ if it holds for all $s \in (0, t^*(\varepsilon)]$. For our purposes, only values of s close to 0 are of interest. The emphasis put on the behavior for small s in the formulation of Lemma 5 also allows us to obtain easier proofs for the next propositions and to organize them in a logical structure that is easier to read.

6. Proof of Theorem 1

In the sequel we prove some auxiliary results for STIT tessellations to provide the proof of Theorem 1.

Lemma 6. *For all W', all $t > 0$, and all $\varepsilon > 0$ and $t^*(\varepsilon)$ such that (13) holds, there exists $t_1 = t_1(\varepsilon, t) \in (0, \min\{t, t^*(\varepsilon)\})$ such that, for all $t_2 \in (0, t_1]$,*

$$\mathbb{P}(Y \wedge W' \text{ has no jump in } [t - t_2, t)) > 1 - \varepsilon.$$

Proof. Let $\{C_t^{\prime i} : i = 1, \ldots, \xi_t'\}$ be the family of cells of the pure jump Markov process $Y_t \wedge W'$. The lifetimes of $C_t^{\prime i}$ are exponentially distributed with parameters $\Lambda([C_t^{\prime i}])$ and they are conditionally independent, conditioned that a certain set of cells is given at time t. Then, given a certain set of cells at t, the minimum of the lifetimes is exponentially distributed with parameter $\zeta_t = \sum_{i=1}^{\xi_t'} \Lambda([C_t^{\prime i}])$.

Note that ζ_t is monotonically increasing in t, because if at some time a cell C' is divided into the cells C'' and C''', we have $C' = C'' \cup C'''$ and $[C'] = [C''] \cup [C''']$. Then, by the subadditivity of Λ,

$$\Lambda([C']) = \Lambda([C''] \cup [C''']) \leq \Lambda([C'']) + \Lambda([C''']).$$

Since the process $Y \wedge W'$ has no explosion, for any fixed $t > 0$, there exists an $x_0 > 0$ such that, for all $s \in [0, t]$, we have $\mathbb{P}(\zeta_s \leq x_0) > \sqrt{1-\varepsilon}$. We fix $t_1 = t_1(\varepsilon, t) \in (0, \min\{t, t^*(\varepsilon)\})$ as a value which also satisfies $\mathrm{e}^{-t_1(\varepsilon)x_0} > \sqrt{1-\varepsilon}$. This yields, for all $t_2 \in (0, t_1]$,

$$\begin{aligned}
&\mathbb{P}(Y \wedge W' \text{ has no jump in } [t - t_2, t)) \\
&\quad \geq \mathbb{P}(Y \wedge W' \text{ has no jump in } [t - t_1, t)) \\
&\quad = \int_0^\infty \mathbb{P}(Y \wedge W' \text{ has no jump in } [t - t_1, t) \mid \zeta_{t-t_1} = x)\mathbb{P}(\zeta_{t-t_1} \in \mathrm{d}x) \\
&\quad \geq \int_0^{x_0} \mathbb{P}(Y \wedge W' \text{ has no jump in } [t - t_1, t) \mid \zeta_{t-t_1} = x)\mathbb{P}(\zeta_{t-t_1} \in \mathrm{d}x) \\
&\quad = \int_0^{x_0} \mathrm{e}^{-t_1 x}\mathbb{P}(\zeta_{t-t_1} \in \mathrm{d}x) \\
&\quad \geq \mathrm{e}^{-t_1 x_0}\mathbb{P}(\zeta_{t-t_1} \leq x_0) \\
&\quad > 1 - \varepsilon.
\end{aligned}$$

This completes the proof.

In the sequel, for $t > 0$ and $\varepsilon > 0$, the quantity $t_1 = t_1(\varepsilon, t)$ is that given by Lemma 6.

To avoid overburdened notation and since there will be no confusion, if $\mathcal{E} \in \mathcal{B}_{W'}(\mathbb{T})$, we will write $\mathcal{E}$ instead of $\mathcal{E} \wedge W'$, and, correspondingly, for all $t > 0$ and all $\mathcal{E} \in \mathcal{B}_{W'}(\mathbb{T})$,

$$\mathbb{P}(Y_t \in \mathcal{E}) = \mathbb{P}((Y_t \wedge W') \in (\mathcal{E} \wedge W')) = \mathbb{P}(Y_t \wedge W' \in \mathcal{E}).$$

This abbreviation also allows us to put, for all $\mathcal{E} \in \mathcal{B}_{W'}(\mathbb{T})$, $\mathcal{D} \in \mathcal{B}_{W^c}(\mathbb{T})$, and $s > 0$,

$$\mathbb{P}(Y_t \in \mathcal{E} \cap \mathcal{D},\ \mathcal{S}(W', W) < s) = \mathbb{P}(Y_t \wedge W' \in \mathcal{E},\ Y_t \in \mathcal{D},\ \mathcal{S}(W', W) < s).$$

Lemma 7. *For all $t > 0$, $\varepsilon > 0$, and $t_2 \in (0, t_1]$, we have*

$$|\mathbb{P}(Y_t \in \mathcal{E}) - \mathbb{P}(Y_{t-s} \in \mathcal{E})| < \varepsilon \quad \text{for all } s \in (0, t_2] \text{ and all } \mathcal{E} \in \mathcal{B}_{W'}(\mathbb{T}).$$

Proof. We have, for all $s \in (0, t_2]$,

$$\{Y \wedge W' \text{ has no jump in } [t - t_2, t)\} \subseteq \{Y \wedge W' \text{ has no jump in } [t - s, t)\}.$$

Then

$$\begin{aligned}\{Y_t \wedge W' \in \mathcal{E},\ Y \wedge W' &\text{ has no jump in } [t - t_2, t)\}\\ &= \{Y_{t-s} \wedge W' \in \mathcal{E},\ Y \wedge W' \text{ has no jump in } [t - t_2, t)\}.\end{aligned}$$

Therefore, $\{Y_t \wedge W' \in \mathcal{E}\}\Delta\{Y_{t-s} \wedge W' \in \mathcal{E}\} \subseteq \{Y \wedge W' \text{ has some jump in } [t - t_2, t)\}$. By using the relation $|\mathbb{P}(\Gamma) - \mathbb{P}(\Theta)| \le \mathbb{P}(\Gamma\Delta\Theta)$, the result follows.

In the following results W is a window such that $W' \subset \operatorname{int}(W)$. We use the notation $\mathcal{S} = \mathcal{S}(W', W)$ for the encapsulation time.

Proposition 1. *For all $t > 0$, $\varepsilon > 0$, and $t_2 \in (0, t_1]$, we have*

$$|\mathbb{P}(Y_t \in \mathcal{E}) - \mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2)| < \varepsilon \quad \text{for all } \mathcal{E} \in \mathcal{B}_{W'}(\mathbb{T}).$$

Proof. Let us first show that

$$\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} = s) = \mathbb{P}(Y_{t-s} \in \mathcal{E}) \quad \text{for all } s \in (0, t) \text{ and all } \mathcal{E} \in \mathcal{B}_{W'}(\mathbb{T}). \tag{14}$$

Let $\vec{Y}' = (Y'^m : m \in \mathbb{N})$ be a sequence of independent copies of Y, and also independent of Y. Relation (8) yields

$$Y_t \wedge W' \sim (Y_s \boxplus \vec{Y}'_{t-s}) \wedge W' \sim (Y_s \wedge W') \boxplus \vec{Y}'_{t-s}.$$

On the event $\mathcal{S} = s$ we have $W' \subseteq C^1_s$, the cell containing the origin at time s, and, thus, $Y_s \wedge W' = \{W'\}$. Hence, on $\mathcal{S} = s$ we have $(Y_s \wedge W') \boxplus \vec{Y}'_{t-s} \sim Y'^1_{t-s} \wedge W'$. Therefore,

$$\mathbb{P}(Y_t \wedge W' \in \mathcal{E} \mid \mathcal{S} = s) = \mathbb{P}(Y'^1_{t-s} \wedge W' \in \mathcal{E} \mid \mathcal{S} = s) = \mathbb{P}(Y'^1_{t-s} \wedge W' \in \mathcal{E}). \tag{15}$$

Since $Y'^1_{t-s} \sim Y_{t-s}$, relation (14) is satisfied. Hence,

$$\begin{aligned}&|\mathbb{P}(Y_t \in \mathcal{E}) - \mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2)|\\ &\quad= \left|\mathbb{P}(Y_t \in \mathcal{E}) - \frac{1}{\mathbb{P}(\mathcal{S} < t_2)} \int_0^{t_2} \mathbb{P}(Y_t \in \mathcal{E} \mid S = s)\mathbb{P}(\mathcal{S} \in \mathrm{d}s)\right|\\ &\quad= \left|\mathbb{P}(Y_t \in \mathcal{E}) - \frac{1}{\mathbb{P}(\mathcal{S} < t_2)} \int_0^{t_2} \mathbb{P}(Y_{t-s} \in \mathcal{E})\mathbb{P}(\mathcal{S} \in \mathrm{d}s)\right|\\ &\quad= \frac{1}{\mathbb{P}(S < t_2)}\left|\int_0^{t_2} \mathbb{P}(Y_t \in \mathcal{E}) - \mathbb{P}(Y_{t-s} \in \mathcal{E})\mathbb{P}(\mathcal{S} \in \mathrm{d}s)\right|\\ &\quad\le \frac{1}{\mathbb{P}(\mathcal{S} < t_2)} \int_0^{t_2} |\mathbb{P}(Y_t \in \mathcal{E}) - \mathbb{P}(Y_{t-s} \in \mathcal{E})|\mathbb{P}(\mathcal{S} \in \mathrm{d}s)\\ &\quad< \varepsilon,\end{aligned}$$

where in the last inequality we used Lemma 7.

Lemma 7, Proposition 1, and relation (14) obviously imply the following result.

Corollary 1. *For all $t > 0$, $\varepsilon > 0$, $t_2 \in (0, t_1]$, and all $s \in (0, t_2)$, it holds that*

$$|\mathbb{P}(Y_{t-s} \in \mathcal{E}) - \mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2)| = |\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} = s) - \mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2)|$$
$$< 2\varepsilon \quad \text{for all } \mathcal{E} \in \mathcal{B}_{W'}(T).$$

Lemma 8. *For all $t > 0$, $\varepsilon > 0$, and $t_2 \in (0, t_1]$, we have, for all $\mathcal{D} \in \mathcal{B}_{W'}(\mathbb{T})$ and $\mathcal{E} \in \mathcal{B}_{W^c}(\mathbb{T})$,*

$$|\mathbb{P}(Y_t \in \mathcal{D} \cap \mathcal{E} \mid \mathcal{S} < t_2) - \mathbb{P}(Y_t \in \mathcal{D} \mid \mathcal{S} < t_2)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2)| < 2\varepsilon.$$

Proof. Firstly, we show the following conditional independence property: for all $\mathcal{D} \in \mathcal{B}_{W'}(\mathbb{T})$ and $\mathcal{E} \in \mathcal{B}_{W^c}(\mathbb{T})$,

$$\mathbb{P}(Y_t \in \mathcal{D} \cap \mathcal{E} \mid \mathcal{S} = s) = \mathbb{P}(Y_t \in \mathcal{D} \mid \mathcal{S} = s)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} = s). \tag{16}$$

We use the notation introduced in the proof of Proposition 1, and we also write $\mathcal{E}$ instead of $\mathcal{E} \wedge W^c$ for short. Also, the arguments are close to those used in the proof of relation (14).

Recall that, on the event $\mathcal{S} = s$, we have $W' \subseteq C_s^1$. By $Y_s \boxplus (Y'^m_{t-s} : m \geq 2)$ we mean that the tessellations Y'^m_{t-s} are nested only into the cells C_s^m of Y_s with $m \geq 2$, and not into the cell C_s^1. From (14), (15), and the independence of the random variables Y_s, Y'^m_{t-s}, $m \geq 1$, we obtain

$$\begin{aligned}
&\mathbb{P}(Y_t \wedge W' \in \mathcal{D},\ Y_t \in \mathcal{E} \mid \mathcal{S} = s) \\
&\quad = \mathbb{P}((Y_s \boxplus \vec{Y}'_{t-s}) \wedge W' \in \mathcal{D},\ (Y_s \boxplus \vec{Y}'_{t-s}) \in \mathcal{E} \mid \mathcal{S} = s) \\
&\quad = \mathbb{P}(Y'^1_{t-s} \wedge W' \in \mathcal{D},\ (Y_s \boxplus (Y'^m_{t-s} : m \geq 2)) \in \mathcal{E} \mid \mathcal{S} = s) \\
&\quad = \mathbb{P}(Y'^1_{t-s} \wedge W' \in \mathcal{D})\mathbb{P}(Y_s \boxplus (Y'^m_{t-s} : m \geq 2) \in \mathcal{E} \mid \mathcal{S} = s) \\
&\quad = \mathbb{P}(Y_t \wedge W' \in \mathcal{D} \mid \mathcal{S} = s)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} = s).
\end{aligned}$$

Then (16) is verified. By using this equality and Corollary 1, we find that

$$\begin{aligned}
&\mathbb{P}(Y_t \in \mathcal{D} \mid \mathcal{S} < t_2)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2) - 2\varepsilon \\
&\quad \leq (\mathbb{P}(Y_t \in \mathcal{D} \mid \mathcal{S} < t_2) - 2\varepsilon)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2) \\
&\quad = \frac{1}{\mathbb{P}(S < t_2)} \int_0^{t_2} (\mathbb{P}(Y_t \in \mathcal{D} \mid \mathcal{S} < t_2) - 2\varepsilon)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} = s)\mathbb{P}(\mathcal{S} \in \mathrm{d}s) \\
&\quad < \frac{1}{\mathbb{P}(\mathcal{S} < t_2)} \int_0^{t_2} \mathbb{P}(Y_t \in \mathcal{D} \mid \mathcal{S} = s)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} = s)\mathbb{P}(\mathcal{S} \in \mathrm{d}s) \\
&\quad = \frac{1}{\mathbb{P}(\mathcal{S} < t_2)} \int_0^{t_2} \mathbb{P}(Y_t \in \mathcal{D} \cap \mathcal{E} \mid \mathcal{S} = s)\mathbb{P}(\mathcal{S} \in \mathrm{d}s) \\
&\quad = \mathbb{P}(Y_t \in \mathcal{D} \cap \mathcal{E} \mid \mathcal{S} < t_2).
\end{aligned}$$

In an analogous way, we can prove that

$$\mathbb{P}(Y_t \in \mathcal{D} \cap \mathcal{E} \mid \mathcal{S} < t_2) < \mathbb{P}(Y_t \in \mathcal{D} \mid \mathcal{S} < t_2)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2) + 2\varepsilon,$$

which completes the proof.

We summarize the argument as follows. The window W' was fixed and there was no loss of generality in assuming that $0 \in \mathrm{int}(W')$. Let $t > 0$ and $\varepsilon > 0$ be fixed. We construct $t_1 = t_1(\varepsilon, t) \in (0, \min\{t, t^*(\varepsilon)\})$. Now let W be as in Lemma 4. Note that, for all $t_2 \in (0, t_1(\varepsilon)]$, we have $\mathrm{e}^{-t_2\Lambda([W'])} > \sqrt{1-\varepsilon}$. Then, from Lemma 5, for all $t_2 \in (0, t_1(\varepsilon)]$, there exists $r(t_2) \ge 1$ such that $\mathbb{P}(\mathcal{S}(W', rW) < t_2) > 1 - \varepsilon$ for all $r \ge r(t_2)$.

Now we fix $t_2 \in (0, t_1(\varepsilon)]$ and in the sequel we write W instead of $r(t_2)W$ ($W := r(t_2)W$) for short. We have $W' \subset \mathrm{int}(W)$. From Lemma 8 we deduce the following result.

Corollary 2. *For all $\mathcal{D} \in \mathcal{B}_{W'}(\mathbb{T})$ and $\mathcal{E} \in \mathcal{B}_{W^{\mathrm{c}}}(\mathbb{T})$, it holds that*

$$|\mathbb{P}(Y_t \in \mathcal{D} \cap \mathcal{E} \mid \mathcal{S} < t_2) - \mathbb{P}(Y_t \in \mathcal{D} \mid \mathcal{S} < t_2)\mathbb{P}(Y_t \in \mathcal{E} \mid \mathcal{S} < t_2)| < 2\varepsilon.$$

Proposition 2. *For the window W', all $t > 0$, and $\varepsilon > 0$, the window W satisfies*

$$|\mathbb{P}(Y_t \in \mathcal{D} \cap \mathcal{E}) - \mathbb{P}(Y_t \in \mathcal{D})\mathbb{P}(Y_t \in \mathcal{E})| < 4\varepsilon \quad \textit{for all } \mathcal{D} \in \mathcal{B}_{W'}(\mathbb{T}) \textit{ and } \mathcal{E} \in \mathcal{B}_{W^{\mathrm{c}}}(\mathbb{T}).$$

Proof. Define

$$\Gamma = \{Y_t \in \mathcal{D}\}, \qquad \Theta = \{Y_t \in \mathcal{E}\}, \qquad \Upsilon = \{\mathcal{S}(W', W) < t_2\}.$$

We have

$$|\mathbb{P}(\Gamma \cap \Theta \mid \Upsilon) - \mathbb{P}(\Gamma \mid \Upsilon)\mathbb{P}(\Theta \mid \Upsilon)| < 2\varepsilon \quad \text{and} \quad \mathbb{P}(\Upsilon) > 1 - \varepsilon. \tag{17}$$

Observe that $\mathbb{P}(\Upsilon) > 1-\varepsilon$ implies that $\mathbb{P}(\Xi) - \mathbb{P}(\Xi \cap \Upsilon) < \varepsilon$ and $\mathbb{P}(\Xi) - \mathbb{P}(\Xi \cap \Upsilon)\mathbb{P}(\Upsilon) < 2\varepsilon$ for all events Ξ, in particular when Ξ is Γ, Θ, or $\Gamma \cap \Theta$.

The first relation in (17) obviously implies that

$$|\mathbb{P}(\Gamma \cap \Theta \mid \Upsilon) - \mathbb{P}(\Gamma \mid \Upsilon)\mathbb{P}(\Theta \mid \Upsilon)|\mathbb{P}(\Upsilon)^2 < 2\varepsilon.$$

Then

$$\begin{aligned} \mathbb{P}(\Gamma \cap \Theta) - 4\varepsilon &< \mathbb{P}(\Gamma \cap \Theta) - \mathbb{P}(\Gamma \cap \Theta \mid \Upsilon)\mathbb{P}(\Upsilon)^2 + \mathbb{P}(\Gamma \mid \Upsilon)\mathbb{P}(\Theta \mid \Upsilon)\mathbb{P}(\Upsilon)^2 - 2\varepsilon \\ &= \mathbb{P}(\Gamma \cap \Theta) - \mathbb{P}(\Gamma \cap \Theta \cap \Upsilon)\mathbb{P}(\Upsilon) + \mathbb{P}(\Gamma \cap \Upsilon)\mathbb{P}(\Theta \cap \Upsilon) - \mathbb{P}(\Gamma)\mathbb{P}(\Theta) \\ &\quad + \mathbb{P}(\Gamma)\mathbb{P}(\Theta) - 2\varepsilon \\ &< \mathbb{P}(\Gamma)\mathbb{P}(\Theta). \end{aligned}$$

In an analogous way, we show that $\mathbb{P}(\Gamma)\mathbb{P}(\Theta) < \mathbb{P}(\Gamma \cap \Theta) + 4\varepsilon$. Hence, the result is proven.

Proposition 2 yields the proof of Theorem 1, by substituting 4ε by ε.

7. Comparison of STIT and Poisson hyperplane tessellations

Intuitively, we expect a gradual difference in the mixing properties of STIT and Poisson hyperplane tessellations (PHTs). Hyperplanes are unbounded, while the maximum faces (also referred to as I-faces) in a STIT tessellation are always bounded. However, these I-faces tend to be very large (it has already been shown in [14] for the planar case that the length of the typical I-segment has a finite expectation but an infinite second moment). Since we have shown that the tail σ-algebra for STIT is trivial, then STIT has a short-range correlation in the sense of [4, Exercise 12.3.4].

One aspect is the following. For PHTs, Schneider and Weil [18, Section 10.5] showed that it is mixing if the directional distribution θ (see (12)) has zero mass on all great subspheres of $\mathbb{S}_+^{\ell-1}$. They contributed an example of a tessellation where this last condition is not fulfilled and which is not mixing. In contrast to this, Lachièze-Rey [12] proved that STIT is mixing for all θ which are not concentrated on a great subsphere.

Concerning the tail σ-algebra, we have Theorem 2 for STIT tessellations. In contrast, for PHTs, the tail σ-algebra is not trivial.

Lemma 9. *Let Y^{PHT} denote a Poisson hyperplane tessellation with intensity measure Λ which is a nonzero, locally finite, and translation invariant measure on $\mathcal{H}$, and assume that the support set of Λ is such that there is no line in $\mathbb{R}^\ell$ with the property that all hyperplanes of the support are parallel to it. Then the tail σ-algebra is not trivial with respect to the distribution of Y^{PHT}.*

Proof. Let $(W_n : n \in \mathbb{N})$ be an increasing sequence of windows such that, for all $n \in \mathbb{N}$, $W_n \subset \operatorname{int} W_{n+1}$ and $\mathbb{R}^\ell = \bigcup_{n\in\mathbb{N}} W_n$. Let B_1 be the unit ball in $\mathbb{R}^\ell$ centered at 0. Consider the following event in $\mathcal{B}(\mathbb{T})$:

$$\mathcal{D} := \{\text{there exists a hyperplane in } Y^{\mathrm{PHT}} \text{ which intersects } B_1\}.$$

Since outside of any bounded window W_n all the hyperplanes belonging to Y^{PHT} can be identified and, thus, it can be decided (in a measurable way) whether there is a hyperplane which intersects B_1, we have $\mathcal{D} \in \mathcal{B}_{W_n^c}(\mathbb{T})$ for all $n \in \mathbb{N}$, and, hence, $\mathcal{D} \in \mathcal{B}_\infty(\mathbb{T})$. We have $\mathbb{P}(\mathcal{D}) = 1 - \mathrm{e}^{-\Lambda([B_1])}$, and this probability is neither 0 nor 1.

Acknowledgements

The authors thank Lothar Heinrich for helpful hints and discussions. They are also very grateful to both of the anonymous referees for their careful reading and fruitful comments, in particular concerning the consequences of the tail triviality. (We partly used these comments in the introduction of the revised paper and added the references they suggested.) The authors are indebted for the support of Program Basal CMM from CONICYT (Chile) and by DAAD (Germany).

References

[1] Bertoin, J. (1996). *Lévy Processes.* Cambridge University Press.
[2] Bradley, R. C. (2007). *Introduction to Strong Mixing Conditions*, Vol. 1. Kendrick Press, Heber City, UT.
[3] Breiman, L. (1992). *Probability*, Corrected reprint of the 1968 original. SIAM, Philadelphia, PA.
[4] Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes*, Vol. II, *General Theory and Structure*, 2nd edn. Springer, New York.
[5] Georgii, H.-O. (1988). *Gibbs Measures and Phase Transitions.* De Gruyter, Berlin.
[6] Heinrich, L. (1994). Normal approximation for some mean-value estimates of absolutely regular tessellations. *Math. Meth. Statist.* **3,** 1–24.
[7] Heinrich, L. (2013). Asymptotic methods in statistics of random point processes. In *Stochastic Geometry, Spatial Statistics and Random Fields* (Lecture Notes Math. **2068**), Springer, Heidelberg, pp. 115–150.
[8] Heinrich, L. and Molchanov, I. S. (1999). Central limit theorem for a class of random measures associated with germ–grain models. *Adv. Appl. Prob.* **31,** 283–314.
[9] Heinrich, L., Schmidt, H. and Schmidt, V. (2007). Limit theorems for functionals on the facets of stationary random tessellations. *Bernoulli* **13,** 868–891.
[10] Ibragimov, I. A. and Linnik, Yu. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters-Noordhoff, Groningen.
[11] Kingman, J. F. C. (1993). *Poisson Processes.* Oxford University Press.
[12] Lachièze-Rey, R. (2011). Mixing properties for STIT tessellations. *Adv. Appl. Prob.* **43,** 40–48.
[13] Martínez, S. and Nagel, W. (2012). Ergodic description of STIT tessellations. *Stochastics* **84,** 113–134.

[14] Mecke, J., Nagel, W. and Weiss, V. (2007). Length distributions of edges in planar stationary and isotropic STIT tessellations. *J. Contemp. Math. Analysis* **42,** 28–43.
[15] 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.
[16] Nagel, W. and Weiss, V. (2005). Crack STIT tessellations: characterization of stationary random tessellations stable with respect to iteration. *Adv. Appl. Prob.* **37,** 859–883.
[17] Parry, W. (1981). *Topics in Ergodic Theory*. Cambridge University Press.
[18] Schneider, R. and Weil, W. (2008). *Stochastic and Integral Geometry.* Springer, Berlin.
[19] 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. Appl. Prob.* **42,** 913–935.
[20] Schreiber, T. and Thäle, C. (2011). Intrinsic volumes of the maximal polytope process in higher dimensional STIT tessellations. *Stoch. Process. Appl.* **121,** 989–1012.
[21] Schreiber, T. and Thäle, C. (2013). Limit theorems for iteration stable tessellations. *Ann. Prob.* **41,** 2261–2278.