A Stochastic Ramsey Theorem

Zibo Xu

A thesis submitted for the degree of
Doctor of Philosophy

Department of Mathematics
London School of Economics
and Political Science

September 2010

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is my own work.

Abstract

A stochastic extension of Ramsey's theorem is established. Any Markov chain generates a filtration relative to which one may define a notion of stopping time. A stochastic colouring is any k-valued ($k < \infty$) colour function defined on all pairs consisting of a bounded stopping time and a finite partial history of the chain truncated before this stopping time. For any bounded stopping time θ and any infinite history ω of the Markov chain, let $\omega|\theta$ denote the finite partial history up to time $\theta(\omega)$. It is first proved for $k = 2$ that for every $\epsilon > 0$ there is an increasing sequence $\theta_1 < \theta_2 < ...$ of bounded stopping times having the property that, with probability greater than $1-\epsilon$, the history ω is such that the values assigned to all pairs $(\omega|\theta_i,\ \theta_j)$, with $i < j$, are the same. Just as with the classical Ramsey theorem, an analogous finitary stochastic Ramsey theorem is obtained. Furthermore, with appropriate finiteness assumptions, the time one must wait for the last stopping time (in the finitary case) is uniformly bounded, independently of the probability transitions. The results are generalised to any finite number k of colours. A stochastic extension is derived for hypergraphs, but with rather weaker conclusions. The stochastic Ramsey theorem can be applied to the expected utility of a Markov chain to conclude that on some infinite increasing sequence of bounded stopping times the expected utility remains the same to within ϵ (also in probability).

Acknowledgements

I would like to thank my supervisors, Robert Simon and Adam Ostaszewski, for their guidance and encouragement. I would also like to thank Nick Bingham and Graham Brightwell for many discussions and suggestions. Finally, my special thanks go to Nan. Without her, this enterprise would have had little meaning.

Contents

1

Introduction

1.1 Stochastic Ramsey Theorem

Let C be a finite set whose elements are colours. Ramsey [13] proved that, for every function that assigns a colour $c(k,l) \in C$ to every two positive integers $k < l$, there is an increasing sequence of integers $n_1 < n_2 < \ldots$ such that $c(n_1, n_2) = c(n_i, n_j)$ for all $i < j$. (For a textbook treatment, see [7].) We prove a stochastic Ramsey theorem.

To state our result, we first establish some notation. We denote the set of natural numbers by $\mathbb{N}$. For all n in $\mathbb{N}$, let J_n be a nonempty countable set of states. We regard the sets J_n as discrete spaces, and define the *history space* to be $\Omega := \prod_n J_n$. For any n in $\mathbb{N}$, denote $\prod_{1 \leq m \leq n} J_m$ by Ω_n. Any q in Ω_n is a *partial history* with *length* n, and we denote the length of q by $\|q\|$. Write $\Omega_{<\infty}$ for $\bigcup_n \Omega_n$. We give Ω the Tychonoff product topology: for a basic open set U in the Tychonoff product Ω, we have $U = \prod_m U_m$ with each U_m open in J_m and with the *support* of U, i.e., $\{m : U_m \neq J_m\}$, finite. Let $\mathcal{F}_n$ be the σ-algebra on Ω of Borel sets generated by the basic open sets with support included in $\{1, \ldots, n\}$. That is, $\mathcal{F}_n$ is the collection of all events included in Ω that are known at stage n to be true or false (so knowable) from the course of history. Then $\{\mathcal{F}_n\}_{n \geq 0}$ is a *filtration*, i.e., an increasing family of sub-σ-algebras of $\mathcal{P}(\Omega)$: $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{P}(\Omega)$, where $\mathcal{P}(\Omega)$ is the power set of Ω. Define $\mathcal{F}_\infty$ to be the σ-algebra of Borel sets on Ω. Let

P be a probability measure on $(\Omega, \mathcal{F}_\infty)$. Call $(\Omega, \mathcal{F}_\infty)$ and $(\Omega, \mathcal{F}_\infty, P)$ the measurable space and the probability space *constructed from the sequence* $\{J_n\}_{n\in\mathbb{N}}$ of sets of states, respectively.

We write an element in Ω as $\omega = \{\omega(n)\}_{n\in\mathbb{N}}$ where $\omega(n)$ is in J_n for each n. We denote by $\omega|n$ the partial history of ω *truncated* at stage n, i.e., $\omega|n := \{\omega(1), ..., \omega(n)\}$. Define the *basic open set determined by* a q in Ω_n by $B(q) := \{\omega \in \Omega : \omega|n = q\}$. A mapping $\tau : \Omega \to \{0, 1, 2, ..., \infty\}$ is called a *stopping time* if

$$\forall n \leq \infty,\ \{\tau = n\} = \{\omega \in \Omega : \tau(\omega) = n\} \in \mathcal{F}_n.$$

That is, by stage n the decision whether or not to execute some action (referred to as 'stopping') is a 'knowable' event in relation to the information so far disclosed by history. For every n in $\mathbb{N}$, we denote by $\mathcal{T}_n$ the collection of all stopping times τ on Ω bounded by n, and put $\mathcal{T} := \bigcup_n \mathcal{T}_n$. For any $\tau \in \mathcal{T}$ and any $\omega \in \Omega$, it is natural to denote the partial history of ω *restricted to* τ by

$$\omega|\tau := \omega|\tau(\omega) = \{\omega(1), ..., \omega(\tau(\omega))\}.$$

We call $\omega|\tau$ the *stopping place* of τ on ω, and denote by $S_\tau := \{\omega|\tau : \omega \in \Omega\}$ the set of stopping places of τ. We say that τ is *consistent* with $q \in \Omega_n$, and write $\tau \in \mathcal{T}(q)$, if $\tau(\omega) > n\ \forall \omega \in B(q)$. Note that the collection $\mathcal{T}$ of bounded stopping times has a natural partial order. Namely, for any two bounded stopping times $\sigma, \tau \in \mathcal{T}$, we say that σ is *ahead of* τ, and write $\sigma < \tau$, if $\sigma(\omega) < \tau(\omega)\ \forall \omega \in \Omega$. Denote the *partial ordered pairs of bounded stopping times* by

$$\mathcal{T}^{(2)}_< := \{(\sigma, \tau) : \sigma, \tau \in \mathcal{T} \text{ and } \sigma < \tau\}.$$

We drop the subscript '$<$' when context allows. We also need the following notion of a *stochastic colouring.*

Definition 1.1. *Given a set C of colours, a stochastic colouring f is a mapping from $Z := \{(q, \tau) : q \in \Omega_{<\infty}$ and $\tau \in \mathcal{T}(q)\}$ to C. The induced stochastic colouring $\hat{f}$ of f is a mapping from $\mathcal{T}^{(2)} \times \Omega$ to C and defined by*

$$\hat{f}_{\sigma,\tau}(\omega) := f(\omega|\sigma, \tau)\ \forall (\sigma, \tau) \in \mathcal{T}^{(2)}\ \forall \omega \in \Omega.$$

Comment 1. The set Z here is the key for Definition 1.1 and later in the proof of Theorem 3.2 and Theorem 3.13. There some of the arguments reduce to the consideration of the case when all J_i are finite. In this circumstance, Z is countable. To see this, first note that, for given n in $\mathbb{N}$, each τ in $\mathcal{T}_n$ is uniquely encodable by reference to its stopping places, which are contained in a finite set of cardinality $\prod_{i \leq n} |J_i|$. By the same argument, for a given τ in $\mathcal{T}_n$, the set of partial histories q with which τ is consistent, i.e., $\tau \in \mathcal{T}(q)$, is finite.

Note that if some J_i is infinite, then Z is uncountable. To see this, suppose that $J_1 = \mathbb{N}$. For any $A \subseteq \mathbb{N}$, define a mapping $\tau_A : \Omega \to \mathbb{N}$ by

$$\tau_A(\omega) := \begin{cases} 1 & \text{if } \omega(1) \in A \\ 2 & \text{if } \omega(1) \notin A. \end{cases}$$

Each τ_A is a stopping time bounded by 2, since $\{\omega \in \Omega : \omega(1) \in A\} \in \mathcal{F}_1$. Furthermore, if $A \neq B$, then $\tau_A \neq \tau_B$. Because there are $2^{\mathbb{N}}$ many such τ_A, $\mathbb{Z}$ is uncountable.

Comment 2. For the definition of $\hat{f}_{\sigma,\tau}(\omega)$, note that if τ and σ are bounded stopping times with $\tau > \sigma$, then τ is consistent with $\omega|\sigma$, as

$$\tau(\omega) > \|\omega|\sigma\| = \sigma(\omega) \ \forall \omega \in \Omega.$$

Comment 3. We can replace 'bounded' τ by 'finite' τ in the definition of stochastic colouring. Indeed, if the replacement were made, the proofs of all theorems in this thesis would still only refer to the action of the stochastic colouring on all bounded stopping times.

We are now ready to state our stochastic Ramsey theorem. It implies the classical result when $|J_i| = 1 \ \forall i \in \mathbb{N}$.

Theorem 1.2. (Stochastic Infinitary Ramsey Theorem). *Given a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence $\{J_i\}_{i\in\mathbb{N}}$ of sets of states and a stochastic colouring f with values in a finite set C, then for every $\epsilon > 0$ there exists an increasing sequence of bounded stopping times*

$\theta_1 < \theta_2 < \theta_3 < \ldots$ *such that*

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j}\ \forall 1 \le i < j) > 1 - \epsilon,$$

where $\hat{f}$ *is the induced stochastic colouring of* f.

Comment. The complete version of the probability formula above is

$$P(\{\omega \in \Omega : \hat{f}_{\theta_1,\theta_2}(\omega) = \hat{f}_{\theta_i,\theta_j}(\omega)\ \forall 1 \le i < j\}) > 1 - \epsilon.$$

Given an increasing sequence of bounded stopping times $\theta_1 < \theta_2 < \theta_3 < \ldots$, we note that the set $\{\omega \in \Omega : \hat{f}_{\theta_1,\theta_2}(\omega) = \hat{f}_{\theta_i,\theta_j}(\omega)\ \forall 1 \le i < j\}$ is in $\mathcal{F}_\infty$. In this thesis, whenever a set S is trivially in $\mathcal{F}_\infty$, we use the shorthand $P(S)$ for $P(\{\omega : \omega \in S\})$, and omit the proof of that $S \in \mathcal{F}_\infty$.

Note that some special cases of the induced stochastic colouring $\hat{f}$ in Definition 1.1 will essentially induce results in the deterministic Ramsey theorem. For example, suppose that $\hat{f}_{\sigma,\tau}$ is a constant function on Ω for each $(\sigma, \tau) \in \mathcal{T}^{(2)}$. Then we can simply confine ourselves to the sequence of bounded stopping times $\{\theta_n\}$ in which $\theta_n(\omega) = n\ \forall \omega \in \Omega\ \forall n > 0$. The problem of finding an infinite sequence of bounded stopping times on which $\hat{f}$ is monochromatic is here simply a problem in classical Ramsey theory.

Shmaya and Solan [16] considered a stochastic Ramsey theorem in a general probability space $(\Omega, \mathcal{F}_\infty, P)$ with a given filtration $\{\mathcal{F}_n\}$ (where $\mathcal{F}_n \subseteq \mathcal{F}_\infty$ $\forall n \in \mathbb{N}$). They defined an NT function in the stochastic setting as follows. An NT function is one that assigns to every nonnegative integer n and every bounded stopping time τ an $\mathcal{F}_n$-measurable function $c_{n,\tau}$ defined over the set $\{\tau > n\}$ with range C. They also imposed an $\mathcal{F}$-consistency requirement: if $\tau_1 = \tau_2 > n$ on F with $F \in \mathcal{F}_n$, then $c_{n,\tau_1} = c_{n,\tau_2}$ on F. When σ and τ are two bounded stopping times with $\sigma < \tau$, they put $c_{\sigma,\tau}(\omega) := c_{\sigma(\omega),\tau}(\omega)$, and thus made $c_{\sigma,\tau}$ an $\mathcal{F}_\sigma$-measurable random variable. Under these conditions, they derived the weaker conclusion of the existence of a stepwise monochromatic path rather than a Ramsey theorem: for every finite set C of colours, every $\mathcal{F}$-consistent NT function c and every $\epsilon > 0$, there exists a sequence of bounded stopping times $0 \le \theta_1 < \theta_2 < \ldots$ such that $P(c_{\theta_1,\theta_2} = c_{\theta_i,\theta_{i+1}}\ \forall i) > 1 - \epsilon$. In their paper they commented that 'the natural stochastic generalisation of Ramsey's theorem requires the stronger condition that $P(c_{\theta_1,\theta_2} = c_{\theta_i,\theta_j}\ \forall 0 < i < j) \geq 1-\epsilon$. We do not know whether this generalisation is correct.' Their hunch turned out to be right, at least in our 'tree' context, as the current thesis demonstrates.

Our stochastic colouring in Definition 1.1 is almost the same as an NT function in the case that the measurable space $(\Omega, \mathcal{F}_\infty)$ is constructed from a sequence of sets of states, except that we do not need the $\mathcal{F}$-consistency requirement in Definition 1.1. Given a partial history $q \in \Omega_{<\infty}$, consider two bounded stopping times $\tau_1, \tau_2 \in \mathcal{T}(q)$ where $\omega|\tau_1 = \omega|\tau_2\ \forall \omega \in B(q)$. In this context, the $\mathcal{F}$-consistency condition of Shmaya and Solan requires that $f(q,\tau_1) = f(q,\tau_2)$. That is, for any $q \in \Omega_{<\infty}$ and any $\tau \in \mathcal{T}(q)$, it is the subset $\{\omega|\tau : \omega \in B(q)\}$ of $\{\omega|\tau : \omega \in \Omega\}$ that determines the value of $f(q,\tau)$. By contrast, our Definition 1.1 places no such requirement: the value of $f(q,\tau)$ in the above context is allowed to be determined by the larger set $\{\omega|\tau : \omega \in \Omega\}$; so Definition 1.1 also includes stochastic colourings with $\mathcal{F}$-consistency. Note that, in the case of $(\Omega, \mathcal{F}_\infty)$ constructed from a sequence of sets of countably many states (Markov chain), the Stochastic Infinitary Ramsey Theorem (Theorem 1.2) not only includes the result of Shmaya and Solan: $P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_{i+1}} \forall i) > 1-\epsilon$, but also answers affirmatively their open question.

1.2 The structure of this thesis

The outline of this thesis is as follows. After an introduction to the stochastic Ramsey theorem in this chapter, we review the classical Ramsey theorem in Chapter 2. Besides the standard proof, we give a detailed proof analogous to the proof of the stochastic Ramsey theorem to be shown in Chapter 3. One can compare these proofs for a deeper understanding of the Ramsey theorem in both deterministic and stochastic formulations.

In Chapter 3, we prove the Stochastic Infinitary Ramsey Theorem (Theorem 1.2) and results analogous to the classical finitary Ramsey theorem. For a more detailed introduction of this chapter, please see Section 3.2.

In Chapter 4, we consider the problems of hypergraphs in the Ramsey theory. In the deterministic context, given a natural number $k > 2$, we colour every set of k natural numbers $n_1 < n_2 < \dots < n_k$ and check for the existence of an infinite monochromatic subset A of $\mathbb{N}$ in the sense that every subset of cardinality k in A is assigned the same colour. While the extension of the deterministic (classical) Ramsey theorem to hypergraphs is a standard result, it is by no means a trivial problem in the stochastic variation. We have to first determine the best suited definition of colouring for hypergraphs in the stochastic context. The complexity of colouring multiple bounded stopping times then denies any easy generalisation of the proofs in Chapter 3, and in Theorem 4.2 (for $k = 3$) we have only a result weaker than the conjectured one.

In Chapter 5, we present an application of the stochastic Ramsey theorem in expected utility theory. We refer to a utility function u to value a random variable X which is regarded as a random monetary result. Our conclusion is that, for any $\epsilon > 0$, we can find an infinite increasing sequence of bounded stopping times such that with probability greater than $1 - \epsilon$, there exists a real number r and the expected utility on those stopping times is always bounded by $[r - \epsilon, r + \epsilon]$. (cf. Theorem 5.1)

2

The classical Ramsey theorem

Given k in $\mathbb{N}$, we denote $\{(n_1, ..., n_k) : n_l \in \mathbb{N}\ \forall 1 \leq l \leq k;\ n_i < n_j\ \forall 1 \leq i < j \leq k\}$ by $\mathbb{N}^{(k)}_{<}$. If no ambiguity, we drop the subscript '$<$'.

For simplicity, we here only apply colour functions c from $\mathbb{N}^{(2)}$ to a set of colours C, but we shall consider the general case that the domain of c is $\mathbb{N}^{(m)}$ for any $m = 2, 3, ...$ in Chapter 4. Recall the classical Ramsey theorem for the case of $\mathbb{N}^{(2)}$.

Theorem 2.1. *Given a set of colours $C = \{c_1, ..., c_k\}$, for every function $c : \mathbb{N}^{(2)} \to C$, there is a sequence of integers $n_1 < n_2 < ...$ such that $c(n_1, n_2) = c(n_i, n_j)\ \forall 1 \leq i < j$.*

We give in Chapter 3 a stochastic generalisation, which includes Theorem 2.1 as a special (deterministic) case. The proof of our more general version may thus be adapted to apply directly to the classical setting, so giving a new proof. Rather than present this 'stripped-down' proof at a later stage, we give it here to aid a better understanding of the proof strategy of Chapter 3 in the more general (so more complex) setting. Our proof here highlights in Lemma 2.3 (see citations in Remark 2) properties arising in the definition of the Ellentuck topology. As a preliminary, we include a standard proof, though not as in [13], for the two-colour case to introduce some 'labelling' terminology.

In what follows, subsets of $\mathbb{N}$ may sometimes usefully be also termed 'subsequences' (of the sequence of natural numbers).

Theorem 2.2. (Two-colour Ramsey Theorem) *Given a set of colours* $C = \{\text{Red}, \text{Blue}\}$, *for every function* $c : \mathbb{N}^{(2)} \to C$, *there is an increasing sequence of integers* $n_1 < n_2 < \ldots$ *such that* $c(n_1, n_2) = c(n_i, n_j)$ $\forall 1 \leq i < j$.

First Proof (Standard Proof). Under the given function c, we shall construct inductively an infinite sequence $\{s_i\}_{i \in \mathbb{N}} \subseteq \mathbb{N}$ and a sequence $\{T_i\}_{i \geq 0}$ of subsets of $\mathbb{N}$ with $T_{i+1} \subset T_i$ $\forall i$. In doing so, we shall label each element in $\{s_i\}_{i \in \mathbb{N}}$ with a symbol r or b. We finally show that any infinite subsequence of $\{s_i\}_{i \in \mathbb{N}}$ with identical label is the desired sequence.

Let $T_0 = \mathbb{N}$ and $s_1 = 1$. We write $R_1 := \{n \in T_0 : c(s_1, n) = \text{Red}\}$ and $B_1 := \{n \in T_0 : c(s_1, n) = \text{Blue}\}$, i.e., $B_1 = T_0 \setminus (\{s_1\} \cup R_1)$. Then at least one of R_1 and B_1 is infinite. If R_1 is infinite, let T_1 be R_1 and label s_1 with r; otherwise let T_1 be B_1 and label s_1 with b. Suppose that we have already obtained the labelled initial subsequence $\{s_1, \ldots, s_k\}$ and a decreasing sequence $\{T_1, \ldots, T_k\}$, i.e., $T_{i+1} \subset T_i$ $\forall 0 < i < k$. The induction assumption is that, for each i with $1 < i \leq k$, $s_i \in T_{i-1}$ and $c(s_i, n)$ is the same colour for all n in T_i. Let s_{k+1} be the minimum element in T_k. Write $R_{k+1} := \{n \in T_k : c(s_{k+1}, n) = \text{Red}\}$ and $B_{k+1} := \{n \in T_k : c(s_{k+1}, n) = \text{Blue}\}$. If R_{k+1} is infinite, let T_{k+1} be R_{k+1} and label s_{k+1} with r; otherwise let T_{k+1} be B_{k+1} and label s_{k+1} with b. Then the new initial subsequence $\{s_1, \ldots, s_{k+1}\}$ and the decreasing sequence $\{T_1, \ldots, T_{k+1}\}$ satisfy the induction assumption. In this way, we obtain a labelled infinite sequence $\{s_i\}_{i \in \mathbb{N}}$. We write $R := \{i \in \mathbb{N} : s_i \text{ labelled with r}\}$ and $B := \{i \in \mathbb{N} : s_i \text{ labelled with b}\}$. For any $i, j \in R$ with $i < j$, $s_j \in T_i$. So $c(s_i, s_j) = \text{Red}$. By similar arguments, $c(s_i, s_j) = \text{Blue}$ $\forall i, j \in B$ with $i < j$. At least one of R and B is infinite, which completes the proof. $\square$

We now prove Theorem 2.2 by a different method, which will shed light on the approach in the proof of the Stochastic Ramsey Theorem introduced in the next chapter. For the preparation of the proof, we analyse the characterisation of one special infinite sequence of natural numbers under a colour function c and obtain a lemma as follows.

Lemma 2.3. *Given a colour function c, there exists an infinite sequence $A = (a_1, a_2, ...) \subseteq \mathbb{N}$ ($a_i < a_j \ \forall i < j$) with a partition $A = R(A) \cup B(A)$ having the following property: for any infinite subsequence $A' \subseteq A$ and any finite subset $S \subset A'$, there exists an infinite subsequence $T = (t_1, t_2, ...) \subset A'$ such that*

$$c(n, t_l) = \text{Red} \ \forall n \in S \cap R(A) \ \forall l > 0 \tag{2.1}$$

and

$$c(n, t_l) = \text{Blue} \ \forall n \in S \cap B(A) \ \forall l > 0. \tag{2.2}$$

Remark 1: The partition above may be degenerate, that is, one of the partitioning sets of A may be empty. This lemma is the deterministic version of Lemma 3.7.

Remark 2: Note that the two conditions of 'S finite' and '$T \subset A$' here are key to the definition of the basic open sets of the Ellentuck topology. (cf. [9])

Proof. We first extract an infinite subsequence $A = \{a_i\}_{i \geq 0}$ of $\mathbb{N}$ by a labelling process, each element being labelled with r or b. We shall then show that the sequence A has the desired property and the two collections of elements labelled with r and b are respectively the two partitioning sets of A, $R(A)$ and $B(A)$.

Step 1: We construct an increasing sequence of finite subsets $\{E_i\}_{i \geq 0}$ of $\mathbb{N}$ with $E_0 = \emptyset$ and $E_i \subseteq E_{i+1} \ \forall i \geq 0$, a sequence of infinite sets $\{G_i\}_{i \geq 0}$ of $\mathbb{N}$ with $G_0 = \mathbb{N}$ and $G_i \supseteq G_{i+1}$, and a partial labelling of each G_i with symbols r and b. In doing so, we pick an infinite sequence $\{a_i\}_{i \geq 0}$ of labelled natural numbers such that $\{a_0, ..., a_i\} \subseteq G_i \ \forall i$, in an inductive procedure.

Set $a_0 = 1$.

Suppose that we have defined $a_0 < ... < a_i$, $\{E_0, ..., E_i\}$ and $\{G_0, ..., G_i\}$. We define E_{i+1} and G_{i+1} by cases.

Case 1. First suppose that for any infinite sub-subsequence $G \subseteq G_i$, there exists an infinite subsequence $T_G = \{t_1, t_2, ...\} \subset G$ with $t_m < t_n \ \forall m < n$ such that

$$c(a_i, t_l) = c(a, t_l) = \text{Red} \ \forall l > 0 \ \forall a \in E_i.$$

Then we label a_i with r, and put $E_{i+1} = E_i \cup \{a_i\}$, $G_{i+1} = G_i$.

Case 2. Now suppose otherwise; then there exists an infinite subsequence, say $G_i(a_i)(E_i)$, of G_i with the following property. Given any $T = \{t_l\}_{l>0} \subset G_i(a_i)(E_i)$ with $c(a, t_l) = \text{Red} \ \forall l > 0 \ \forall a \in E_i$, there exists $N \in \mathbb{N}$ such that

$$c(a_i, t_l) = \text{Blue} \ \forall l \geq N.$$

Indeed, otherwise $c(a_i, t_l) = \text{Red}$ infinitely often, say on $\{t'_l\}_{l \in \mathbb{N}} \subseteq T$; then on $\{t'_l\}_{l \in \mathbb{N}}$ we have

$$c(a_i, t'_l) = c(a, t'_l) = \text{Red} \ \forall l > 0 \ \forall a \in E_i.$$

Choose such a $G_i(a_i)(E_i)$. We let G_{i+1} be $G_i(a_i)(E_i) \cup \{a_0, ..., a_i\}$, so $G_{i+1} \subseteq G_i$; then label a_i with b, and keep $E_{i+1} = E_i$.

We let $a_{i+1} := \min\{n \in G_{i+1} : n > a_i\}$. This completes the induction.

Denote the sequence $\{a_0, a_1, ...\}$ obtained above by A. Hence, A is the infinite sequence all of whose elements are labelled with either r or b. The union of the finite sets $\{E_i\}_{i \geq 0}$ is the union of all numbers labelled with r in A.

Step 2: We denote the two subsets of A labelled with r and b respectively, by S_r and S_b respectively. That is, $S_r = \bigcup_{i \geq 0} E_i$ and $S_b = A \setminus S_r$. We show that the two partitioning sets S_r and S_b of A can fulfill the roles of $R(A)$ and $B(A)$. Suppose $A' = \{a'_1, a'_2, ...\}$ is an infinite subsequence of A and S is a finite subset of A'. We show the conditions (2.1) and (2.2) hold for an infinite subsequence $T' = \{t'_1, t'_2, ...\} \subset A'$, via an intermediate infinite subset $T_{A'} = \{t_1, t_2, ...\} \subset A'$ with the property $T' \subset T_{A'}$. ($T_{A'}$ will take the form T_G when Case 1 above applies.)

We define for a finite S

$$d(S) := \min\{n : S \subseteq \{a_i : 0 \leq i < n\}\}.$$

To obtain the result (2.1) in Lemma 2.3, we consider a fixed S and the finite sequence $E_{d(S)}$ in A. Recall that $E_i = \{a_l : 0 \leq l < i;\ a_l \text{ is labelled with r}\}$. It follows that $S_r \cap S \subseteq E_{d(S)}$.

On the one hand, if $E_{d(S)} \neq \emptyset$, by the definition of label r in respect of the last a in $E_{d(S)}$, for any infinite subsequence in $G_{d(S)}$, and hence in particular for A' (note $A' \subseteq A \subseteq G_{d(S)}$), there exists an infinite subsequence $T_{A'} = \{t_i\}_{i>0} \subset A'$ (as in Case 1 above) such that

$$c(a, t_i) = \text{Red } \forall a \in E_{d(S)}\ \forall i > 0. \tag{2.3}$$

Therefore,

$$c(a, t_i) = \text{Red } \forall a \in S_r \cap S\ \forall i > 0,$$

i.e. condition (2.1) holds.

If on the other hand $E_{d(S)} = \emptyset$, then $S_r \cap S = \emptyset$ and (2.1) will hold vacuously, provided we define $T_{A'}$ to be, say A'. Note that result (2.3) still holds in this case, albeit vacuously.

We now show that the condition (2.2) in Lemma 2.3 also holds by reference to the same infinite sequence $T_{A'}$. Without loss of generality, we suppose $S_b \cap S \neq \emptyset$ and consider an element indexed as a_m. Since a_m is labelled with b, by the definition of G_{m+1}, for any infinite sequence $K = (k_1, k_2, \ldots) \subset G_{m+1}$ with

$$c(a, k_i) = \text{Red } \forall i > 0\ \forall a \in E_m,$$

there exists $N_m \in \mathbb{N}$ such that

$$c(a_m, k_i) = \text{Blue } \forall i \geq N_m.$$

Since $T_{A'} \subset A' \subseteq A \subseteq G_{m+1}$, $E_m \subseteq E_{d(S)}$ and (2.3) holds, we can take $T_{A'} = \{t_i\}_{i>0}$ for K above. Moreover, for any element $a_{m'}$ in $S_b \cap S$, we can likewise take $T_{A'}$ again for K above. Hence, there exists a corresponding $N_{m'} \in \mathbb{N}$ such that (with $t_i = k_i$)

$$c(a_{m'}, t_i) = \text{Blue } \forall i \geq N_{m'}.$$

Because $S_b \cap S$ is finite, we can validly define N to be the maximum of all the numbers $N_{m'}$. Therefore, for $T_{A'} = \{t_i\}_{i \in \mathbb{N}}$ as above

$$c(a, t_i) = \text{Blue } \forall a \in S_b \cap S\ \forall i \geq N.$$

Define a new infinite sequence $T' = \{t'_i\}_{i \in \mathbb{N}} \subset T_{A'}$ such that $t'_i = t_{i+N}$ $\forall i \in \mathbb{N}$. This is the desired infinite subsequence for the set S in the infinite sequence A'. □

Remark: In fact, we do not need the full strength of Lemma 2.3 in the proof of Theorem 2.2: we only need the special cases $A' = R(A)$, or $B(A)$. We prefer to establish a more informative result, mainly for the analogy with the situation in the stochastic Ramsey theorems later.

Proof of Theorem 2.2 from Lemma 2.3. In the infinite sequence A obtained from Lemma 2.3, at least one partitioning set of $R(A)$ and $B(A)$ is infinite. Without loss of generality, assume $R(A)$ is infinite.

Let $n_1 = a_1$. We construct the desired infinite subsequence $\{n_i\}_{i \in \mathbb{N}} \subseteq R(A) \subseteq \mathbb{N}$ inductively. Suppose that we have already obtained the initial subsequence $(n_1, ..., n_k)$. Since $(n_1, ..., n_k) \subset R(A)$, by (2.1) in Lemma 2.3, there exists an infinite subsequence $T^k = \{t_i^k\}_{i \in \mathbb{N}} \subset R(A)$ with the property that

$$c(a, t_i^k) = \text{Red } \ \forall a \in \{n_1, ..., n_k\} \ \forall i > 0.$$

Denote n_{k+1} to be t_1^k. In this way, we find an infinite sequence $\{n_i\}_{i \in \mathbb{N}} \subseteq R(A) \subseteq \mathbb{N}$ where

$$c(n_i, n_j) = \text{Red } \ \forall i < j.$$

Similarly, if $B(A)$ is infinite, we can find an infinite sequence $\{n_i\}_{i \in \mathbb{N}} \subseteq B(A) \subseteq \mathbb{N}$ where

$$c(n_i, n_j) = \text{Blue } \ \forall i < j.$$

□

For completeness, we include the standard derivation of Theorem 2.1 from Theorem 2.2.

Proof of Theorem 2.1. We prove this by induction on the number of colours. Recall that $\mathbb{N}^{(2)} = \{(m, n) : m < n; m, n \in \mathbb{N}\}$. Suppose that given the natural number $k > 2$, for every function that assigns a colour $c : \mathbb{N}^{(2)} \to \{c_1, ..., c_{k-1}\}$, there is a sequence of integers $n_1 < n_2 < ...$ such that $c(n_1, n_2) = c(n_i, n_j)$ $\forall 1 \leq i < j$. For any function $c : \mathbb{N}^{(2)} \to C = \{c_1, ..., c_k\}$, we define a new function $c' : \mathbb{N}^{(2)} \to \{c_1, ..., c_{k-2}, \bar{c}\}$ such that for any $(m, n) \in \mathbb{N}^{(2)}$

$$c'(m,n) := \begin{cases} c(m,n) & \text{if } c(m,n) \in \{c_1, ..., c_{k-2}\} \\ \bar{c} & \text{if } c(m,n) \in \{c_{k-1}, c_k\}. \end{cases}$$

Then by induction assumption, there is an increasing sequence of integers $n_1 < n_2 < ...$ such that $c'(n_i, n_j)$ has the same colour $\tilde{c} \in \{c_1, ..., c_{k-2}, \bar{c}\}$ for all $(i, j) \in \mathbb{N}^{(2)}$. If $\tilde{c} \in \{c_1, ..., c_{k-2}\}$, then $c(n_i, n_j) = \tilde{c}\ \forall (i, j) \in \mathbb{N}^{(2)}$. If $\tilde{c} = \bar{c}$, then by Lemma 2.2, there is an infinite subsequence $\{n'_1, n'_2, ...\} \subset \{n_1, n_2, ...\}$ such that $c(n'_1, n'_2) = c(n'_i, n'_j)\ \forall (i, j) \in \mathbb{N}^{(2)}$. In either case, we have obtained an infinite sequence $\{s_1, s_2, ...\} \subset \mathbb{N}$ such that $c(s_1, s_2) = c(s_i, s_j)\ \forall (i, j) \in \mathbb{N}^{(2)}$, which completes the proof. □

We can go further to deduce the finitary version of the Ramsey theorem, which says that, given a set of finitely many colours, no matter how large a target size for a monochromatic subset, there exists a bound such that under whatever colour function defined on $\mathbb{N}^{(2)}$ we can find a monochromatic subset of the target cardinality within that bound. (See [7] for reference.)

Theorem 2.4. (Finitary version) *Given a set of colours $C = \{c_1, ..., c_k\}$ and the natural number m, there exists $\bar{n} = n(m) \in \mathbb{N}$ such that for every function $c : \mathbb{N}^{(2)} \to C$, there is a sequence of integers $n_1 < n_2 < ... < n_m \leq \bar{n}$ such that $c(n_1, n_2) = c(n_i, n_j)\ \forall 1 \leq i < j \leq m$.*

Proof. The set of all colourings f is $[C]^{\mathbb{N}^{(2)}}$: the finite set $[C]$ is given the discrete topology, and $[C]^{\mathbb{N}^{(2)}}$ the product topology. By Tychonov's theorem, this space is compact. For each subset in $\mathbb{N}$ of size m, say $Q = \{n_1, n_2, ...n_m\}$ where $n_i < n_j\ \forall 1 \leq i < j \leq m$, the set C_Q of colourings in $[C]^{\mathbb{N}^{(2)}}$ in which

$$c(n_1, n_2) = c(n_i, n_j)\ \forall 1 \leq i < j \leq m$$

is an open set (by definition of the product topology). The Infinitary Ramsey Theorem (Theorem 2.1) asserts that these sets cover the whole of $[C]^{\mathbb{N}^{(2)}}$. By compactness, some finite collection $\{C_{Q_1}, ..., C_{Q_t}\}$ of these sets also covers $[C]^{\mathbb{N}^{(2)}}$. This implies that every colouring on $\bigcup_{i=1}^{t} Q_i$ will have

$$c(n_1, n_2) = c(n_i, n_j)\ \forall 1 \leq i < j \leq m,$$

on at least one of the Q_i. Then we may take for $\bar{n}$ any bound of $\bigcup_{i=1}^{t} Q_i$ such that $\forall 1 \leq l \leq t \; \forall n_m \in Q_l$, $n_m < \bar{n}$. $\square$

Remark: There are alternatives to this proof: for instance, one can reorganise the proof using 'sequential compactness' or apply the 'compactness theorem for first-order logic'. (cf. [7])

3

The main stochastic theorem

3.1 Examples

We first give several examples of stochastic colouring all based on the same probability space $(\Omega, \mathcal{F}_\infty, P)$ which we now define. Referring to Section 1.1, let $J_n = \{\text{Head}, \text{Tail}\}$ $\forall n \in \mathbb{N}$. Let P be the usual product probability measure associated with elements of J_n being given equal measure. So, for any partial history $q \in \Omega_n$,

$$P\big(\omega : \omega|n = q \text{ and } \omega(n+1) = \text{Head}\,|B(q)\big) = 1/2,$$

and

$$P\big(\omega : \omega|n = q \text{ and } \omega(n+1) = \text{Tail}\,|B(q)\big) = 1/2.$$

We can view this model as tossing a fair coin infinitely many times, and visualise $\Omega_{<\infty}$ as generating a binary tree $\mathbb{F}$. A partial history is a finite sequence of Head and Tail; a bounded stopping time is an instruction of when to stop referring only to the revealed partial history. An example of what is not a bounded stopping time is the instruction to stop at the first time when you get a Head. That is because the stopping time would be infinity if the randomly chosen history turned out to be Tails forever. Below is an example of an increasing sequence of bounded stopping times.

$\theta_1 = \min\{5,$ the first time Head shows up$\}$,

$\theta_2 = \min\{10,$ the third time Head shows up$\}$,

$\theta_3 = \min\{12,$ the fourth time Head shows up$\}$.

Examples of stochastic colouring are as follows.

1. Let $C = \{\text{Red}, \text{Blue}\}$. For any partial history $q = (q(1), ..., q(n), q(n+1) = \text{Head})$, let $f(q, \tau) = \text{Red}\ \forall \tau \in \mathcal{T}(q)$; for any partial history $q = (q(1), ..., q(n), q(n+1) = \text{Tail})$, let $f(q, \tau) = \text{Blue}\ \forall \tau \in \mathcal{T}(q)$. That is, the value of $f(q, \tau)$ is determined by the last coordinate of the partial history q only.

   This example is similar to the one given by Shmaya and Solan in [16], which shows that, in the stochastic Ramsey theorem, the condition $\epsilon > 0$ is indispensable. Consider any three bounded stopping times $\theta_1 < \theta_2 < \theta_3$ in which $\theta_2 \in \mathcal{T}_N$ for some N. Note that there exists a history ω such that the last coordinates of the two partial histories of ω restricted to θ_1 and θ_2 are different, i.e., $\omega|\theta_1(\theta_1(\omega)) \neq \omega|\theta_2(\theta_2(\omega))$. To be specific, for the history $\omega_1 = (\omega_1(1), \omega_1(2), ...)$ in which $\omega_1(k) = \text{Head}\ \forall k \in \mathrm{N}$, it follows that $f(\omega_1|\theta_1, \theta_2) = f(\omega_1|\theta_1, \theta_3) = \text{Red}$. Now consider $\omega_2 = (\omega_2(1), \omega_2(2), ...)$ in which $\omega_2(k) = \text{Head}\ \forall 1 \leq k \leq \|\omega_1|\theta_1\|$ and $\omega_2(k) = \text{Tail}\ \forall k > \|\omega_1|\theta_1\|$. It follows that $f(\omega_2|\theta_2, \theta_3) = \text{Blue}$, and $P(B(\omega_2|\theta_2)) \geq 1/2^N$. Note that $B(\omega_2|\theta_1) \supset B(\omega_2|\theta_2)$, from which we may infer that $P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_1,\theta_3} = \hat{f}_{\theta_2,\theta_3}) \leq 1 - 1/2^N$.

2. Let $C = \{\text{Red}, \text{Blue}\}$. Given any function $g : \Omega_{<\infty} \to C$, define a stochastic colouring f such that, for every partial history q, $f(q, \theta) = g(q)\ \forall \theta \in \mathcal{T}(q)$. Thus, the stochastic colouring only depends on the partial history, but it is still a more general stochastic colouring than the one in the preceding example. For convenience, with this f, we *label* a partial history q with r if $f(q, \theta) = \text{Red}\ \forall \theta \in \mathcal{T}(q)$, and with b if $f(q, \theta) = \text{Blue}\ \forall \theta \in \mathcal{T}(q)$. If the Stochastic Infinitary Ramsey Theorem (Theorem 1.2) is to be true in this context, then we should be able to find a partition $\Omega = R \cup B$ with $R, B \in \mathcal{F}_N$ for some N

in $\mathbb{N}$, and an infinite increasing sequence of bounded stopping times $N \leq \theta_1 < \theta_2 < \ldots$ such that

$$P(\omega \in \Omega : \omega|\theta_i \text{ is labelled with r } \forall i \in \mathbb{N}|R) > 1 - \epsilon$$

and

$$P(\omega \in \Omega : \omega|\theta_i \text{ is labelled with b } \forall i \in \mathbb{N}|B) > 1 - \epsilon.$$

To find a sequence $\{\theta_i\}$ with this property, recall in measure theory that a set $S_R = \{\omega : \omega|n \text{ is labelled with r infinitely often}\}$ is in $\mathcal{F}_\infty$ and can be approximated in probability by a Borel set R in $\bigcup_n \mathcal{F}_n$. This means that in the set R, with very large probability (i.e., $> 1 - \epsilon$), we will come across infinitely many partial histories labelled with r along a randomly selected history. If R is ϵ-maximal, then in its complement one may similarly find a Borel set B in which partial histories are labelled with b eventually. We will find an increasing sequence of natural numbers $\{N_i\}$ from which we can then define θ_i simply for each i to be the bounded stopping time which stops at N_i unless we come across the i^{th} partial history in R labelled with r or the i^{th} partial history in B labelled with b. More detailed and rigorous analysis will appear in Lemma 3.8.

3. Let $C = \{\text{Red}, \text{Blue}, \text{Green}\}$. For each pair (q, τ) in Z, i.e., $q \in \Omega_{<\infty}$ and $\tau \in \mathcal{T}(q)$, we denote $P(\{\omega \in \Omega : \omega(\tau(\omega)) = \text{Head}\}|B(q))$ by $P_{q,\tau}(\text{Head})$ and $P(\{\omega \in \Omega : \omega(\tau(\omega)) = \text{Tail}\}|B(q))$ by $P_{q,\tau}(\text{Tail})$. Let

$$f(q,\tau) := \begin{cases} \text{Red} & \text{if } P_{q,\tau}(\text{Head}) > P_{q,\tau}(\text{Tail}); \\ \text{Blue} & \text{if } P_{q,\tau}(\text{Head}) < P_{q,\tau}(\text{Tail}); \\ \text{Green} & \text{if } P_{q,\tau}(\text{Head}) = P_{q,\tau}(\text{Tail}). \end{cases}$$

4. Let $C = \{\text{Red}, \text{Blue}\}$. For a pair (q, τ) in Z, we say that τ is *uniform under* q if there exists n in $\mathbb{N}$ such that $\tau(\omega) = n$ for all ω in $B(q)$. For each pair (q, τ) in Z, let

$$f(q,\tau) := \begin{cases} \text{Red} & \text{if } \tau \text{ is uniform under } q; \\ \text{Blue} & \text{otherwise.} \end{cases}$$

5. We can release the constraint that f is defined conditional on $B(q)$ in both example 2 and 3 above. For instance, in example 2, given a partial history q in $\Omega_{<\infty}$, we can simply define $P_{q,\tau}(\text{Head}) := P(\{\omega \in \Omega : \omega(\tau(\omega)) = \text{Head}\})$ and $P_{q,\tau}(\text{Tail})$ similarly, for all $\tau \in \mathcal{T}(q)$.

6. Let $C = \{\text{Red}, \text{Blue}\}$. For any bounded stopping time $\tau \in \mathcal{T}$, define b_τ to be the minimum bound of τ such that $\tau \in \mathcal{T}_{b_\tau}$ but $\tau \notin \mathcal{T}_{b_\tau - 1}$. For each pair (q, τ) in Z, let
$$f(q,\tau) := \begin{cases} \text{Red} & \text{if } b_\tau \text{ is even;} \\ \text{Blue} & \text{if } b_\tau \text{ is odd.} \end{cases}$$

See Chapter 5 for an application of the stochastic Ramsey theorem in expected utility theory.

We show two mappings below that are *not* stochastic colourings. Let $C = \{\text{Red}, \text{Blue}\}$ and $\Omega = J^{\mathbf{N}}$ for a finite J in both two examples.

1. For any partial history $q = (q(1), ..., q(n_q))$, we say a partial history $h = (h(1), ..., h(n_q), h(n_q + 1), ..., h(n_h))$ is *consistent with* q if $h(i) = q(i)\ \forall 1 \leq i \leq n_q$. That is, for any ω in $B(h)$, $\omega|n_q = q$ and $\omega|n_h = h$, with $n_h > n_q$. Denote the set $\{(q,h) : q \in \Omega_{<\infty};\ h \text{ is consistent with } q\}$ by X. We define a mapping $g : X \to C$ and an induced mapping $\hat{g} : \mathcal{T}^{(2)} \times \Omega \to C$ by
$$\hat{g}_{\sigma,\tau}(\omega) := g(\omega|\sigma, \omega|\tau)\ \forall(\sigma,\tau) \in \mathcal{T}^{(2)}\ \forall \omega \in \Omega.$$

2. For any bounded stopping time $\tau \in \mathcal{T}$, we say a partial history q is *unrevealed by* τ if $\|q\| > \tau(\omega)\ \forall \omega \in B(q)$. Denote the set $\{(\tau, q) : \tau \in \mathcal{T};\ q \text{ is unrevealed by } \tau\}$ by Y. We define a mapping $\gamma : Y \to C$ and an induced mapping $\hat{\gamma} : \mathcal{T}^{(2)} \times \Omega \to C$ by
$$\hat{\gamma}_{\sigma,\tau}(\omega) := g(\sigma, \omega|\tau)\ \forall(\sigma,\tau) \in \mathcal{T}^{(2)}\ \forall \omega \in \Omega.$$

Recall that for any $(\sigma, \tau) \in \mathcal{T}^{(2)}$ the induced stochastic colouring $\hat{f}_{\sigma,\tau}(\omega) = f(\omega|\sigma, \tau)$ is determined at stage $\omega|\sigma$. In examples above, both the formulations of $\hat{g}_{\sigma,\tau}(\omega)$ and $\hat{\gamma}_{\sigma,\tau}(\omega)$ include the partial history $\omega|\tau$, which means one has to determine $\hat{g}_{\sigma,\tau}(\omega)$ and $\hat{\gamma}_{\sigma,\tau}(\omega)$ at a later stage $\omega|\tau$. More rigorously $\hat{f}_{\sigma,\tau}$ is $\mathcal{F}_\sigma$-measurable but both $\hat{g}_{\sigma,\tau}$ and $\hat{\gamma}_{\sigma,\tau}$ are only $\mathcal{F}_\tau$-measurable.

3.2 Outline of the proof

It is clear that a proof similar to the standard proof of Lemma 2.2 is not enough for the two-colour case in the Stochastic Infinitary Ramsey Theorem (Theorem 1.2), as $\hat{f}_{\theta_i,\theta_j}$ is randomly selected for each pair (i, j). We defer its proof to Section 3.5 and firstly focus on the following theorem, which contains the essence of the argument.

Theorem 3.1. (Finite-state Two-colour Stochastic Partition Theorem). *For a set $C = \{c_1, c_2\}$, a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence of sets of finitely many states, and a stochastic colouring f with values in C, for every $\epsilon > 0$, there exists a natural number N, two sets $S_1, S_2 \in \mathcal{F}_N$ with $S_1 \cup S_2 = \Omega$, and a sequence of bounded stopping times $\theta_1 < \theta_2 < \ldots$ such that*

$$P(\hat{f}_{\theta_i,\theta_j} = c_1,\ \forall 1 \leq i < j | S_1) > 1 - \epsilon, \quad \textit{if } P(S_1) > 0,$$

and

$$P(\hat{f}_{\theta_i,\theta_j} = c_2,\ \forall 1 \leq i < j | S_2) > 1 - \epsilon, \quad \textit{if } P(S_2) > 0.$$

The main idea in the proof of Theorem 3.1 is to view $\Omega_{<\infty}$ as a tree. (See [10] for an account of probability theory on trees.) In this tree setting, which we develop in Section 3.3, we can define new finer filtrations by modifications (pruning) of the tree structure. Our approach can be interpreted, but only as a matter of convenience, in the language of *Markov chains*, via the projection process $X_n(\omega) := \omega|n$. (For other aspects of Markov chains and Ramsey theory see e.g. [6].)

After Theorem 3.1 proved in Section 3.4, the remaining steps to prove Theorem 1.2 are straightforward. We shall give a brief proof of a generalised version of Theorem 3.1 for any set C of finite colours in the latter part of Section 3.4. In Section 3.5, we consider the case that some J_i in $\{J_n\}_{n\in\mathbf{N}}$ are countably infinite sets: we sacrifice a very small probability relative to ϵ, and ignore all but finitely many states in each J_i. This method reduces the problem of countably many states to a problem of finitely many states.

Similarly to Ramsey's theorem, a finite version of Theorem 1.2 can be proved by compactness arguments. This is done in Section 3.6. Surprisingly, in the case that J_i is a finite set for every i, we obtain in Theorem 3.2 below a strong finite version of Theorem 1.2 for all stochastic colourings f and all probability measures P defined on the measurable space; that is, $n(m, \epsilon)$ mentioned below does not depend on f or P.

Theorem 3.2. (Strong Stochastic Finitary Ramsey Theorem). *For a measurable space $(\Omega, \mathcal{F}_\infty)$ constructed from a sequence of sets each containing finitely many states, a set C of finitely many colours, a natural number $m \geq 2$ and any $\epsilon > 0$, there exists $n = n(m, \epsilon) \in \mathbb{N}$ such that, for every probability measure P defined on $(\Omega, \mathcal{F}_\infty)$ and every stochastic colouring f with values in C, there exist m bounded stopping times $\theta_1 < \theta_2 < \ldots < \theta_m \leq n$ with*

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j}\ \forall 1 \leq i < j \leq m) > 1 - \epsilon.$$

3.3 The model

The set $\Omega_{<\infty}$ ordered by sequence extension forms a directed tree $\mathbb{F}$. (This is a graph theoretic tree. We regard it as directed downwards and call it the *tree of partial histories.* For convenience, denote the vertex set of $\mathbb{F}$ by $V_\mathbb{F}$, i.e. $V_\mathbb{F} = \Omega_{<\infty}$. A vertex in $\mathbb{F}$ is a partial history, and in particular, the root of $\mathbb{F}$ is the empty partial history. In the tree $\mathbb{F}$, we see that a directed edge is an extension of a partial history, while an infinite directed path from the root of $\mathbb{F}$ is a history. Define a *covering set* $\bar{V}$ in $\mathbb{F}$ to be a subset of $V_\mathbb{F}$ such that any path in $\mathbb{F}$ goes through infinitely many vertices in $\bar{V}$. Hence, at each vertex $q \in \bar{V}$, along any path ω with $\omega \in B(q)$, there exists a unique vertex $q^+_{\bar{V}}(\omega)$ which is the first vertex in $\bar{V}$ gone through by path ω beyond q. Define a *covering subforest* of $\mathbb{F}$ generated from a covering set $\bar{V}$ in $\mathbb{F}$ to be a forest with vertex set $\bar{V}$ and directed edge set $\{(q, q^+_{\bar{V}}(\omega)) : q \in \bar{V}, \omega \in B(q)\}$. For any covering subforest $\mathbb{G}$ of $\mathbb{F}$, we denote its vertex set by $V_\mathbb{G}$. Any covering set $\bar{V}$ in a covering subforest $\mathbb{G}$ determines a covering subforest of the covering subforest $\mathbb{G}$. Note that a covering subforest of $\mathbb{F}$ is a subtree of $\mathbb{F}$ if and only if $\bar{V}$ includes the root of $\mathbb{F}$.

Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$, we say that a stopping time θ is *adapted* to $\mathbb{G}$ if $\omega|\theta \in V_{\mathbb{G}}\ \forall \omega \in \Omega$. Given a stopping time θ adapted to $\mathbb{G}$, we denote the set of vertices $\{\omega|\theta : \omega \in \Omega\}$ by $S_\theta(\mathbb{G})$, and call it the set of stopping places of θ in $\mathbb{G}$, so extending the notation in Section 1.1 (by emphasising restriction to certain subforest). Intuitively, a stopping time θ is adapted to $\mathbb{G}$ if, according to θ, we always stop at a partial history which is a vertex in $\mathbb{G}$, and $S_\theta(\mathbb{G})$ is the set of all such vertices determined by θ in $\mathbb{G}$.

The definition of a covering subforest is closely related to bounded stopping times. Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$, for any $n \geq 0$, define $\sigma_n(\mathbb{G})$ to be a bounded stopping time adapted to $\mathbb{G}$ such that, for any $\omega \in \Omega$, $\omega|\sigma_n(\mathbb{G})$ is the $(n+1)^{\text{st}}$ vertex in $\mathbb{G}$ along the directed path ω in $\mathbb{F}$. Denote $S_{\sigma_n(\mathbb{G})}(\mathbb{G})$ by $L_n(\mathbb{G})$, and call it the n^{th} *level set* of $\mathbb{G}$. We see that $V_{\mathbb{G}} = \bigcup_{n\geq 0} L_n(\mathbb{G})$, from which we may infer that the sequence $\{L_n(\mathbb{G})\}$ generates the covering subforest $\mathbb{G}$ of $\mathbb{F}$. The following lemma shows a way of constructing a covering subforest of $\mathbb{F}$ from a sequence of covering subforests of $\mathbb{F}$. We call this a fusion lemma by analogy with set theory usage (see e.g. [8], Chapter 15).

Lemma 3.3. (Fusion Lemma). *Given a sequence of covering subforests* $\{\mathbb{G}_n\}$ *of* $\mathbb{F}$ *with* $V_{\mathbb{G}_n} \supseteq V_{\mathbb{G}_{n+1}}\ \forall n \geq 0$, *the sequence* $\{L_n(\mathbb{G}_n)\}_{n\geq 0}$ *generates a covering subforest* $\mathbb{G}'$ *of* $\mathbb{F}$.

Proof. Since $V_{\mathbb{G}_n} \supseteq V_{\mathbb{G}_{n+1}}$, $\sigma_n(\mathbb{G}_n) < \sigma_{n+1}(\mathbb{G}_n) \leq \sigma_{n+1}(\mathbb{G}_{n+1})$. That implies $L_i(\mathbb{G}_i) \cap L_j(\mathbb{G}_j) = \emptyset\ \forall i \neq j$. It then follows immediately that any path ω in $\mathbb{F}$ goes through infinitely many vertices in $V_{\mathbb{G}'} := \bigcup_{n\geq 0} L_n(\mathbb{G}_n)$. Thus, $V_{\mathbb{G}'}$ is a covering set in $\mathbb{F}$, which completes the proof. □

Let $\mathbb{G}$ be a covering subforest of $\mathbb{F}$. At each $n \geq 0$, define the σ-algebra generated by $L_n(\mathbb{G})$ to be

$$\mathcal{G}_n := \sigma\{B(q) : q \in L_n(\mathbb{G})\}.$$

If $\mathbb{G} = \mathbb{F}$, then $\mathcal{G}_n$ is exactly $\mathcal{F}_n$ for every n.

Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$ generated from $\bar{V}$ and a covering set S in $\mathbb{G}$, we say that S generates a covering subforest $\mathbb{G}'$ *pruned below* $\sigma_i(\mathbb{G})$ if $S \supset \bigcup_{0 \le n \le i} L_n(\mathbb{G})$. (The n^{th} level sets of $\mathbb{G}$ and $\mathbb{G}'$ are the same for each n with $0 \le n \le i$.) If a vertex q is in $L_i(\mathbb{G})$, a covering subforest of $\mathbb{G}$ pruned below $\sigma_i(\mathbb{G})$ is also called a covering subforest of $\mathbb{G}$ pruned below q's level. Note that this last definition is meant to be a quick way of saying that pruning occurs below all the vertices in $\mathbb{G}$ of the same level as q.

3.4 The stochastic Ramsey theorem for sets of finitely many states

In this section we prove the Finite-state Two-colour Stochastic Partition Theorem (Theorem 3.1) and its generalisation (Theorem 3.9). We begin with a simple but important observation. In the case that J_i is a set of finitely many states for each i, for any covering subforest $\mathbb{G}$ of $\mathbb{F}$ and any bounded stopping time θ adapted to $\mathbb{G}$, the set of stopping places of θ in $\mathbb{G}$ introduced in Section 3.3, i.e. $S_\theta(\mathbb{G})$, is finite.

To prove Theorem 3.1, we shall firstly define and find a *well-structured subtree* $\mathbb{A}$ of $\mathbb{F}$ under the stochastic colouring f. Then we shall extract a covering subforest $\bar{\mathbb{F}}$ of $\mathbb{A}$, which has a 'nice' partition of its vertex set $V_{\bar{\mathbb{F}}}$. Finally, we shall show that there exists a sequence of bounded stopping times $\theta_1 < \theta_2 < ...$ adapted to $\bar{\mathbb{F}}$ which satisfies Theorem 3.1.

We write Red and Blue for c_1 and c_2. Denote $\{0, 1, 2, ...\}$ by $\mathbb{Z}^+$. Now we introduce some preliminary technical definitions and results.

Definition 3.4. *A well-structured subtree* $\mathbb{A}$ *under the stochastic colouring* f *is a covering subtree of* $\mathbb{F}$ *with a partition of its vertex set* $V_{\mathbb{A}} = R(\mathbb{A}) \cup B(\mathbb{A})$ *such that for any covering subforest* $\mathbb{G}$ *of* $\mathbb{A}$ *and any finite subset* $S \subset V_{\mathbb{G}}$, *there exists an infinite sequence of bounded stopping times* $\delta_1 < \delta_2 < ...$ *adapted to* $\mathbb{G}$ *with the property*

$$f(q, \delta_l) = \text{Red} \;\; \forall q \in S \cap R(\mathbb{A}) \; \forall l > 0$$

and

$$f(q, \delta_l) = \text{Blue} \;\; \forall q \in S \cap B(\mathbb{A}) \; \forall l > 0.$$

Remark: The partition above may be degenerate, that is, one of the partitioning sets of $V_{\mathbb{A}}$ may be empty.

We intend to find a well-structured subtree $\mathbb{A}$ of $\mathbb{F}$ under f by breadth-first search. When we define inductively a sequence of vertices $\{q_i\}$, each vertex in it will be labelled with a symbol r or b. We will show that the sequence $\{q_i\}$ generates a well-structured subtree $\mathbb{A}$ of $\mathbb{F}$ under f and the two collections of vertices labelled with r and b are respectively the two partitioning sets of vertices, $R(\mathbb{A})$ and $B(\mathbb{A})$. In doing so, we will need the following *marking scheme* and a related lemma.

Definition 3.5. *Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$ and a finite set M of vertices with $M \subseteq V_{\mathbb{G}}$, for every vertex q in $V_{\mathbb{G}} \setminus M$, mark q with symbol r relative to M in $\mathbb{G}$, if for any covering subforest $\mathbb{G}'$ of $\mathbb{G}$, there exists an infinite sequence of bounded stopping times $\delta_1 < \delta_2 < \delta_3 < ...$ adapted to $\mathbb{G}'$ such that*

$$f(q, \delta_l) = f(\bar{q}, \delta_l) = \text{Red} \ \ \forall l > 0 \ \forall \bar{q} \in M.$$

Mark q with symbol b relative to M in $\mathbb{G}$ if q cannot be marked with r relative to M in $\mathbb{G}$.

Note that in the condition 'for any covering subforest $\mathbb{G}'$ of $\mathbb{G}$' above we do not need either $q \in V_{\mathbb{G}'}$ or $M \subset V_{\mathbb{G}'}$.

As a consequence of the definition above, a vertex q is marked with b relative to M in $\mathbb{G}$ in either of the following circumstances.

I. There exists a covering subforest $\bar{\mathbb{G}}$ of $\mathbb{G}$ such that no infinite sequence of bounded stopping times $\delta_1 < \delta_2 < \delta_3 < ...$ adapted to $\bar{\mathbb{G}}$ has the property that $f(\bar{q}, \delta_l) = \text{Red} \ \forall l > 0 \ \forall \bar{q} \in M$.

II. There exists a covering subforest $\mathbb{G}'$ of $\mathbb{G}$ such that, for any sequence of bounded stopping times $\delta_1 < \delta_2 < \delta_3 < ...$ adapted to $\mathbb{G}'$ which satisfies $f(\bar{q}, \delta_l) = \text{Red} \ \forall l > 0 \ \forall \bar{q} \in M$, there exists $N \in \mathbb{N}$ such that $f(q, \delta_i) = \text{Blue} \ \forall i \geq N$.

For any such $\mathbb{G}'$ in the condition II, we say that q is marked with b relative to M *with witness* $\mathbb{G}'$ in $\mathbb{G}$. ($\mathbb{G}'$ is a witness to the condition II.) Note that the finite set M can be empty. In such a case, the condition II simplifies down to the existence of a covering subforest $\mathbb{G}'$ of $\mathbb{G}$ such that, for any sequence of bounded stopping times $\delta_1 < \delta_2 < \delta_3 < ...$ adapted to $\mathbb{G}'$, there exists $N \in \mathbb{N}$ such that $f(q, \delta_i) = \text{Blue } \forall i \geq N$.

Remark: In the proof of Lemma 3.7 later, the fact that a vertex q is labelled with b always follows from the condition II above.

Lemma 3.6. *If a vertex q is marked with b relative to M with witness $\mathbb{G}'$ in $\mathbb{G}$, then there exists a covering subforest $\mathbb{G}(q)(M)$ of $\mathbb{G}$ pruned below q's level such that q is marked with b relative to M with witness $\mathbb{G}(q)(M)$ in $\mathbb{G}$.*

Proof. Suppose $q \in L_i(\mathbb{G})$ for some i. Consider a new covering subforest $\bar{\mathbb{G}}$ of $\mathbb{G}$ generated from a covering set $(\bigcup_{0 \leq n \leq i} L_n(\mathbb{G})) \cup V_{\mathbb{G}'}$, which means that $\bar{\mathbb{G}}$ is a covering subforest of $\mathbb{G}$ pruned below q's level. Because we are only concerned with the existence of an infinite sequence of bounded stopping times, and $L_n(\bar{\mathbb{G}}) = L_n(\mathbb{G}') \; \forall n > i$, q is also marked with b relative to M with witness $\bar{\mathbb{G}}$ in $\mathbb{G}$. Hence $\bar{\mathbb{G}}$ satisfies to be one desired $\mathbb{G}(q)(M)$. □

We use the word 'marking' only in the sense above, in distinction to 'labelling' used as a generic term.

Lemma 3.7. *Given a stochastic colouring f, there exists in the tree of histories $\mathbb{F}$ a well-structured subtree $\mathbb{A}$ under f.*

Proof. We first extract a subtree $\mathbb{A}$ of $\mathbb{F}$ by a labelling process, then prove this subtree $\mathbb{A}$ is a well-structured subtree of $\mathbb{F}$ under f.

Step 1. We construct an increasing sequence of collections of vertices $\{E_i\}$ with $E_0 = \emptyset$ and $E_i \subseteq E_{i+1} \; \forall i \geq 0$, a sequence of covering subforests $\{\mathbb{G}_i\}$ of $\mathbb{F}$ with $\mathbb{G}_0 = \mathbb{F}$ and $V_{\mathbb{G}_i} \supseteq V_{\mathbb{G}_{i+1}} \; \forall i \geq 0$, and a partial labelling of $\mathbb{G}_i$ with symbols r and b. We shall see that E_i for each i is the set of vertices having been labelled with r at induction stage i. The induction hypothesis is that given the current covering subforest $\mathbb{G}_i$ of $\mathbb{F}$, for every covering subforest $\mathbb{G}$ of

$\mathbb{G}_i$, there exists an infinite sequence of bounded stopping times $\delta_1 < \delta_2 < ...$ adapted to $\mathbb{G}$ such that $f(q, \delta_l) = \text{Red}$ $\forall q \in E_i$ $\forall l$. Given E_i and $\mathbb{G}_i$, we will check one vertex q_i which has not been labelled in $\mathbb{G}_i$ to define E_{i+1} and $\mathbb{G}_{i+1}$. If q_i is marked with r relative to E_i in $\mathbb{G}_i$, we will let $E_{i+1} = E_i \cup \{q_i\}$, $\mathbb{G}_{i+1} = \mathbb{G}_i$, and label q_i with r. Otherwise, by inductive hypothesis, we will see that q is marked with b relative to E_i in $\mathbb{G}_i$ by the condition II below Definition 3.5. It follows from Lemma 3.6 that there exists a covering subforest $\mathbb{G}_i(q_i)(E_i)$ of $\mathbb{G}_i$ pruned below q_i's level. Let $\mathbb{G}_{i+1}$ be this $\mathbb{G}_i(q_i)(E_i)$, let $E_{i+1} = E_i$, and label q_i with b.

We now explain how to determine the vertices q_i which are to be checked. Denote the root of $\mathbb{F}$ by q_0. After obtaining E_1 and $\mathbb{G}_1$ from q_0, E_0 and $\mathbb{G}_0$, we enumerate all vertices in $L_1(\mathbb{G}_1)$. Note as before that $L_1(\mathbb{G}_1)$ is finite. Suppose $|L_1(\mathbb{G}_1)| = l_1$; enumerate the elements of $L_1(\mathbb{G}_1)$ as $q_1, q_2, ...q_{l_1}$. Applying the above process to $q_1, q_2, ...q_{l_1}$ in turn, we obtain two finite sequences $\{E_i\}$ and $\{\mathbb{G}_i\}$ in which $1 < i \leq l_1 + 1$. Note that $L_1(\mathbb{G}_1) = L_1(\mathbb{G}_{i+1})$ $\forall 1 \leq i \leq l_1$, since if q_i is labelled with b, $\mathbb{G}_{i+1}$ is a covering subforest of $\mathbb{G}_i$ pruned below $\sigma_1(\mathbb{G}_i)$. It follows that all q_1 to q_{l_1} indexed in $\mathbb{G}_1$ are in $L_1(\mathbb{G}_i)$ for all $0 < i \leq l_1 + 1$. We then enumerate all vertices in $L_2(\mathbb{G}_{l_1+1})$ as $q_{l_1+1}, q_{l_1+2}, ..., q_{l_1+l_2}$. By applying the process to $q_{l_1+1}, q_{l_1+2}, ..., q_{l_1+l_2}$, we obtain two finite sequences $\{E_i\}$ and $\{\mathbb{G}_i\}$ in which $l_1 + 1 < i \leq l_1 + l_2 + 1$. We can then define all $L_{i+1}(\mathbb{G}_{1+\sum_{m\leq i} l_m})$ for all i in a similar way.

Let $l_0 = 1$. Define $z(k) = \sum_{0\leq i\leq k} l_i$. Given k, it follows that $L_n(\mathbb{G}_{z(k)}) = L_n(\mathbb{G}_i)$ $\forall 0 \leq n \leq k$ $\forall i > z(k)$, and in any $\mathbb{G}_i$ with $i \geq z(k)$, all vertices in $\bigcup_{0\leq n\leq k} L_n(\mathbb{G}_i)$ have already been labelled. Note that $V_{\mathbb{G}_{z(k)}} \supseteq V_{\mathbb{G}_{z(k+1)}}$ $\forall k \geq 0$. So, by Fusion Lemma (Lemma 3.3), the sequence $\{L_k(\mathbb{G}_{z(k)})\}_{k\geq 0}$ generates a covering subforest of $\mathbb{F}$. Denote this covering subforest by $\mathbb{A}$. Note that since $L_0(\mathbb{G}_{z(0)})$ only contains the root of $\mathbb{F}$, $\mathbb{A}$ is actually a covering subtree of $\mathbb{F}$. From the definition of $z(k)$, $\mathbb{A}$ is the covering subtree that contains exactly the vertices labelled with r or b. The union of the collections $\{E_i : i \in \mathbb{Z}^+\}$ is the union of all vertices labelled with r in $\mathbb{A}$.

Step 2. We keep the index order of $V_{\mathbb{A}} = \{q_i\}_{i\in\mathbb{Z}^+}$ as in the labelling process above, and denote two collections of vertices labelled with r and b by S_r and S_b, respectively. That is $S_r = \bigcup_{i\in\mathbb{Z}^+} E_i$ and $S_b = V_{\mathbb{A}} \setminus S_r$. We show that the two partitioning sets S_r and S_b of $V_{\mathbb{A}}$ can be regarded as $R(\mathbb{A})$ and $B(\mathbb{A})$, respectively, in Definition 3.4 and hence the covering subtree $\mathbb{A}$ is a well-structured subtree under f.

For any covering subforest $\mathbb{G}$ of $\mathbb{A}$ and any finite set $S \subset V_{\mathbb{A}}$, we define

$$d(S) := \min\{n : \bigcup_{0\leq i\leq n} L_i(\mathbb{A}) \supseteq S\}.$$

If no ambiguity, we abbreviate $d(S)$ to d. By the definition of $z(k)$ in the labelling process, we see that in fact

$$\bigcup_{0\leq i\leq d} L_i(\mathbb{A}) = \{q_i : 0 \leq i < z(d)\}.$$

Hence $S \subseteq \{q_i : 0 \leq i < z(d)\}$. Recall that

$$E_i = \{q_l : 0 \leq l < i,\ q_l \text{ is labelled with r}\}.$$

It follows that $S_r \cap S \subseteq E_{z(d)}$. On the other hand, if $E_{z(d)} \neq \emptyset$, then for any covering subforest of $\mathbb{G}_{z(d)}$, and so in particular for $\mathbb{G}$ (since $V_{\mathbb{G}} \subseteq V_{\mathbb{A}} \subseteq V_{\mathbb{G}_{z(d)}}$), there exists an infinite increasing sequence $\{\delta_i\}$ of bounded stopping times adapted to $\mathbb{G}$ such that

$$f(q, \delta_i) = \text{Red } \ \forall q \in E_{z(d)}\ \forall i > 0,$$

by the definition of r. Therefore,

$$f(q, \delta_i) = \text{Red } \ \forall q \in S_r \cap S\ \forall i > 0.$$

If $E_{z(d)} = \emptyset$, pick any infinite increasing sequence of bounded stopping times adapted to $\mathbb{G}$ to be $\{\delta_i\}_{i>0}$.

Suppose that one vertex q_m is in $S_b \cap S$. By the definition of $\mathbb{G}_m(q_m)(E_m)$ and the condition II of the definition of b, for any infinite increasing sequence $\{\tau_i\}$ of bounded stopping times adapted to $\mathbb{G}_m(q_m)(E_m)$ with

$$f(q, \tau_i) = \text{Red } \ \forall i > 0\ \forall q \in E_m,$$

there exists $N_m \in \mathbb{N}$ such that $f(q_m, \tau_i) = \text{Blue } \forall i \geq N_m$. Since each bounded stopping time in $\{\delta_i\}_{i\in\mathbb{N}}$ is adapted to $\mathbb{G}$, and $\mathbb{G}$ is a covering subforest of $\mathbb{G}_m(q_m)(E_m)$, each bounded stopping time in $\{\delta_i\}_{i\in\mathbb{N}}$ is also adapted to $\mathbb{G}_m(q_m)(E_m)$. Because $E_{z(d)} \supseteq E_m$, we see

$$f(q, \delta_i) = \text{Red} \ \forall i > 0 \ \forall q \in E_m.$$

Therefore, we can take the sequence $\{\delta_i\}$ for the sequence $\{\tau_i\}$ above. Furthermore, for every vertex q_m in $S_b \cap S$, we can likewise take $\{\delta_i\}$ again for the sequence $\{\tau_i\}$ above. Hence there exists a corresponding $N_m \in \mathbb{N}$ such that

$$f(q_m, \delta_i) = \text{Blue} \ \forall i \geq N_m.$$

Because $S_b \cap S$ is finite, we can define N to be the maximum of those N_m. Hence,

$$f(q, \delta_i) = \text{Blue} \ \forall q \in S_b \cap S \ \forall i \geq N.$$

Define a new infinite sequence of bounded stopping time $\{\delta_i'\}_{i\in\mathbb{N}}$ such that $\delta_i' = \delta_{i+N} \ \forall i \in \mathbb{N}$. This is the desired sequence of bounded stopping times as per Definition 3.4, for the finite set S in the covering subforest $\mathbb{G}$. □

Comment on the labelling process. The covering subforest $\mathbb{A}$ is generated from the covering set $\{q_i\}_{i\in\mathbb{Z}^+}$. In practice, we achieve $\mathbb{A}$ by building the sequence $\{L_n(\mathbb{A})\}$ from the sequence $\{q_i\}$. For all i, it is important to replace the current covering subforest $\mathbb{G}_i$ by a covering subforest pruned below q_i's level, if q_i is labelled with b. If we simply replace $\mathbb{G}_i$ by an arbitrary covering subforest $\mathbb{G}'$ when q_i is marked with b relative to E_i with witness $\mathbb{G}'$ in $\mathbb{G}_i$, then we may fail to achieve the desired sequence $\{L_n(\mathbb{A})\}$. Consider the extreme case of every q_i labelled with b, and assume that we are now building $L_m(\mathbb{A})$ in which $m > 0$, and that the minimum number of vertices in $\mathbb{G}_i$ to form a $L_m(\mathbb{A})$ in $\mathbb{G}_i$ with the current subsequence $\{q_k\}_{0\leq k\leq i}$ is l. The arbitrary covering subforest $\mathbb{G}'$ may stretch out so wildly that to form a $L_m(\mathbb{A})$ in $\mathbb{G}'$, the minimum number of vertices in $\mathbb{G}'$ with the current $\{q_k\}_{0\leq k\leq i}$ is far greater than $l+1$. We can, however, only include one vertex q_{i+1} into the sequence $\{q_k\}$. If we regard this $\mathbb{G}'$ as $\mathbb{G}_{i+1}$, and achieve a sequence $\{q_k\}_{k\geq 0}$ in this way, we may never obtain the desired $L_m(\mathbb{A})$. In that case, the sequence $\{q_k\}_{k\geq 0}$ is not a covering set, and we cannot generate any covering subforest from it.

We fix one well-structured subtree $\mathbb{A}$ of $\mathbb{F}$ under f obtained by the labelling process. As in the definition after Lemma 3.3, we obtain the filtration $\{\mathcal{A}_n\}$ from this $\mathbb{A}$. That is, for each $n \in \mathbb{Z}^+$, $\mathcal{A}_n = \sigma\{B(q) : q \in L_n(\mathbb{A})\}$.

For our next result, we need a further notation. In any covering subforest $\mathbb{G}$ of $\mathbb{A}$, for any bounded stopping time θ adapted to $\mathbb{G}$, we denote

$$r_\theta(\mathbb{G}) := \cup\{B(q) : q \in S_\theta(\mathbb{G}) \cap R(\mathbb{A})\}$$

and

$$b_\theta(\mathbb{G}) := \cup\{B(q) : q \in S_\theta(\mathbb{G}) \cap B(\mathbb{A})\}.$$

Recall the definition of events happening *infinitely often* or *eventually*. Suppose that $(E_n : n \in \mathbb{N})$ is a sequence of events. We define

$$(E_n\ i.o.) := (E_n \text{ infintely often}) := \limsup E_n := \bigcap_m \bigcup_{n \geq m} E_n$$

and

$$(E_n\ ev.) := (E_n \text{ eventually}) := \liminf E_n := \bigcup_m \bigcap_{n \geq m} E_n.$$

The following lemma says that, for any $\epsilon > 0$, we can find a 'nice' covering subforest $\bar{\mathbb{F}}$ of $\mathbb{A}$ with the associated filtration $\{\bar{\mathcal{F}}_n\}$. In $\{\bar{\mathcal{F}}_n\}$, we can approximate the set $S \in \bar{\mathcal{F}}_\infty$ where for any $\omega \in S$ the vertices in $\mathbb{A}$ along the path ω are labelled with r infinitely often (b eventually, respectively), by a set R (B, respectively) in $\bar{\mathcal{F}}_1$. Furthermore, we can arrange that $R \cup B = \Omega$, and the probability that a path includes no vertex labelled with b (r, respectively) in $\bar{\mathbb{F}}$ conditional on R (B, respectively) is greater than $1 - 5\epsilon/8$. We prove this lemma by a method adapted from one used in [16].

Lemma 3.8. *For any $\epsilon > 0$, there exists a covering subforest $\bar{\mathbb{F}}$ of $\mathbb{A}$ with the associated filtration $\{\bar{\mathcal{F}}_n\}$ which has two sets $R, B \in \bar{\mathcal{F}}_1$ such that $R \cup B = \Omega$ and, for any sequence of bounded stopping times $\theta_1 < \theta_2 < \dots$ adapted to $\bar{\mathbb{F}}$,*

$$P(\bigcap_{i \geq 1} r_{\theta_i}(\bar{\mathbb{F}}) | R) > 1 - 5\epsilon/8, \quad \textit{if } P(R) > 0,$$

and

$$P(\bigcap_{i \geq 1} b_{\theta_i}(\bar{\mathbb{F}}) | B) > 1 - 5\epsilon/8, \quad \textit{if } P(B) > 0.$$

Proof. Adopt the abbreviation $\sigma_i := \sigma_i(\mathbb{A})$ in this proof. Then $\{r_{\sigma_i}(\mathbb{A})\}_{i\in\mathbb{Z}^+}$ and $\{b_{\sigma_i}(\mathbb{A})\}_{i\in\mathbb{Z}^+}$ are well defined in $\mathbb{A}$. Because $r_{\sigma_i}(\mathbb{A}), b_{\sigma_i}(\mathbb{A}) \in \mathcal{A}_i$ and $L_i(\mathbb{A})$ is finite for every i, $\{r_{\sigma_i}(\mathbb{A})\ i.o.\}$ and $\{b_{\sigma_i}(\mathbb{A})\ ev.\}$ are both Borel measurable sets in $\mathcal{A}_\infty$. Let $Y_R := \{r_{\sigma_i}(\mathbb{A})\ i.o.\}$ and $Y_B := \{b_{\sigma_i}(\mathbb{A})\ ev.\}$. Note that the topological space associated with $\mathbb{A}$ is a Cantor space, in particular a metric space. Since any measure on a metric space is regular (cf. [12], Theorem II.1.2), P is regular in $\mathcal{A}_\infty$. Therefore, there exists $N \in \mathbb{N}$ such that we can find two sets R,$B \in \mathcal{A}_N$ to approximate Y_R, Y_B respectively, i.e.

(a) $R \bigcup B = \Omega$,

(b) $P(Y_R|R) > 1 - \epsilon/2$, if $P(R) > 0$,

(c) $P(Y_B|B) > 1 - \epsilon/2$, if $P(B) > 0$.

Without loss of generality, assume both $P(R)$ and $P(B)$ are positive. Note that '$b_{\sigma_i}(\mathbb{A})$ eventually' implies '$b_{\sigma_i}(\mathbb{A})$ infinitely often'. Hence, for any $\epsilon' > 0$ and any $x \in \mathbb{N}$, there exists $y \in \mathbb{N}$ with $y > x$ such that

$$P(\bigcup_{x\le i<y} r_{\sigma_i}(\mathbb{A})|Y_R) > 1 - \epsilon'$$

and

$$P(\bigcup_{x\le i<y} b_{\sigma_i}(\mathbb{A})|Y_B) > 1 - \epsilon'.$$

Thus we can find a sequence of integers $n_0 = N < n_1 < n_2 < \ldots$ and two sequences of sets, to be denoted by $\{SR_l\}_{l\ge1}$ and $\{SB_l\}_{l\ge1}$ with $SR_l := \bigcup_{n_{l-1}\le i<n_l} r_{\sigma_i}(\mathbb{A})$ and $SB_l := \bigcup_{n_{l-1}\le i<n_l} b_{\sigma_i}(\mathbb{A})$, such that for every $l \ge 1$

$$P(SR_l|Y_R) > 1 - (1/8)(\epsilon/2^l)$$

and

$$P(SB_l|Y_B) > 1 - (1/8)(\epsilon/2^l).$$

It follows that

$$P(\bigcap_{l\ge1} SR_l|Y_R) > 1 - \epsilon/8$$

and

$$P(\bigcap_{l\ge1} SB_l|Y_B) > 1 - \epsilon/8.$$

From $P(Y_R|R) > 1 - \epsilon/2$ and $P(Y_B|B) > 1 - \epsilon/2$, we may infer that

$$P(\bigcap_{l\geq 1} SR_l|R) > 1 - 5\epsilon/8$$

and

$$P(\bigcap_{l\geq 1} SB_l|B) > 1 - 5\epsilon/8.$$

For each $l \geq 1$, define a stopping time δ_l adapted to $\mathbb{A}$ so that on $SR_l \cap R$

$$\delta_l(\omega) = \min\{i : \sigma_{n_{l-1}}(\omega) \leq i < \sigma_{n_l}(\omega);\ \omega|i \in R(\mathbb{A})\},$$

on $SB_l \cap B$

$$\delta_l(\omega) = \min\{i : \sigma_{n_{l-1}}(\omega) \leq i < \sigma_{n_l}(\omega);\ \omega|i \in B(\mathbb{A})\},$$

and on $(R \setminus SR_l) \cup (B \setminus SB_l)$,

$$\delta_l(\omega) = n_l - 1.$$

Define a covering subforest $\bar{\mathbb{F}}$ so that

$$\sigma_l(\bar{\mathbb{F}}) := \delta_{l+1}\ \forall l \geq 0.$$

For any infinite increasing sequence of bounded stopping times $\{\theta_i\}$ adapted to $\bar{\mathbb{F}}$,

$$\left((\bigcap_{i\geq 1} r_{\theta_i}(\bar{\mathbb{F}})) \bigcap R\right) \supseteq \left((\bigcap_{l\geq 1} SR_l) \bigcap R\right)$$

and

$$\left((\bigcap_{i\geq 1} b_{\theta_i}(\bar{\mathbb{F}})) \bigcap B\right) \supseteq \left((\bigcap_{l\geq 1} SB_l) \bigcap B\right).$$

This completes the proof. □

We are now ready to prove the Finite-state Two-colour Stochastic Partition Theorem (Theorem 3.1).

Proof of Theorem 3.1. In a covering subforest $\bar{\mathbb{F}}$ of $\mathbb{A}$ obtained from Lemma 3.8, we define θ_1 to be $\sigma_0(\bar{\mathbb{F}})$. We construct the desired infinite increasing sequence of bounded stopping times inductively. Suppose that we have obtained $\langle\theta_1, \theta_2, ..., \theta_k\rangle$. Denote $\bigcup_{1\le i\le k} S_{\theta_i}(\bar{\mathbb{F}})$ by S^k. Since $\bar{\mathbb{F}}$ is a covering subforest of $\mathbb{A}$ and S^k is finite, by Definition 3.4, there exists an infinite increasing sequence of bounded stopping times $\{\delta_i^k\}_{i\in\mathbb{N}}$ adapted to $\bar{\mathbb{F}}$ with the property

$$f(q, \delta_i^k) = \text{Red } \forall q \in S^k \cap R(\mathbb{A}) \ \forall i > 0$$

and

$$f(q, \delta_i^k) = \text{Blue } \forall q \in S^k \cap B(\mathbb{A}) \ \forall i > 0.$$

Define θ_{k+1} to be δ_1^k. In this way, we find an infinite increasing sequence of bounded stopping times $\{\theta_n\}_{n\in\mathbb{N}}$. It follows from Lemma 3.8 that, for all initial subsequences $\{\theta_1, ..., \theta_i\}$ $(i \ge 2)$ in it,

$$P(\hat{f}_{\theta_l,\theta_m} = \text{Red } \forall 1 \le l < m \le i | R) > 1 - 5\epsilon/8, \quad \text{if } P(R) > 0,$$

and

$$P(\hat{f}_{\theta_l,\theta_m} = \text{Blue } \forall 1 \le l < m \le i | B) > 1 - 5\epsilon/8, \quad \text{if } P(B) > 0,$$

which completes the proof. □

Comment: The desired sequence $\{\theta_n\}_{n\in\mathbb{N}}$ in Theorem 3.1 is by no means unique. When we label each vertex q_i in $\mathbb{F}$, the covering subforest $\mathbb{G}_{i+1} = \mathbb{G}_i(q_i)(E_i)$ is not unique, if q_i is labelled with b. Suppose the q_i is in $L_m(\mathbb{G}_i)$ for some m. The assertion follows from the same idea in Lemma 3.6: because we are only concerned with the existence of an infinite sequence of bounded stopping times, any deletion or addition of finite vertices in $\bigcup_{n>m} L_n(\mathbb{G}_i(q_i)(E_i))$ is allowed for the construction of $\mathbb{G}_{i+1}$. Then, the well-structured subtree $\mathbb{A}$ under f is not unique, as the generating sequence of covering subforests $\{\mathbb{G}_i\}_{i\in\mathbb{N}}$ can vary. The extraction of $\bar{\mathbb{F}}$ from $\mathbb{A}$ is not fixed either. To see this, any covering subforest of $\bar{\mathbb{F}}$ still suffices to be one $\bar{\mathbb{F}}$, by the definition of $\mathbb{A}$ (Definition 3.4). When we build θ_{k+1} based on $\bigcup_{1\le i\le k} S_{\theta_i}(\bar{\mathbb{F}})$, the definition of $\mathbb{A}$ only assures the existence of $\{\delta_i^k\}_{i\in\mathbb{N}}$ from which θ_{k+1} can be chosen, but there may well be many qualified sequences $\{\delta_i^k\}_{i\in\mathbb{N}}$ in $\mathbb{F}$. In summary, the collection of infinite increasing sequences of bounded stopping times $\{\theta_n\}_{n\in\mathbb{N}}$ which satisfy Theorem 3.1 is 'rich'.

Given the proof of Theorem 3.1, the multiple-colour case is straightforward except some minor modification in the marking process. For completeness, we give below this more general theorem for the set of colours $C = \{c_1, ..., c_k\}$ with $k \geq 2$. (Of course, for $k = 1$, the problem is trivial.)

Theorem 3.9. (Finite-state Finite-colour Stochastic Partition Theorem). *Given a set $C = \{c_1, c_2, ...c_k\}$, a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence of sets of finitely many states, and a stochastic colouring f with values in C, then for every $\epsilon > 0$, there exists a natural number N, k sets $Q_m \in \mathcal{F}_N$ $\forall 1 \leq m \leq k$ with $\bigcup_{1 \leq m \leq k} Q_i = \Omega$, and a sequence of bounded stopping times $\theta_1 < \theta_2 < ...$ such that*

$$\forall 1 \leq m \leq k, P(\hat{f}_{\theta_i, \theta_j} = c_m,\ \forall 1 \leq i < j | Q_m) > 1 - \epsilon,\ \text{if } P(Q_m) > 0.$$

We begin by the definition of a well-structured subtree in this multiple-colour context.

Definition 3.10. *For a set $C = \{c_1, c_2, ..., c_k\}$ of finitely many colours, a well-structured subtree $\mathbb{A}$ under the stochastic colouring f is a covering subtree of $\mathbb{F}$ with a partition of its vertex set $V_\mathbb{A} = \bigcup_{1 \leq i \leq k} S_i(\mathbb{A})$ such that for any covering subforest $\mathbb{G}$ of $\mathbb{A}$ and any finite subset $S \subset V_\mathbb{G}$, there exists an infinite sequence of bounded stopping times $\delta_1 < \delta_2 < ...$ adapted to $\mathbb{G}$ with the property*

$$\forall 1 \leq i \leq k,\ f(q, \delta_l) = c_i\ \forall q \in S \cap S_i(\mathbb{A})\ \forall l > 0.$$

Remark: The partition above may be degenerate, that is, some partitioning sets of $V_\mathbb{A}$ may be empty.

Similarly as in the case of $|C| = 2$, we are going to find a well-structured subtree $\mathbb{A}$ of $\mathbb{F}$ under f by breadth-first search. When we define inductively a sequence of vertices $\{q_i\}$, each vertex in it will be labelled with a symbol in $\{s_1, ..., s_k\}$. We will show that the sequence $\{q_i\}$ generates a well-structured subtree $\mathbb{A}$ of $\mathbb{F}$ under f and, for each i with $1 \leq i \leq k$, the collection of vertices labelled with s_i is exactly the partitioning set $S_i(\mathbb{A})$.

In doing so, we need a more elaborate *multi-colour* marking scheme below. As with the two-colour marking scheme, here vertices will be 'marked' using distinct symbols $s_1, ..., s_k$ corresponding to the colours $c_1, ..., c_k$. A vertex marked with s_i can be labelled with s_i in the partial labelling scheme later. The marking scheme is done by means of a definition that begins with a symbol s_1 by referring to a property involving only the colour c_1, and proceeds to a symbol s_m $(1 < m < k)$ by reference to a property involving only the colours $c_1, ..., c_m$ (in such a way that s_m is the earliest available to mark with – see condition b3 below). The final symbol s_k (corresponding to c_k) is 'complementary' to all the preceding cases, as we shall see. To avoid confusion, we have denoted the symbols by $s_1, ..., s_k$ rather than $c_1, ..., c_k$.

Suppose that we are given a covering subforest $\mathbb{G}$ of $\mathbb{F}$ and a $(k-1)$ sequence $\langle M_1, M_2, ..., M_{k-1} \rangle$ of sets of nodes with the following properties.

I. $M_i \subseteq V_{\mathbb{G}}\ \forall 1 \leq i < k$.

II. M_i is finite (perhaps empty) for each $1 \leq i < k$.

III. For any covering subforest $\mathbb{G}'$ of $\mathbb{G}$, there exists a sequence of bounded stopping times $\delta_1 < \delta_2 < ...$ adapted to $\mathbb{G}'$ such that, for any i with $1 \leq i < k$, $f(q^i, \delta_l) = c_i\ \forall l > 0\ \forall q^i \in M_i$.

We call such a sequence $\langle M_1, M_2 ..., M_{k-1} \rangle$ a $(k-1)$-*matched collection* in $\mathbb{G}$.

For every vertex q in $V_{\mathbb{G}} \setminus (\bigcup_{1 \leq i < k} M_i)$, mark q as follows.

a. Mark q with s_1 relative to $\langle M_1, M_2, ..., M_{k-1} \rangle$ in $\mathbb{G}$ if, for any covering subforest $\mathbb{G}'$ of $\mathbb{G}$, there exists a sequence of bounded stopping times $\delta_1 < \delta_2 < ...$ adapted to $\mathbb{G}'$ such that, for any i with $1 \leq i < k$, $f(q^i, \delta_l) = c_i\ \forall l > 0\ \forall q^i \in M_i$ and $f(q, \delta_l) = c_1\ \forall l > 0$.

b. Mark q with s_m $(1 < m < k)$ relative to $\langle M_1, M_2, ..., M_{k-1} \rangle$ in $\mathbb{G}$ if q satisfies the three conditions below with respect to s_m.

b1. There exists a covering subforest $\mathbb{G}^m$ of $\mathbb{G}$ pruned below q's level such that for any sequence of bounded stopping times $\delta_1 < \delta_2 < ...$ adapted to $\mathbb{G}^m$ which has the property that, for any i with $1 \leq i < k$, $f(q^i, \delta_l) = c_i \ \forall l > 0 \ \forall q^i \in M_i$, there exists $N \in \mathbb{N}$ such that $f(q, \delta_l) \notin \{c_1, ..., c_{m-1}\} \ \forall l \geq N$. Denote any one such $\mathbb{G}^m$ by $\mathbb{G}(q)(\langle M_1, M_2, ..., M_{k-1}\rangle)(m)$.

b2. For any covering subforest $\mathbb{G}'$ of $\mathbb{G}(q)(\langle M_1, M_2, ..., M_{k-1}\rangle)(m)$, there exists a sequence of bounded stopping times $\tau_1 < \tau_2 < ...$ adapted to $\mathbb{G}'$ such that, for any i with $1 \leq i < k$, $f(q^i, \tau_l) = c_i \ \forall l > 0 \ \forall q^i \in M_i$ and $f(q, \tau_l) = c_m \ \forall l > 0$.

b3. For all $1 < i < m$, q does not satisfy the conjunction of 1 and 2 above with respect to s_i. (Thus s_m is the earliest symbol to satisfy both 1 and 2.)

c. Mark q with s_k relative to $\langle M_1, M_2, ..., M_{k-1}\rangle$ in $\mathbb{G}$ if q cannot be marked with any of $\{s_1, ..., s_{k-1}\}$ relative to $\langle M_1, M_2, ..., M_{k-1}\rangle$ in $\mathbb{G}$.

According to the definition above, a vertex q is marked with s_k relative to $\langle M_1, M_2, ..., M_{k-1}\rangle$ if the following circumstance holds. There exists a covering subforest $\mathbb{G}'$ of $\mathbb{G}$ pruned below q's level such that for any sequence of bounded stopping times $\delta_1 < \delta_2 < ...$ adapted to $\mathbb{G}'$ there exists $N \in \mathbb{N}$ with $f(q, \delta_l) = c_k \ \forall l \geq N$, provided that, for all i with $1 \leq i < k$, $f(q^i, \delta_l) = c_i \ \forall l > 0 \ \forall q^i \in M_i$. Denote any one such $\mathbb{G}'$ by $\mathbb{G}(q)(\langle M_1, M_2, ..., M_{k-1}\rangle)(k)$.

Remark: For a vertex q marked with symbol s_m $(1 < m \leq k)$ relative to $\langle M_1, M_2, ..., M_{k-1}\rangle$ in $\mathbb{G}$, we can find a covering subforest $\mathbb{G}^m$ of $\mathbb{G}$ pruned below q's level with the attached conditions by the same arguments in Lemma 3.6.

Lemma 3.11. *Given a set $C = \{c_1, c_2, ..., c_k\}$ of finitely many colours, there exists in the tree of histories $\mathbb{F}$ a well-structured subtree $\mathbb{A}$ under f.*

Proof. As in the two-colour case, we first extract a subtree $\mathbb{A}$ of $\mathbb{F}$ by a labelling process, then prove this subtree $\mathbb{A}$ is a well-structured subtree of $\mathbb{F}$ under f.

Step 1. We construct:

i. $k-1$ families $\{E_i^1\}$,...,$\{E_i^{k-1}\}$ (consisting of vertices) with $E_0^m = \emptyset$ and $E_i^m \subseteq E_{i+1}^m$ $\forall 0 < m < k$ $\forall i \geq 0$;

ii. a sequence of covering subforests $\{\mathbb{G}_i\}$ of $\mathbb{F}$ with $\mathbb{G}_0 = \mathbb{F}$ and $V_{\mathbb{G}_i} \supseteq V_{\mathbb{G}_{i+1}}$ $\forall i \geq 0$; and

iii. a partial labelling of $\mathbb{G}_i$ with symbols in $\{s_1, ..., s_k\}$.

The inductive hypothesis ensures that the $(k-1)$ sequence $\langle E_i^1$,...,$E_i^{k-1}\rangle$ to be constructed forms a $(k-1)$-matched collection in $\mathbb{G}_i$ for each i. Given E_i^1,...,E_i^{k-1} and $\mathbb{G}_i$, we will, by an analogous procedure to that in the earlier proof of Lemma 3.7, pick some vertex q_i which has not been labelled in $\mathbb{G}_i$ to define E_{i+1}^1,...,E_{i+1}^{k-1} and $\mathbb{G}_{i+1}$. We proceed in parallel to the definition of a multi-colour marking scheme:

a. If q_i is marked with s_1 relative to $\langle E_i^1, ..., E_i^{k-1}\rangle$ in $\mathbb{G}_i$, let $E_{i+1}^1 = E_i^1 \cup \{q_i\}$, $E_{i+1}^m = E_i^m$ $\forall 1 < m < k$, $\mathbb{G}_{i+1} = \mathbb{G}_i$, and label q_i with s_1.

b. If q_i is marked with s_l $(1 < l < k)$ relative to $\langle E_i^1, ..., E_i^{k-1}\rangle$ in $\mathbb{G}_i$, then there exists a covering subforest $\mathbb{G}_i(q_i)(\langle E_i^1, ..., E_i^{k-1}\rangle)(l)$ of $\mathbb{G}_i$ pruned below q_i's level, as in b1. Let $\mathbb{G}_{i+1}$ be this $\mathbb{G}_i(q_i)(\langle E_i^1, ..., E_i^{k-1}\rangle)(l)$, let $E_{i+1}^l = E_i^l \cup \{q_i\}$, $E_{i+1}^m = E_i^m$ $\forall 0 < m < l$ & $l < m < k$, and label q_i with s_l.

c. If q_i is marked with s_k relative to $\langle E_i^1, ..., E_i^{k-1}\rangle$ in $\mathbb{G}_i$, then there exists a covering subforest $\mathbb{G}_i(q_i)(\langle E_i^1, ..., E_i^{k-1}\rangle)(k)$ of $\mathbb{G}_i$ pruned below q_i's level, as in the paragraph following c above. Take this $\mathbb{G}_i(q_i)(\langle E_i^1, ..., E_i^{k-1}\rangle)(k)$ to be $\mathbb{G}_{i+1}$, let $E_{i+1}^m = E_i^m$ $\forall 0 < m < k$, and label q_i with s_k.

We now explain how to determine the vertices q_i which are to be picked. Denote the root of $\mathbb{F}$ by q_0. After obtaining $\langle E_1^1, ..., E_1^{k-1}\rangle$ and $\mathbb{G}_1$ from q_0, $\langle E_0^1, ..., E_0^{k-1}\rangle$ and $\mathbb{G}_0$, we enumerate all vertices in $L_1(\mathbb{G}_1)$, which is the 1[st] level set of $\mathbb{G}_1$ defined in Section 3.3. As noted at the beginning of this section, $L_1(\mathbb{G}_1)$ is finite. We suppose $|L_1(\mathbb{G}_1)| = l_1$, and enumerate the elements of $L_1(\mathbb{G}_1)$ as $q_1, q_2, ...q_{l_1}$. We now apply the above process to

$q_1, q_2, \ldots q_{l_1}$ in turn, and obtain finite sequences $\{E_i^1\}, \ldots, \{E_i^{k-1}\}$ and $\{\mathbb{G}_i\}$ with $1 < i \leq l_1 + 1$. Note that $L_1(\mathbb{G}_i) = L_1(\mathbb{G}_{i+1})$ $\forall 1 \leq i \leq l_1$, since if q_i is labeled with s_m $(1 < m \leq k)$, $\mathbb{G}_{i+1}$ is a covering subforest of $\mathbb{G}_i$ pruned below $\sigma_1(\mathbb{G}_i)$. It follows that all q_1 to q_{l_1} indexed in $\mathbb{G}_1$ are in $L_1(\mathbb{G}_i)$ for all $0 < i \leq l_1 + 1$. We then enumerate all vertices in $L_2(\mathbb{G}_{l_1+1})$ as $q_{l_1+1}, q_{l_1+2}, \ldots, q_{l_1+l_2}$. By applying the process to $q_{l_1+1}, q_{l_1+2}, \ldots, q_{l_1+l_2}$, we obtain finite sequences $\{E_i^1\}, \ldots, \{E_i^{k-1}\}$ and $\{\mathbb{G}_i\}$ for all $l_1 + 1 < i \leq l_1 + l_2 + 1$. We can then define all $L_{i+1}(\mathbb{G}_{1+\sum_{m\leq i} l_m})$ for all i in a similar way.

Let $l_0 = 1$. Define $z(k) := \sum_{0\leq i\leq k} l_i$. Given k, it follows that

$$L_n(\mathbb{G}_{z(k)}) = L_n(\mathbb{G}_i) \ \forall 0 \leq n \leq k \ \forall i > z(k),$$

and in any $\mathbb{G}_i$ with $i \geq z(k)$, all vertices in $\bigcup_{0\leq n\leq k} L_n(\mathbb{G}_i)$ have already been labeled. Note $V_{\mathbb{G}_{z(k)}} \supseteq V_{\mathbb{G}_{z(k+1)}}$ $\forall k \geq 0$. So, by Fusion Lemma (Lemma 3.3), the sequence $\{L_k(\mathbb{G}_{z(k)})\}_{k\geq 0}$ generates a covering subforest of $\mathbb{F}$. Denote this covering subforest by $\mathbb{A}$. Note that since $L_0(\mathbb{G}_{z(0)})$ only contains the root of $\mathbb{F}$, $\mathbb{A}$ is actually a covering subtree of $\mathbb{F}$. From the definition of $z(k)$, $\mathbb{A}$ is the covering subtree that contains exactly the vertices labeled. For every $1 \leq n < k$, $\bigcup_{i>0} E_i^n$ is the union of all vertices labeled with s_n in $\mathbb{A}$.

Step 2. We keep the index order of $V_{\mathbb{A}} = \{q_i\}_{i\in\mathbb{Z}^+}$ as in the labelling process above. Denote the collection of vertices labelled with s_i by S_i. That is, $S_i = \bigcup_{m>0} E_m^i$ $\forall 1 \leq i < k$ and $S_k = V_{\mathbb{A}} \setminus (\bigcup_{1\leq i<k} S_i)$. We show the partition $\{S_1, \ldots, S_k\}$ of $V_{\mathbb{A}}$ satisfies Definition 3.10 and hence the covering subtree $\mathbb{A}$ is a well-structured subtree under f.

For any covering subforest $\mathbb{G}$ of $\mathbb{A}$ and any finite set $S \subset V_{\mathbb{A}}$, we define

$$d(S) := \min\{n : \bigcup_{0\leq i\leq n} L_i(\mathbb{A}) \supseteq S\}.$$

If no ambiguity, we abbreviate $d(S)$ to d. By the definition of $z(k)$ in the labelling process, we see that in fact

$$\bigcup_{0\leq i\leq d} L_i(\mathbb{A}) = \{q_i : 0 \leq i < z(d)\}.$$

Hence $S \subseteq \{q_i : 0 \leq i < z(d)\}$. Recall that

$$E_m^i = \{q_l : 0 \leq l < m, \ q_l \text{ is labelled with } s_i\} \ \forall 1 \leq i < k.$$

So, for each i with $1 \leq i < k$, we have $S_i \cap S \subseteq E^i_{z(d)}$. On the other hand, if $\bigcup_{1\leq i<k} E^i_{z(d)} \neq \emptyset$, then for any covering subforest of $\mathbb{G}_{z(d)}$ and so in particular for $\mathbb{G}$ (since $V_{\mathbb{G}} \subseteq V_{\mathbf{A}} \subseteq V_{\mathbb{G}_{z(d)}}$), there exists an infinite increasing sequence $\{\delta_i\}$ of bounded stopping times adapted to $\mathbb{G}$ such that

$$f(q^i, \delta_l) = c_i \quad \forall 1 \leq i < k \; \forall q^i \in E^i_{z(d)} \; \forall l > 0,$$

by the labelling process. Therefore,

$$f(q^i, \delta_l) = c_i \quad \forall 1 \leq i < k \; \forall q^i \in S_i \cap S \; \forall l > 0.$$

If $\bigcup_{1\leq i<k} E^i_{z(d)} = \emptyset$, pick any infinite increasing sequence of bounded stopping times adapted to $\mathbb{G}$ to be $\{\delta_i\}_{i>0}$.

Suppose that one vertex q_m is in $S_k \cap S$. By the labelling process, for any infinite increasing sequence $\{\tau_l\}$ of bounded stopping times adapted to $\mathbb{G}_m(q_m)(\langle E^1_m, ..., E^{k-1}_m\rangle)(k)$ with

$$f(q^i, \tau_l) = c_i \quad \forall 1 \leq i < k \; \forall q^i \in E^i_m \; \forall l > 0,$$

there exists $N_m \in \mathbb{N}$ such that $f(q_m, \tau_l) = c_k \; \forall l \geq N_m$. Since each bounded stopping time in $\{\delta_l\}_{l\in\mathbb{N}}$ is adapted to $\mathbb{G}$, and $\mathbb{G}$ is a covering subforest of $\mathbb{G}_m(q_m)(\langle E^1_m, ..., E^{k-1}_m\rangle)(k)$, each bounded stopping time in $\{\delta_i\}_{i\in\mathbb{N}}$ is also adapted to $\mathbb{G}_m(q_m)(\langle E^1_m, ..., E^{k-1}_m\rangle)(k)$. Because $E^i_{z(d)} \supseteq E^i_m \; \forall 1 \leq i < k$, we see

$$f(q^i, \delta_l) = c_i \quad \forall 1 \leq i < k \; \forall q^i \in E^i_m \; \forall l > 0.$$

Therefore, we can take $\{\delta_l\}_{l>0}$ for the sequence $\{\tau_l\}_{l>0}$ above. Furthermore, for every vertex q_m in $S_k \cap S$, we can likewise take $\{\delta_l\}_{l>0}$ for the sequence $\{\tau_l\}_{l>0}$ above. Hence there exists a corresponding $N_m \in \mathbb{N}$ such that

$$f(q_m, \delta_l) = c_k \; \forall l \geq N_m.$$

Because $S_k \cap S$ is finite, we can define N to be the maximum of those N_m. Hence,

$$f(q^k, \delta_l) = c_k \; \forall q^k \in S_k \cap S \; \forall l \geq N.$$

Define a new infinite sequence of bounded stopping time $\{\delta'_l\}_{l\in\mathbb{N}}$ such that $\delta'_l = \delta_{l+N} \; \forall l \in \mathbb{N}$. This is the desired sequence of bounded stopping times as per Definition 3.10, for the finite set S in the covering subforest $\mathbb{G}$. $\square$

We fix one well-structured subtree $\mathbb{A}$ of $\mathbb{F}$ under f. As in the definition after Lemma 3.3, we obtain the filtration $\{\mathcal{A}_n\}$ from this $\mathbb{A}$. That is, for each $n \in \mathbb{Z}^+$, $\mathcal{A}_n = \sigma\{B(q) : q \in L_n(\mathbb{A})\}$. For any bounded stopping time θ adapted to $\mathbb{A}$, we define

$$\bar{S}^i_\theta = \cup\{B(q) : q \in S_\theta(\mathbb{A}) \cap S_i(\mathbb{A})\} \ \forall 1 \le i \le k.$$

We prove the generalised Lemma 3.8 for $|C| = k$.

Lemma 3.12. *For any $\epsilon > 0$, there exists a covering subforest $\bar{\mathbb{F}}$ of $\mathbb{A}$ with the associated filtration $\{\bar{\mathcal{F}}_n\}$ which has k sets $Q_i \in \bar{\mathcal{F}}_1$ $\forall 1 \le i \le k$ such that $\bigcup_{1 \le i \le k} Q_i = \Omega$ and, for any sequence of bounded stopping times $\theta_1 < \theta_2 < \ldots$ adapted to $\bar{\mathbb{F}}$ and for each i with $1 \le i \le k$,*

$$P(\bigcap_{m \ge 1} \bar{S}^i_{\theta_m} | Q_i) > 1 - 5\epsilon/8, \quad \text{if } P(Q_i) > 0.$$

Proof. Adopt the abbreviation $\sigma_i := \sigma_i(\mathbb{A})$ in this proof. Then $\{\bar{S}^i_{\sigma_n} : 1 \le i \le k, n \ge 0\}$ are well defined in $\mathbb{A}$. Because $\bar{S}^i_{\sigma_n} \in \mathcal{A}_n$ $\forall 1 \le i \le k$ and $L_n(\mathbb{A})$ is finite for every $n \ge 0$, $\{\bar{S}^i_{\sigma_n}\ i.o.\}$ and $\{\bar{S}^i_{\sigma_n}\ ev.\}$ with respect to n are both Borel measurable sets in $\mathcal{A}_\infty$, for all $1 \le i \le k$. Let $Y_i := \{\bar{S}^i_{\sigma_n}\ i.o.\}$ $\forall 1 \le i \le k$ and $Y_k := \{\bar{S}^k_{\sigma_n}\ ev.\}$. Note that the topological space associated with $\mathbb{A}$ is a Cantor space, in particular a metric space. From Theorem 1.2 on page 27 in [12], P is regular in $\mathcal{A}_\infty$. So, given any $\epsilon' > 0$, for any set $S \in \mathcal{A}_\infty$, we can find a set $S' \in \bigcup_n \mathcal{A}_n$ such that $P(S \bigtriangleup S') < \epsilon'$. Therefore, there exists $N \in \mathbb{N}$ such that we can find k sets $Q_1, \ldots, Q_k \in \mathcal{A}_N$ with the property that

(a) $\bigcup_{1 \le i \le k} Q_i = \Omega$,

(b) $\forall 1 \le i \le k$, $P(Y_i | Q_i) > 1 - \epsilon/2$, if $P(Q_i) > 0$.

Let us note that '$\bar{S}^k_{\sigma_n}$ eventually' implies '$\bar{S}^k_{\sigma_n}$ infinitely often'. Hence, For any $\epsilon' > 0$, for any $x \in \mathbb{N}$, there exists $y \in \mathbb{N}$ with $y > x$ such that

$$P(\bigcup_{x \le n < y} \bar{S}^i_{\sigma_n} | Y_i) > 1 - \epsilon' \ \forall 1 \le i \le k.$$

Thus we can find a sequence of integers $n_0 = N < n_1 < n_2 < \ldots$ and k sequences of sets to be denoted by $\{W^i_l\}_{l \ge 1}$ $\forall 1 \le i \le k$ with $W^i_l :=$

$\bigcup_{n_{l-1}\le m<n_l} \bar{S}^i_{\sigma_m}$ such that for every $l \ge 1$

$$P(W^i_l|Y_i) > 1-(1/8)(\epsilon/2^l).$$

It follows that

$$P(\bigcap_{l\ge 1} W^i_l|Y_i) > 1-\epsilon/8\ \forall 1\le i\le k.$$

From $P(Y_i|Q_i) > 1-\epsilon/2\ \forall 1 \le i \le k$, we may infer that

$$P(\bigcap_{l\ge 1} W^i_l|Q_i) > 1-5\epsilon/8\ \forall 1\le i\le k.$$

For each $l \ge 1$, define a stopping time δ_l adapted to $\mathbb{A}$ such that, for every $1 \le i \le k$, on $W^i_l \cap Q_i$

$$\delta_l(\omega) = \min\{m : \sigma_{n_{l-1}}(\omega) \le m < \sigma_{n_l}(\omega);\ \omega|m \text{ is labeled with } s_i \text{ in } \mathbb{A}\},$$

on $\bigcup_{1\le i\le k}(Q_i \setminus W^i_l)$,

$$\delta_l(\omega) = n_l - 1.$$

Define a covering subforest $\bar{\mathbb{F}}$ such that

$$\sigma_l(\bar{\mathbb{F}}) := \delta_{l+1}\ \forall l \ge 0.$$

For any infinite increasing sequence of bounded stopping times $\{\theta_m\}$ adapted to $\bar{\mathbb{F}}$,

$$\left((\bigcap_{m\ge 1} \bar{S}^i_{\theta_m})\bigcap Q_i\right) \supseteq \left((\bigcap_{l\ge 1} W^i_l)\bigcap Q_i\right)\ \forall 1\le i\le k.$$

This completes the proof. □

As in the proof of Theorem 3.1, we prove the Finite-state Finite-colour Stochastic Partition Theorem (Theorem 3.9) by induction.

Proof of Theorem 3.9. In a covering subforest $\bar{\mathbb{F}}$ of $\mathbb{A}$ obtained from Lemma 3.12, we define θ_1 to be $\sigma_0(\bar{\mathbb{F}})$. We construct the desired infinite increasing sequence of bounded stopping times inductively. Suppose that we have obtained $\langle\theta_1, \theta_2, ..., \theta_t\rangle$. Denote $\bigcup_{1\le i\le t} S_{\theta_i}(\bar{\mathbb{F}})$ by S^t. Since $\bar{\mathbb{F}}$ is a covering subforest of $\mathbb{A}$ and S^t is finite, by Definition 3.10, there exists an infinite increasing sequence of bounded stopping times $\{\delta_i^t\}_{i\in\mathrm{N}}$ adapted to $\bar{\mathbb{F}}$ such that

$$f(q^i,\delta_l^t)=c_i \quad \forall 1\leq i\leq k\ \forall q^i\in S^t\cap S_i(\mathbb{A})\ \forall l>0.$$

Define θ_{k+1} to be δ_1^k. In this way, we find an infinite increasing sequence of bounded stopping times $\{\theta_n\}_{n\in\mathrm{N}}$. It follows from Lemma 3.12 that for all initial subsequences $\{\theta_1,...,\theta_t\}$ $(t\geq 2)$ in it and each i with $1\leq i\leq k$

$$P(\hat{f}_{\theta_l,\theta_m}=c_i\ \forall 1\leq l<m\leq t|Q_i)>1-5\epsilon/8, \quad \text{if } P(Q_i)>0,$$

which completes the proof. □

3.5 The stochastic Ramsey theorem for sets of countably many states

Before proving the Stochastic Infinitary Ramsey Theorem (Theorem 1.2), we comment on the approach to be taken. Suppose that in the sequence $\{J_i\}_{i\in\mathrm{N}}$ of sets of states, some J_i are countably infinite sets, and we arrive at the probability space $(\Omega,\mathcal{F}_\infty,P)$ constructed from $\{J_i\}_{i\in\mathrm{N}}$. If we represent Ω as the tree $\mathbb{F}$ defined in Section 3.3, then for some bounded stopping times $\theta>0$ in $\mathbb{F}$, the set of stopping places of θ in $\mathbb{F}$, i.e. $S_\theta(\mathbb{F})$, is infinite. That may invalidate the approach in Section 3.4. The extreme case is that J_i is a countably infinite set for all i. Then, for any covering subforest $\mathbb{G}$ of $\mathbb{F}$ and any $i>0$, $L_i(\mathbb{G})$ is infinite. Hence, we would never finish the labelling process, if we insisted on following exactly the same procedure as in Section 3.4.

To prove Theorem 1.2, we follow the strategy mentioned in the Introduction: ignore all but finitely many vertices at each level of the tree $\mathbb{F}$ so that the probability of the ignored part of $\mathbb{F}$ is 'small', and the remaining tree looks like one constructed from a sequence of sets of finitely many states; then we can prove the result by the Finite-state Finite-colour Stochastic Partition Theorem (Theorem 3.9). This is in the nature of routine. For a detailed proof, there is a technical problem: the finite set of successor vertices at level $n+1$ depends on its predecessor vertex at level n. Recall that our setting in Theorem 3.9 and Theorem 3.1 requires a 'uniform' J_{n+1}, that is, every vertex at level n has the same finite set of successor vertices at level $n+1$. To this end, we shall construct an adjusted probability space $(\Omega', \mathcal{F}'_\infty, P')$ from a sequence $\{J'_i\}_{i\in\mathbb{N}}$ of sets of finitely many states. By Theorem 3.9, we then find a 'solution sequence' $\{\theta'_n\}_{n\in\mathbb{N}}$ of increasing bounded stopping times after defining a new stochastic colouring f' on $(\Omega', \mathcal{F}'_\infty, P')$. We will eventually show that an image sequence $\{\theta_n\}_{n\in\mathbb{N}}$ defined on the original probability space $(\Omega, \mathcal{F}_\infty, P)$ satisfies the condition of Theorem 1.2.

For any partial history $q \in \Omega_n$, denote the collection of extensions of q at stage $n+1$, i.e., $\{\omega|(n+1) : \omega \in B(q)\}$, by $e(q)$.

Proof of Theorem 1.2. Let $\epsilon > 0$ be given. To obtain the adjusted probability space constructed from a sequence of sets of finitely many states, we first construct a 'bridging' probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}_\infty, \tilde{P})$ from $(\Omega, \mathcal{F}_\infty, P)$ by defining the collection $\tilde{\Omega}_n$ of partial histories for each n.

We place an *absorbing state* $*$ in the yet-to-be-constructed $\tilde{J}_n$ for all n with the property that for any partial history $\tilde{q} = (\tilde{q}(1), ..., \tilde{q}(n) = *)$, the only extension of $\tilde{q}$ in the yet-to-be-constructed $\tilde{\Omega}_{n+1}$ is $(\tilde{q}(1), ..., \tilde{q}(n) = *, \tilde{q}(n+1) = *)$. Let $\tilde{\Omega}_0 = \Omega_0$, i.e., the set of empty partial history. Assume that we have already defined $\tilde{\Omega}_n$ for all $0 \le n \le k$. For every partial history $\tilde{q} = (\tilde{q}(1), ..., \tilde{q}(k))$, if $\tilde{q}(k) = *$, then we already know the only extension of $\tilde{q}$ at stage $k+1$ is $(\tilde{q}(1), ..., \tilde{q}(k) = *, \tilde{q}(k+1) = *)$. If $\tilde{q}(k) \neq *$, the inductive assumption is $\tilde{q} \in \Omega_k$ and $\tilde{P}(B(\tilde{q})) = P(B(\tilde{q})) > 0$. Now we define a set $c(\tilde{q}) \subseteq e(\tilde{q})$ with the following properties.

1.
$$e(\tilde{q}) \setminus c(\tilde{q}) \text{ is finite.}$$

2.
$$\forall q \in \big(e(\tilde{q}) \setminus c(\tilde{q})\big), P\big(B(q)\big) > 0.$$

3.
$$\sum_{q \in c(\tilde{q})} P(B(q)|B(\tilde{q})) < \frac{\epsilon}{3(2^{k+1})}.$$

Denote $\{\tilde{q}(1), ..., \tilde{q}(k), *\}$ by $\tilde{q}^\wedge *$. We define the collection of extensions of $\tilde{q}$ at stage $k+1$ to be $\tilde{e}(\tilde{q}) := (e(\tilde{q}) \setminus c(\tilde{q})) \cup \{\tilde{q}^\wedge *\}$. Define

$$\tilde{P}\big(B(q)|B(\tilde{q})\big) := P\big(B(q)|B(\tilde{q})\big) \ \forall q \in \big(e(\tilde{q}) \setminus c(\tilde{q})\big)$$

and

$$\tilde{P}\big(B(\tilde{q}^\wedge *)|B(\tilde{q})\big) := \sum_{q \in c(\tilde{q})} P\big(B(q)|B(\tilde{q})\big).$$

We see that every partial history in $\tilde{e}(\tilde{q})$ satisfies the inductive assumption at stage $k+1$. Then we can achieve $\tilde{\Omega}_n$ for all n and hence the history space $\tilde{\Omega}$. The existence of a probability measure $\tilde{P}$ follows from the Kolmogorov Extension Theorem (cf. [3], Theorem 36.2). We then construct the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}_\infty, \tilde{P})$ in a similar way to $(\Omega, \mathcal{F}_\infty, P)$.

We can modify $(\tilde{\Omega}, \tilde{\mathcal{F}}_\infty, \tilde{P})$ to a further probability space $(\Omega', \mathcal{F}'_\infty, P')$ (the adjusted probability space) which will be constructed from a sequence of sets of finitely many states. For all $\tilde{q} = (\tilde{q}(1), ..., \tilde{q}(k)) \in \tilde{\Omega}_k$, define the collection of the last coordinates of partial histories in $\tilde{e}(\tilde{q})$ by $\tilde{L}(\tilde{q})$. That is,

$$\tilde{L}(\tilde{q}) := \{h : \big(\tilde{q}(1), ..., \tilde{q}(k), h\big) \in \tilde{e}(\tilde{q})\}.$$

Define

$$J'_{k+1} := \bigcup_{\tilde{q} \in \tilde{\Omega}_k} \tilde{L}(\tilde{q}).$$

Note that for every $\tilde{q} \in \tilde{\Omega}_{<\infty}$, $\tilde{e}(\tilde{q})$ is finite, and hence $\tilde{L}(\tilde{q})$ is finite. It follows that J'_{k+1} is finite for every k, because $\tilde{\Omega}_k$ is finite. Define

$$\Omega' := \prod_k J'_k, \quad \Omega'_k := \prod_{m \leq k} J'_k, \quad \Omega'_{<\infty} := \bigcup_k \Omega'_k.$$

We see that the new measurable space $(\Omega', \mathcal{F}'_\infty)$ is constructed from the sequence $\{J'_k\}_{k \in \mathbb{N}}$ of sets of finitely many states. For every $q' \in \Omega'_{<\infty}$, let

$$P'\big(B(q')\big) := \begin{cases} \tilde{P}\big(B(q')\big) & \text{if } q' \in \tilde{\Omega}_{<\infty} \\ 0 & \text{if } q' \notin \tilde{\Omega}_{<\infty}. \end{cases}$$

The existence of probability measure P' again follows from the Kolmogorov Extension Theorem (cf. [3], Theorem 36.2).

Denote the collection of all bounded stopping times on $(\Omega', \mathcal{F}'_\infty, P')$ by $\mathcal{T}'$. For any $\omega = (\omega(1), \omega(2), ...)$ in Ω, define a *shadow history* $s(\omega)$ in Ω' as follows. If $\omega \in \Omega'$, let $s(\omega) = \omega$. If $\omega \notin \Omega'$, then there exists a unique new history $\omega' = (\omega'(1), \omega'(2), ..., \omega'(m), ...) \in \Omega'$ such that $\omega'(i) = \omega(i)\ \forall 1 \leq i < m$ and $\omega'(i) = *\ \forall i \geq m$, and we let $s(\omega) = \omega'$. Define y to be an injective mapping from $\mathcal{T}'$ to $\mathcal{T}$ such that $y(\theta')(\omega) = \theta'(s(\omega))$.

We still assume that $C = \{c_1, ..., c_k\}$. From the stochastic colouring f with values in C, we define the new stochastic colouring f' on $(\Omega', \mathcal{F}'_\infty)$ such that for all pairs (q', τ') with $q' \in \Omega'_{<\infty}$ and $\tau' \in \mathcal{T}'(q')$

$$f'(q', \tau') := \begin{cases} f\big(q', y(\tau')\big) & \text{if } q' = (h'_1, ..., h'_n \neq *) \\ c_1 & \text{if } q' = (h'_1, ..., h'_n = *). \end{cases} \tag{3.1}$$

Apply the Finite-state Finite-colour Stochastic Partition Theorem (Theorem 3.9) to obtain a sequence of bounded stopping times $0 \leq \theta'_1 < \theta'_2 < \theta'_3 < ...$ defined on $(\Omega', \mathcal{F}'_\infty, P')$ such that

$$P'(\hat{f}'_{\theta'_1, \theta'_2} = \hat{f}'_{\theta'_i, \theta'_j}\ \forall 1 \leq i < j) > 1 - 5\epsilon/8.$$

As

$$P(\Omega \cap \Omega') = P'(\Omega \cap \Omega') > 1 - \epsilon/3,$$

and for any Borel measurable set $S \subseteq \Omega \cap \Omega'$, $P(S) = P'(S)$, we find that

$$P(\hat{f}_{\theta_1, \theta_2} = \hat{f}_{\theta_i, \theta_j}\ \forall 1 \leq i < j) > 1 - \epsilon$$

where $\theta_n := y(\theta'_n)\ \forall n \in \mathbb{N}$. □

3.6 Finite stochastic Ramsey theorem

Before proving the Strong Stochastic Finitary Ramsey Theorem (Theorem 3.2), we use the Stochastic Infinitary Ramsey Theorem (Theorem 1.2) to prove the following result. This time $n(m, \epsilon, P)$ mentioned below depends on P, but not on f.

Theorem 3.13. (Stochastic Finitary Ramsey Theorem). *For a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence of sets of countably many* *states, a set C of finitely many colours, a natural number $m \geq 2$ and any $\epsilon > 0$, there exists $n = n(m, \epsilon, P) \in \mathbb{N}$ such that for every stochastic colouring f with values in C, there exist m bounded stopping times $\theta_1 < \theta_2 < ... < \theta_m \leq n$ which satisfy*

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j} \ \forall 1 \leq i < j \leq m) > 1 - \epsilon.$$

Proof. Recall from Definition 1.1 that

$$Z = \{(q, \tau) : q \in \Omega_{<\infty}, \tau \in \mathcal{T}(q)\}.$$

We firstly suppose that $(\Omega, \mathcal{F}_\infty, P)$ is constructed from the sequence $\{J_n\}_{n\in\mathbb{N}}$ where each J_n is finite. The set of all stochastic colourings f is $[C]^Z$: the finite set $[C]$ has the discrete topology, and $[C]^Z$ the product topology. By Tychonov's theorem, this space is compact. For each collection Q of m bounded stopping times $Q = \{\theta_1, \theta_2, ...\theta_m\}$ with $\theta_i < \theta_j \ \forall i < j$, we define the subset C_Q of stochastic colourings included in $[C]^Z$ to comprise those f with

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j} \ \forall 1 \leq i < j \leq m) > 1 - 5\epsilon/8.$$

We claim and now prove that C_Q is open in $[C]^Z$ for every such $Q = \{\theta_1, \theta_2, ..., \theta_m\}$.

For a stochastic colouring $f \in [C]^Z$, we write

$$A(f, Q) := \bigcap_{\omega\in\Omega, 1\leq i<j\leq m} \{g \in [C]^Z : g(\omega|\theta_i, \theta_j) = f(\omega|\theta_i, \theta_j)\}.$$

If f is in C_Q, then any stochastic colouring f' which agrees with f on Q, i.e., $f' \in A(f, Q)$, is also in C_Q. Given an ω in Ω and a pair of i and j with $1 \leq i < j \leq m$, it follows from the definition of open sets in a product space that $\{g \in [C]^Z : g(\omega|\theta_i, \theta_j) = f(\omega|\theta_i, \theta_j)\}$ is open in $[C]^Z$. As each J_n in $\{J_n\}_{n\in\mathbb{N}}$ is finite, $A(f, Q)$ is actually a finite intersection of these open sets in $[C]^Z$; this is becasue there are only finitely many vertices of the form $\omega|\theta_i$ for each bounded θ_i. Since $A(f, Q)$ is a subset of C_Q for each f in C_Q, $\bigcup_{f\in C_Q} A(f, Q) = C_Q$ is also open, which completes the proof of the claim.

The Stochastic Infinitary Ramsey Theorem (Theorem 1.2) asserts that these sets C_Q cover the whole of $[C]^Z$. By compactness, some finite collection $\{C_{Q_1}, ..., C_{Q_t}\}$ of these sets also covers $[C]^Z$. This implies that every stochastic colouring f on $\bigcup_{i=1}^{t} Q_i$ will have

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j} \ \forall 1 \le i < j \le m) > 1 - 5\epsilon/8$$

on at least one of the Q_i. Then we may take for n any bound of $\bigcup_{i=1}^{t} Q_i$ such that $\forall 1 \le l \le t \ \forall \theta_m \in Q_l$, $\theta_m < n$.

For the case that at least one J_n is countably infinite, we construct an adjusted probability space $(\Omega', \mathcal{F}'_\infty, P')$ from the sequence $\{J'_n\}_{n\in\mathbb{N}}$ of sets of finitely many states, as in the proof of Theorem 1.2 in Section 3.5. Recall the definitions of the new stochastic colouring f' on $(\Omega', \mathcal{F}'_\infty)$, the shadow history $s(\omega)$ and the injective mapping $y : \mathcal{T}' \to \mathcal{T}$ in the proof of Theorem 1.2. From the result for the case of finite J_n for all n above, the existence of an $n' = n'(m, 5\epsilon/8, P') \in \mathbb{N}$ follows such that for every new stochastic colouring g' defined in $(\Omega', \mathcal{F}'_\infty)$, there exists m bounded stopping times $\theta'_1 < \theta'_2 < ... < \theta'_m \le n'$ on Ω' with the property that

$$P'(\hat{g}'_{\theta'_1,\theta'_2} = \hat{g}'_{\theta'_i,\theta'_j} \ \forall 1 \le i < j \le m) > 1 - 5\epsilon/8. \tag{3.2}$$

We now show that we can regard this n' as a qualified bound $n(m, \epsilon, P)$ for $(\Omega, \mathcal{F}_\infty, P)$. Given a stochastic colouring f defined in $(\Omega, \mathcal{F}_\infty)$, we obtain the corresponding new stochastic colouring f' defined on $(\Omega', \mathcal{F}'_\infty)$ and a sequence of bounded stopping times $\theta'_1 < \theta'_2 < ... < \theta'_m \le n'$ which satisfies the property (3.2). Define a sequence of bounded stopping times $\theta_1 < \theta_2 < ... < \theta_m$ on Ω where $\theta_i := y(\theta'_i) \ \forall 1 \le i \le m$. Recall that

$$f(q', y(\tau')) = f'(q', \tau') \ \forall q' \in \Omega_{<\infty} \cap \Omega'_{<\infty}, \ \tau' \in \mathcal{T}'(q'),$$

as shown in (3.1). So for any ω with $s(\omega)|\theta'_{m-1} \in \Omega_{<\infty}$,

$$\hat{f}_{\theta_i,\theta_j}(\omega) = \hat{f}'_{\theta'_i,\theta'_j}(s(\omega)), \ \forall 1 \le i < j \le m.$$

Recall at the end of the proof of Theorem 1.2 that $P(\Omega \cap \Omega') = P'(\Omega \cap \Omega') > 1 - \epsilon/3$, and for any Borel measurable set $S \subseteq \Omega \cap \Omega'$, we have $P(S) = P'(S)$. Hence

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j} \ \forall 1 \le i < j \le m) > 1 - \epsilon,$$

and it follows from the definition of y that $\theta_m \leq n'$. $\square$

To prove the Strong Stochastic Finitary Ramsey Theorem (Theorem 3.2), we first review the weak-* topology defined in a space of probability measures. (cf. [9], [5]) This will help us to verify the following lemma and later to check easily that certain sets are open in the weak-* topology.

Recall that a topological space X admitting a countable dense set is called *separable*; a space X is *complete metrizable* if it admits a compatible metric d such that (X, d) is complete; a separable complete metrizable space is called *Polish*. Let X be a compact metrizable space and Y a metrizable space. We denote by $C(X, Y)$ the space of continuous functions from X to Y with the topology induced by the *sup* or *uniform metric*

$$d_u(f, g) = \sup_{x \in X} d_Y\big(f(x), g(x)\big) \ \forall f, g \in C(X, Y).$$

where d_Y is a compatible metric for Y. If Y is a Polish space, then $C(X, Y)$ is also Polish (cf. (4.19) in [9]). Recall that a *Banach space* is a normed linear space which is complete in the metric defined by its norm. So, for a compact metrizable space X, $C(X, \mathbb{R})$ is a Banach space with norm $\|f\|_\infty = \sup_{x \in X} |f(x)|$, whose associated metric in $C(X, \mathbb{R})$ is $d_u(f, g) = \|f - g\|_\infty$.

For separable Banach spaces X and Y, we denote by $L(X, Y)$ the Banach space of bounded linear operators $T : X \to Y$ with norm $\|T\| = \sup\{\|Tx\| : x \in X, \|X\| \leq 1\}$. Denote by $L_1(X, Y)$ the unit ball $\{T \in L(X, Y) : \|T\| \leq 1\}$ of $L(X, Y)$. The *strong topology* on $L(X, Y)$ is the topology generated by the family of point evaluation maps $e_x(T) = Tx$, $e_x : L(X, Y) \to Y$, for $x \in X$. It has a basis consisting of the sets of the form

$$V_{x_1, \ldots, x_n; \epsilon; T} = \{S \in L(X, Y) : \|Sx_1 - Tx_1\| < \epsilon, \ldots, \|Sx_n - Tx_n\| < \epsilon\},$$

for $x_1, \ldots, x_n \in X$, $\epsilon > 0$, $T \in L(X, Y)$. The unit ball $L_1(X, Y)$ with the strong topology is Polish.

Denote by $C_b(X)$ the set of bounded continuous real-valued functions on X. Recall that a *linear functional* Λ on $C_b(X)$ is a map $\Lambda : C_b(X) \to \mathbb{R}$ such that for any two constants α and β and any two elements f and g of $C_b(X)$ the equation

$$\Lambda(\alpha f + \beta g) = \alpha \Lambda(f) + \beta \Lambda(g)$$

holds. A linear functional Λ is called *positive* if $\Lambda(f) \geq 0$ whenever $f \geq 0$. Let X be a separable Banach space. The *dual space* X^* of X is the Banach space of all bounded linear functionals $x^* : X \to \mathbb{R}$, with norm $\|x^*\| = \sup\{x^*(x) : x \in X, \|x\| \leq 1\}$. So $X^* = L(X, \mathbb{R})$. The strong topology on X^* is the topology generated by functions $x^* \mapsto x^*(x)$, for $x \in X$, and is called the *weak-* topology of* X^*. Let $B_1(X^*) = L_1(X, \mathbb{R})$ be the unit ball of X^*, then $B_1(X^*)$ with the weak-* topology is Polish and compact. (See Theorem 4.7 in [9].)

Back to the condition in Theorem 3.2, the measurable space $(\Omega, \mathcal{F}_\infty)$ is supposed to be constructed from a sequence $\{J_i\}_{i \in \mathrm{N}}$ where each J_i is finite. Then Ω is a compact Polish space. We denote by $\mathcal{M}$ the space of all probability measures defined on this $(\Omega, \mathcal{F}_\infty)$. Let I denote the function which takes the value 1 everywhere, i.e., a constant function. Given any probability measure $\mu \in \mathcal{M}$ the functional $\Lambda_\mu : f \mapsto \int f d\mu$ can be seen as a positive linear functional on $C_b(\Omega)$ with $\Lambda_\mu(I) = 1$. We endow $\mathcal{M}$ with the topology generated by the maps $\mu \mapsto \int f d\mu$ where f varies over $C_b(\Omega)$. A base of open neighbourhoods for any μ_0 in $\mathcal{M}$ is of the form

$$V_{\mu_0}(f_1, f_2, ..., f_k; \epsilon_1, ..., \epsilon_k) = \{\mu \in \mathcal{M} : |\int f_i d\mu - \int f_i d\mu_0| < \epsilon_i \ \forall 1 \leq i \leq k\}$$

where $f_1, f_2, ..., f_k$ are in $C_b(\Omega)$ and $\epsilon_1, \epsilon_2, ..., \epsilon_k$ are positive numbers. We shall call this the weak topology in $\mathcal{M}$, since we can view it as the relative topology of the weak-* topology of $B_1(C(\Omega, \mathbb{R})^*)$ restricted to $\mathcal{M}$.

Lemma 3.14. *If the measurable space $(\Omega, \mathcal{F}_\infty)$ is constructed from $\{J_n\}_{n \in \mathrm{N}}$ where each J_n is finite, then the space $\mathcal{M}$ of probability measures on $(\Omega, \mathcal{F}_\infty)$ is compact metrizable.*

Proof. From the condition in the lemma, we find that Ω is a compact Polish space and $C(\Omega, \mathbb{R})$ is a separable Banach space. The unit ball $B_1(C(\Omega, \mathbb{R})^*)$ with the weak-* topology is compact metrizable. Let K be the set of positive linear functionals on $C_b(\Omega)$ with $\Lambda(I) = 1 \ \forall \Lambda \in K$. Then K is closed in $B_1(C(\Omega, \mathbb{R})^*)$, and hence compact metrizable. By Theorem 5.8 on Page 38 of [12], for any $\Lambda \in K$, there exists a unique measure $\mu \in \mathcal{M}$ such that

$$\Lambda(f) = \int f d\mu \ \forall f \in C_b(\Omega).$$

In fact, it is a homeomorphism of K with $\mathcal{M}$. Thus $\mathcal{M}$ is compact metrizable. $\square$

Comment: We can also use the Riesz Representation Theorem (cf. Theorem 2.14 in [14]) to prove the existence of a bijection $\Lambda \leftrightarrow \mu$ between K and $\mathcal{M}$ in the sence that $\Lambda(f) = \int f d\mu \ \forall f \in C_b(\Omega)$.

We now employ a similar procedure to prove the Strong Stochastic Finitary Ramsey Theorem (Theorem 3.2).

Proof of Theorem 3.2. Recall again the definition of Z. The set of all stochastic colourings f is $[C]^Z$: the finite set $[C]$ has the discrete topology, and $[C]^Z$ the countable product topology. By Tychonov's theorem, this space is compact. Ω is a Cantor space, and hence also a compact metrizable space. By Lemma 3.14, the set $\mathcal{M}$ of probability measures on $(\Omega, \mathcal{F}_\infty)$ with the weak topology is a compact metrizable space. Hence, the product $\mathcal{M} \times [C]^Z$ is also compact. For each collection Q of m bounded stopping times $Q = \{\theta_1, \theta_2, ..., \theta_m\}$ with $\theta_i < \theta_j \ \forall i < j$, we define the subset S_Q included in $\mathcal{M} \times [C]^Z$ such that for each (P, f) in S_Q

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j} \ \forall 1 \le i < j \le m) > 1 - \epsilon.$$

We now prove that S_Q is open in $\mathcal{M} \times [C]^Z$ for every such $Q = \{\theta_1, ..., \theta_m\}$. We copy the definition of the class $A(f, Q)$ for a stochastic colouring f in $[C]^Z$ from the proof of Theorem 3.13. That is, we again put

$$A(f, Q) := \bigcap_{\omega \in \Omega, 1 \le i < j \le m} \{g \in [C]^Z : g(\omega | \theta_i, \theta_j) = f(\omega | \theta_i, \theta_j)\}.$$

From the claim in Theorem 3.13, it follows that $A(f, Q)$ is open in $[C]^Z$ for each f. For a stochastic colouring f, we write $U_f = \{\omega \in \Omega : \hat{f}_{\theta_1,\theta_2}(\omega) = \hat{f}_{\theta_i,\theta_j}(\omega) \ \forall 1 \le i < j \le m\}$, and denote $\{P \in \mathcal{M} : P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j} \ \forall 1 \le i < j \le m) = 1\}$ by K_f. That means $\int \mathbf{1}_{U_f} dP = 1 \ \forall p \in K_f$, where $\mathbf{1}$ is the indicator function. By the definition of the weak-* topology in $\mathcal{M}$, for any P in K_f, $S(P) := \{\mu \in \mathcal{M} : |\int \mathbf{1}_{U_f} d\mu - \int \mathbf{1}_{U_f} dP| < \epsilon\}$ is an open set in $\mathcal{M}$, because $\mathbf{1}_{U_f}$ is a continuous function from Ω to $\{0, 1\}$. Thus $M(f) :=$

$\bigcup_{P\in K_f} S(P)$ is open in $\mathcal{M}$. It follows that $S_Q = \bigcup_{f\in[C]^Z}(M(f) \times A(f,Q))$ is open in $\mathcal{M} \times [C]^Z$.

The Stochastic Infinitary Ramsey Theorem (Theorem 1.2) asserts that these sets cover the whole of $\mathcal{M} \times [C]^Z$. By compactness, some finite collection $\{S_{Q_1}, S_{Q_2}, ..., S_{Q_t}\}$ of these sets also covers $\mathcal{M} \times [C]^Z$. This implies that every pair (P, f) comprising of a probability measure P and a stochastic colouring f on $\bigcup_{i=1}^t Q_i$ will have

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j} \ \forall 1 \leq i < j \leq m) > 1 - \epsilon$$

on at least one of the Q_i. Then we may take for n any bound of $\bigcup_{i=1}^t Q_i$ such that $\forall 1 \leq l \leq t \ \forall \theta_m \in Q_l$, $\theta_m < n$. □

Comment: It is unknown whether Theorem 3.2 is still true for a measurable space $(\Omega, \mathcal{F}_\infty)$ constructed from a sequence of sets of countably many states. However, we make a conjecture that it is false with the following two reasons. For simplicity, suppose $\Omega = \prod_n J_n$ where $J_n = \mathbb{N}$ for all n. Then Ω is no longer a compact metrizable space. So the set $\mathcal{M}$ of probability measures on $(\Omega, \mathcal{F}_\infty)$ with the weak topology is not a compact metric space, which invalidates the approach of the proof of Theorem 3.2. Furthermore, so far all the proofs of stochastic Ramsey statements with respect to this $\Omega = \prod_n J_n$ need to refer to an adjusted probability space $(\Omega', \mathcal{F}'_\infty, P')$, where the P' entirely depends on the probability measure P on $(\Omega, \mathcal{F}_\infty)$. Since P is arbitrary in the conjecture, we cannot go further to reduce the problem to the one in a measurable space constructed from a sequence of sets of finitely many states as in Theorem 3.13 before.

We can see this conjecture from another point of view. For each i in $\mathbb{N}$, we suppose $\alpha_i := \prod_n J_n$ where $J_n = \{1, 2, ..., i\}$ $\forall n$. Given a natural number $m > 1$ and an $\epsilon > 0$, we define n_i to be $\min[n(m, \epsilon)]$ associated with α_i in the condition of Theorem 3.2. That is, for every probability measure P on $(\Omega_i, \mathcal{F}_{i_\infty})$ and a stochastic colouring f, there exist m bounded stopping times $\theta_1 < \theta_2 < ... < \theta_m < n_i$ on Ω_i such that

$$P(\hat{f}_{\theta_1,\theta_2} = \hat{f}_{\theta_i,\theta_j} \ \forall 1 \leq i < j \leq m) \leq 1 - \epsilon.$$

It is an open problem whether for each pair (m, ϵ) the sequence $\{n_i\}_{i\in\mathbb{N}}$ is strictly increasing. An affirmative answer would be enough to prove the conjecture.

3.7 Further comments

1. (An alternative marking process which is symmetric) It may be a slightly more elegant way to define the marking process of Section 3.4 as follows. Under a stochastic colouring f, for a vertex q which satisfies condition II of Definition 3.5, we adopt a more natural formulation in analogy with the standard proof of the classical Ramsey theorem. Instead of searching for a bound N for each infinite increasing sequence of bounded stopping times $\{\delta_i\}$ such that $f(q, \delta_i) = \text{Blue}\ \forall i \geq N$, we describe the colouring properties of such a $\{\delta_i\}_{i\in\mathbb{N}}$ in a symmetric way, by replacing the original condition II by a statement of the following format (in which M is an appropriate finite subset of $V_{\mathbb{G}}$).

   There exists a covering subforest $\bar{\mathbb{G}}$ of $\mathbb{G}$ such that for every covering subforest $\mathbb{G}'$ of $\bar{\mathbb{G}}$, there exists an infinite increasing sequence of bounded stopping times $\{\delta_i\}$ adapted to $\mathbb{G}'$ such that

   $$f_{\bar{q},\delta_l} = \text{Red}\ \forall \bar{q} \in M\ \forall l > 0$$

   and

   $$f_{q,\delta_l} = \text{Blue}\ \forall l > 0.$$

   To put this on a rigorous footing, we give below a new marking process following this approach. Note that this necessitates an alternative definition of a well-structured subtree (with a more informative name) and some associated definitions and lemmas detailed below. The general idea after that is still the same as that in Section 3.4.

   Definition 3.15. *Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$ and a vertex q in $V_{\mathbb{G}}$, say that a sequence of bounded stopping times $\delta_1 < \delta_2 < \ldots$ makes q c-rich over $\mathbb{G}$, if all δ_l are consistent with q, adapted to $\mathbb{G}$, and satisfy $f(q, \delta_l) = c\ \forall l > 0$.*

Definition 3.16. *For a covering subforest $\mathbb{G}$ of $\mathbb{F}$, suppose given two finite sets M_R, M_B of vertices included in $V_{\mathbb{G}}$, for every covering subforest $\mathbb{G}'$ of $\mathbb{G}$, there exists an infinite sequence of bounded stopping times $\delta_1 < \delta_2 < \ldots$ adapted to $\mathbb{G}'$ that makes all vertices in M_R Red-rich over $\mathbb{G}$ and all vertices in M_B Blue-rich over $\mathbb{G}$. We then say (M_R, M_B) is {Red,Blue}-richly layered over $\mathbb{G}$.*

Definition 3.17. *We call a covering subforest $\mathbb{A}$ {Red,Blue}-richly layered if there is a partition of its vertex set $V_{\mathbb{A}} = R(\mathbb{A}) \cup B(\mathbb{A})$ such that for any finite subset $S \subset V_{\mathbb{A}}$, $(S \cap R(\mathbb{A}), S \cap B(\mathbb{A}))$ is {Red,Blue}-richly layered over $\mathbb{A}$.*

Lemma 3.18. *Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$ and two finite subsets $M_R, M_B \subset V_{\mathbb{G}}$ with (M_R, M_B) {Red,Blue}-richly layered over $\mathbb{G}$, for any vertex $q \in V_{\mathbb{G}} \setminus (M_R \cup M_B)$, it is true either that $(\{q\} \cup M_R, M_B)$ is {Red,Blue}-richly layered over $\mathbb{G}$, or that, for each covering subforest $\mathbb{G}'$ of $\mathbb{G}$, $(M_R, \{q\} \cup M_B\})$ is {Red,Blue}-richly layered over $\mathbb{G}'$.*

Proof. This is straightforward as an equivalent statement to condition II in Definition 3.5. □

Definition 3.19 (Marking process). *Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$, two finite subsets M_R and M_B of $V_{\mathbb{G}}$ with (M_R, M_B) {Red,Blue}-richly layered over $\mathbb{G}$, for a vertex $q \in V_{\mathbb{G}} \setminus (M_R \cup M_B)$, mark q with symbol r relative to (M_R, M_B) in $\mathbb{G}$ if $(\{q\} \cup M_R, M_B)$ is {Red,Blue}-richly layered over $\mathbb{G}$. Otherwise mark q with symbol b relative to (M_R, M_B) in $\mathbb{G}$.*

2. (Stochastic Ramsey numbers) As indicated in Section 1.1, the stochastic Ramsey theorem is an extension of Ramsey's theorem. So it is natural to ask whether the stochastic versions of other theorems from Ramsey theory exist. Besides, we may enquire about the stochastic version of Ramsey numbers. According to the Strong Stochastic Finitary Ramsey Theorem (Theorem 3.2), a uniform bound n exists for all stochastic colourings f and all probability measures defined on $(\Omega, \mathcal{F}_\infty)$, once C, $(\Omega, \mathcal{F}_\infty)$, m and $\epsilon > 0$ are fixed. It may be difficult

to determine the lowest bound in a general case at the first attempt, but we can try a basic example first. Assume $C = \{\text{Red}, \text{Blue}\}$ and assume $(\Omega, \mathcal{F}_\infty)$ as in the model of tossing a coin infinitely many times in Section 3.1, which implies that $\mathbb{F}$ is a binary tree. For $m > 2$ and $\epsilon > 0$, we define $R(m, \epsilon)$ to be the lowest bound n with respect to these C, $(\Omega, \mathcal{F}_\infty)$, m and ϵ. It is an interesting problem to try determining $R(m, \epsilon)$.

4

On a stochastic extension for hypergraphs

4.1 Stochastic extension for hypergraphs

Before the discussion, we first establish some definitions.

Recall that in Section 1.1, we firstly introduce Ω to be $\prod_n J_n$, and assign it with the Tychonoff product topology. Then we define $\mathcal{F}_n$ to be the σ-algebra on Ω of Borel sets generated by the basic open sets with support included in $\{1, 2, ..., n\}$. There is another quick way to construct $\mathcal{F}_n$ without directly referring to the topology of Ω.

The coordinate process $X = \{X_1, X_2, ...\}$ on Ω is just the sequence of coordinate functions defined, for $\omega = (\omega(1), \omega(2), ...)$ and $n = 1, 2, ...$, by $X_n(\omega) = \omega(n)$. Then we find, for each n, $\mathcal{F}_n$ is exactly the σ-algebra generated by $(X_1, X_2, ..., X_n)$. Furthermore, $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$. We can generalize this concept with respect to any stopping time $\theta : \Omega \to \{0, 1, 2, ..., +\infty\}$. For a stopping time θ, define $X_n^\theta(\omega) := X_{\min\{n, \theta(\omega)\}}(\omega)$. Then define $\mathcal{F}_\theta := \sigma(X_1^\theta, X_2^\theta, ...)$. When θ is a bounded stopping time, $\mathcal{F}_\theta = \sigma(X_\theta)$, where X_θ is a random variable $\omega \mapsto X_{\theta(\omega)}(\omega)$. From another perspective, $\mathcal{F}_\theta$ is the σ-algebra generated by the partial histories truncated at θ, i.e.

$$\mathcal{F}_\theta = \sigma\{B(q) : q \in S_\theta\}.$$

For a set X, denote by $X_<^{(k)}$ the collection of ordered subsets of size k of X, so specifically,

$$\mathbb{N}_<^{(k)} = \{(n_1, ..., n_k) : n_l \in \mathbb{N}\ \forall 1 \leq l \leq k;\ n_i < n_j\ \forall 1 \leq i < j \leq k\};$$

$$\mathcal{T}_<^{(k)} = \{(\delta_1, ..., \delta_k) : \delta_l \in \mathcal{T}\ \forall 1 \leq l \leq k;\ \delta_i < \delta_j\ \forall 1 \leq i < j \leq k\}.$$

We drop the subscript '$<$' when context allows.

Let us recall the classical infinitary Ramsey theorem. Let C be a finite set whose elements are colours. The classical Ramsey theorem says that, given a natural number k, for every function that assigns a colour $c(m_1, ..., m_k) \in C$ to every element $(m_1, ..., m_k)$ in $\mathbb{N}^{(k)}$, there is a sequence of integers $n_1 < n_2 < ...$ such that $c(n_1, ..., n_k) = c(n_{l_1}, ...n_{l_k})$ for all $(l_1, ..., l_k) \in \mathbb{N}^{(k)}$. To represent this result in a complete graph $G = (V, E)$, we let k be the number of vertices in each edge, i.e., the size of each edge, so that $E = V^{(k)}$. Theorem 2.1 tells us that if we have a C-colouring of the edges of an infinite complete graph $G = (V, V^{(2)})$, then we can find an infinite set of vertices $S \subset V$ spanning a monochromatic complete subgraph (clique) $(S, S^{(2)})$. As mentioned at the start of Chapter 2, Theorem 2.1 is only the special case $k = 2$ of the classical Ramsey theorem. Recall that when $|J_i| = 1\ \forall j \in \mathbb{N}$, the Stochastic Infinitary Ramsey Theorem (Theorem 1.2) reduces to the classical Ramsey theorem (Theorem 2.1), which is the case $k = 2$ here. One natural question is whether a stochastic extension of Ramsey's theorem exists for $k > 2$. That is, can we find a stochastic Ramsey theorem corresponding to the classical Ramsey theorem for hypergraphs?

To answer this question, we need first to define appropriately a more general stochastic colouring f.

Definition 4.1. *Given a set C of colours, a stochastic hyper-colouring f^k of order k for $k \in \{2, 3, ...\}$ is a mapping from $Z^k := \{(q, \delta_1, ..., \delta_{k-1}) : q \in \Omega_{<\infty};\ \delta_l \in \mathcal{T}(q)\ \forall 1 \leq l < k;\ (\delta_1, ..., \delta_{k-1}) \in \mathcal{T}^{(k-1)}\}$ to C. (We omit the words 'of order k' whenever the order is clear from context, e.g. from the superscript in use.) The induced stochastic hyper-colouring $\hat{f}^k$ of f^k is a mapping from $\mathcal{T}^{(k)} \times \Omega$ to C and defined by*

$$\hat{f}^k_{\delta_1, ..., \delta_k}(\omega) := f^k(\omega | \delta_1, \delta_2, ..., \delta_k)\ \forall (\delta_1, ..., \delta_k) \in \mathcal{T}^{(k)}\ \forall \omega \in \Omega. \tag{4.1}$$

So, f^2 corresponds directly to the stochastic colouring f of Definition 1.1. Here note that for the bounded stopping time $\delta_l > \delta_1$ (with $1 < l \leq k$) δ_l is again consistent with $\omega|\delta_1$, as

$$\delta_l(\omega) > \|\omega|\delta_1\| = \delta_1(\omega) \ \forall \omega \in \Omega.$$

There are other possible ways to define the notion of a stochastic hyper-colouring f^k. For instance, we can require f^k to be a mapping from Z'^k to C with Z'^k defined as

$$Z'^k := \{(q_1, ..., q_{k-1}, \delta) : \ B(q_i) \supseteq B(q_j) \ \forall 1 \leq i < j < k; \ \delta \in \mathcal{T}(q_{k-1})\}. \tag{4.2}$$

That is, for any $(q_1, ..., q_{k-1}, \delta) \in Z'^k$, there exists ω' in Ω and a finite sequence $(n_1, ..., n_{k-1})$ in $\mathbb{N}^{(k-1)}$ such that $q_l = \omega'|n_l \ \forall 1 \leq l < k$. The corresponding induced mapping is then

$$\hat{f}^k_{\delta_1,...,\delta_k}(\omega) := f^k(\omega|\delta_1, ..., \omega|\delta_{k-1}, \delta_k) \ \forall(\delta_1, ..., \delta_k) \in \mathcal{T}^{(k)} \ \forall \omega \in \Omega. \tag{4.3}$$

An intermediate version combining both notions is also possible: for $1 < i < k-1$, one may consider mappings from the collection $\{(q_1, ..., q_i, \delta_{i+1}, ..., \delta_k)\}$ to C with the 'consistency' condition:

$$B(q_m) \supseteq B(q_n) \ \forall 1 \leq m < n \leq i; \ \delta_{i+1} \in \mathcal{T}(q_i); \ (\delta_{i+1}, ..., \delta_k) \in \mathcal{T}^{(k-i)}.$$

However, we prefer Definition 4.1 for the following reasons. By Definition 4.1, given k bounded stopping times $(\delta_1, ..., \delta_k) \in \mathcal{T}^{(k)}$, the induced stochastic hyper-colouring $\hat{f}^k$ is an $\mathcal{F}_{\delta_1}$-measurable random variable from Ω to C. That is, the value of $\hat{f}^k_{\delta_1,...,\delta_k}(\omega)$ is determined at stage $\|\omega|\delta_1\|$. If we had found an infinite sequence of bounded stopping times $\theta_1 < \theta_2 < ...$ defined in the probability space $(\Omega, \mathcal{F}_\infty, P)$ such that

$$P(\hat{f}^k_{\theta_1,...,\theta_k} = \hat{f}^k_{\theta_{n_1},...,\theta_{n_k}} \ \forall(n_1, n_2, ..., n_k) \in \mathbb{N}^{(k)}_<) > 1 - \epsilon,$$

then, with probability greater than $1 - \epsilon$, the random colours assigned to $(\theta_{n_1}, ..., \theta_{n_k})$ for all $(n_1, ..., n_k) \in \mathbb{N}^{(k)}_<$ by $\hat{f}^k$ would be revealed to be one fixed colour in C at stage $\theta_1(\omega)$.

In the other formulations of f^k, we have to wait for bounded stopping time θ_i with $i > 1$ to see the random colour determined. For instance, consider $f^k : Z'^k \to C$ in (4.2). Given k bounded stopping times $(\delta_1, ..., \delta_k) \in \mathcal{T}^{(k)}$, its induced stochastic hyper-colouring $\hat{f}^k$ is an $\mathcal{F}_{\delta_{k-1}}$-measurable random variable. So, if we had found an infinite sequence of bounded stopping times $\theta_1 < \theta_2 < ...$ defined in the probability space $(\Omega, \mathcal{F}_\infty, P)$ such that

$$P(\hat{f}^k_{\theta_1,...,\theta_k} = \hat{f}^k_{\theta_{n_1},...,\theta_{n_k}} \ \forall (n_1, n_2, ..., n_k) \in \mathbb{N}^{(k)}) > 1 - \epsilon,$$

then, with probability greater than $1 - \epsilon$, the random colours assigned to $(\theta_{n_1}, ..., \theta_{n_k})$ for all $(n_1, ..., n_k) \in \mathbb{N}^{(k)}$ by $\hat{f}^k$ would be revealed to be one fixed colour in C at stage $\theta_{k-1}(\omega)$. That means the initial segment of partial histories $(\omega|\theta_1, ..., \omega|\theta_{k-2})$ is not enough to determine the colour $\hat{f}^k_{\theta_{n_1},...,\theta_{n_k}}$ for any $(n_1, ..., n_k) \in \mathbb{N}^{(k)}$.

Of course, one can replace this $\{\theta_i\}_{i\in\mathbb{N}}$ by $\{\theta'_i\}_{i\in\mathbb{N}}$ with $\theta'_i := \theta_{i+(k-2)} \ \forall i$. Then, with probability greater than $1 - \epsilon$, the random colour assigned to the sequence $\{\theta'_i\}_{i\in\mathbb{N}}$ by $\hat{f}^k$ is fixed from stage $\theta'_1(\omega)$ this time. However, this trick may run into trouble when we consider the lowest bound on the Stochastic Finitary Ramsey Theorem (Theorem 3.13). Suppose that the following claim could be proved. For a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence of sets of states, a set C of finitely many colours, a natural number $m \geq 2$ and any $\epsilon > 0$, there exists $n = n(m, \epsilon, P) \in \mathbb{N}$ such that for every stochastic hyper-colouring f^k with values in C, there exists m bounded stopping times $\theta_1 < \theta_2 < ... < \theta_m \leq n$ which satisfy

$$P(\hat{f}^k_{\theta_1,...,\theta_k} = \hat{f}^k_{\theta_{n_1},...,\theta_{n_k}} \ \forall 1 \leq n_1 < n_2 < ... < n_k \leq m) > 1 - \epsilon.$$

Then we need to set a standard for θ_1 in this sequence: does it need the requirement that with probability greater than $1-\epsilon$, the random colour can be determined from stage $\theta_1(\omega)$, i.e., $\hat{f}^k_{\theta_1,...,\theta_k} : \Omega \to C$ is $\mathcal{F}_{\theta_1}$-measurable?

After clearing this ambiguity, we still need to bear in mind that, under the definition $f^k : Z'^k \to C$ in (4.2), when we calculate $\hat{f}^k$ from f^k, a bounded stopping time θ_i in $\{\theta_i\}_{1\leq i<k}$ can only provide a partial history truncated by θ_i in f^k. That is, for any $(\theta_{n_1}, ..., \theta_{n_k}) \subseteq \{\theta_i\}_{i\in\mathbb{N}}$, in the formulation

$$\hat{f}^k_{\theta_{n_1},...,\theta_{n_k}}(\omega) = f^k(\omega|\theta_{n_1}, ..., \omega|\theta_{n_{k-1}}, \theta_{n_k}) \ \forall \omega \in \Omega,$$

n_k has to be at least k.

Thus, given an increasing sequence $\{\theta_i\}_{i\in\mathbb{N}}$ of bounded stopping times, the sequence $\{\omega|\theta_i\}_{i<k-1}$ of partial histories generated from $\{\theta_i\}_{i<k-1}$ is not enough to determine the random colour assigned to any $(\theta_{n_1},...,\theta_{n_k}) \subset \{\theta_i\}_{i\in\mathbb{N}}$, and any element in $\{\theta_i\}_{1\leq i<k}$ can never present as a bounded stopping times in the formulation of f^k for $\hat{f}^k$. Therefore, it is natural to adopt Definition 4.1 for further study on the existence of stochastic Ramsey's theorem on f^k with $k > 2$.

The natural generalisation of the stochastic Ramsey theorem on f^k is as follows. Given a natural number $k > 1$, a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence $\{J_i\}_{i\in\mathbb{N}}$ of sets of states and a stochastic hyper-colouring f^k with values in a finite set C, for every $\epsilon > 0$ there exists an increasing sequence of bounded stopping times $\theta_1 < \theta_2 < \theta_3 <...$ such that

$$P(\hat{f}^k_{\theta_1,\theta_2,...,\theta_k} = \hat{f}^k_{\theta_{n_1},\theta_{n_2},...,\theta_{n_k}} \ \forall(n_1,n_2,...,n_k) \in \mathbb{N}^{(k)}) > 1-\epsilon. \qquad (4.4)$$

Unfortunately, proving this generalisation does not seem to be straightforward, or maybe the proof needs more technique than those we have used so far. We leave it as an open problem. However, we can still obtain some results weaker than (4.4). For example, when $k = 3$, we have the following theorem.

Theorem 4.2. *Given a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence $\{J_i\}_{i\in\mathbb{N}}$ of sets of states and a stochastic hyper-colouring f^3 with values in a finite set C, then for every $\epsilon > 0$ there exists an increasing sequence of bounded stopping times $\theta_1 < \theta_2 < \theta_3 <...$ such that*

$$P(\hat{f}^3_{\theta_1,\theta_2,\theta_3} = \hat{f}^3_{\theta_i,\theta_{2j},\theta_{2j+1}} \ \forall i,j \in \mathbb{N} \text{ with } 1 \leq i < 2j) > 1-\epsilon.$$

We shall explain later in Section 4.2 why we can prove the stochastic Ramsey theorem for f^2, but not for any f^k with $k > 2$ without the weakening above.

For the proof of Theorem 4.2, we adopt the same tree model $\mathbb{F}$ as in Section 3.3, and confine ourselves to the case $|C| = 2$ and J_i finite for all i, since, by the approach introduced earlier in Theorem 3.9 and Theorem 1.2, it can be extended to the full result. So, we only need to prove the following theorem analogous to Theorem 3.1.

Theorem 4.3. *Given a set $C = \{\text{Red}, \text{Blue}\}$, a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence of sets each containing finitely many states and a stochastic hyper-colouring f^3 with values in C, then for every $\epsilon > 0$, there exists a natural number N, two sets $S_1, S_2 \in \mathcal{F}_N$ with $S_1 \cup S_2 = \Omega$, and a sequence of bounded stopping times $\theta_1 < \theta_2 < \ldots$ such that*

$$P(\hat{f}^3_{\theta_m,\theta_{2n},\theta_{2n+1}} = \text{Red}, \ \forall 1 \leq m < 2n | S_1) > 1 - \epsilon, \quad \textit{if } P(S_1) > 0,$$

and

$$P(\hat{f}^3_{\theta_m,\theta_{2n},\theta_{2n+1}} = \text{Blue}, \ \forall 1 \leq m < 2n | S_2) > 1 - \epsilon, \quad \textit{if } P(S_2) > 0.$$

Now we need to find an analogous well-structured subtree $\mathbb{A}$ from the tree of histories $\mathbb{F}$ under the stochastic hyper-colouring f^3 by an adaption of the marking scheme of Definition 3.5.

Definition 4.4. *A well-structured subtree $\mathbb{A}$ under the stochastic hyper-colouring f^3 is a covering subtree $\mathbb{A}$ of $\mathbb{F}$ with a partition of its vertex set $V_\mathbb{A} = R(\mathbb{A}) \cup B(\mathbb{A})$ such that for any covering subforest $\mathbb{G}$ of $\mathbb{A}$ and any finite subset $S \subset V_\mathbb{G}$, there exists an infinite sequence of bounded stopping times $\delta_1 < \delta_2 < \ldots$ adapted to $\mathbb{G}$ with the property*

$$f^3(q, \delta_{2l-1}, \delta_{2l}) = \text{Red} \ \ \forall q \in S \cap R(\mathbb{A}) \ \forall l > 0$$

and

$$f^3(q, \delta_{2l-1}, \delta_{2l}) = \text{Blue} \ \ \forall q \in S \cap B(\mathbb{A}) \ \forall l > 0.$$

Definition 4.5. (Marking scheme) *Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$ and a finite set M of vertices with $M \subseteq V_\mathbb{G}$, for every vertex q in $V_\mathbb{G} \setminus M$, mark q with symbol r relative to M in $\mathbb{G}$, if for any covering subforest $\mathbb{G}'$ of $\mathbb{G}$,* *there exists an infinite sequence of bounded stopping times* $\delta_1 < \delta_2 < \delta_3 < \ldots$ *adapted to* $\mathbb{G}'$ *such that*

$$f^3(q, \delta_{2l-1}, \delta_{2l}) = f^3(\bar{q}, \delta_{2l-1}, \delta_{2l}) = \text{Red } \forall l > 0 \ \forall \bar{q} \in M.$$

Mark q *with symbol* b *relative to* M *in* $\mathbb{G}$ *if* q *cannot be marked with* r *relative to* M *in* $\mathbb{G}$.

Comment: One way to understand the similarity between Definition 3.4 and Definition 4.4 is by regarding every successive pair of bounded stopping times $(\delta_{2i-1}, \delta_{2i}) \subset \{\delta_i\}_{i\in\mathbb{N}}$ as a special 'bounded stopping strip' δ_i', and to consider the hyper-colouring $f'^2(q, \delta_i') := f^3(q, \delta_{2i-1}, \delta_{2i})$. The same idea applies to the transition from Definition 3.5 to Definition 4.5.

As a consequence of the Definition 4.4, a vertex q is marked with b relative to M in $\mathbb{G}$ in either of the following circumstances.

I. There exists a covering subforest $\bar{\mathbb{G}}$ of $\mathbb{G}$ such that no infinite sequence of bounded stopping times $\delta_1 < \delta_2 < \delta_3 < \ldots$ adapted to $\bar{\mathbb{G}}$ has the property that $f^3(\bar{q}, \delta_{2i-1}, \delta_{2i}) = \text{Red } \forall i > 0 \ \forall \bar{q} \in M$.

II. There exists a covering subforest $\mathbb{G}'$ of $\mathbb{G}$ such that, for any sequence of bounded stopping times $\delta_1 < \delta_2 < \delta_3 < \ldots$ adapted to $\mathbb{G}'$ which satisfies $f^3(\bar{q}, \delta_{2i-1}, \delta_{2i}) = \text{Red } \forall i > 0 \ \forall \bar{q} \in M$, there exists $N \in \mathbb{N}$ such that $f^3(q, \delta_{2i-1}, \delta_{2i}) = \text{Blue } \forall i \geq N$.

For any such $\mathbb{G}'$ in condition II, we say that q is marked with b relative to M with witness $\mathbb{G}'$ in $\mathbb{G}$. ($\mathbb{G}'$ is a witness to condition II.) Note that the finite set M can be empty. In such a case, condition II simplifies down to the existence of a covering subforest $\mathbb{G}'$ of $\mathbb{G}$ such that, for any sequence of bounded stopping times $\delta_1 < \delta_2 < \delta_3 < \ldots$ adapted to $\mathbb{G}'$, there exists $N \in \mathbb{N}$ such that $f^3(q, \delta_{2i-1}, \delta_{2i}) = \text{Blue } \forall i \geq N$. As in the case $k = 2$ analysed in Section 3.4, (for the proof later of Lemma 3.7,) it always follows from condition II above that a vertex q is labelled with b.

Lemma 3.6 is still true under the current assumption with exactly the same proof. That is, if a vertex q is marked with b relative to M with witness $\mathbb{G}'$

in $\mathbb{G}$, then there exists a covering subforest $\mathbb{G}(q)(M)$ of $\mathbb{G}$ pruned below q's level such that q is marked with b relative to M with witness $\mathbb{G}(q)(M)$ in $\mathbb{G}$.

To prove the existence of a well-structured subtree $\mathbb{A}$ in $\mathbb{F}$, we follow a similar procedure to that in the proof of Lemma 3.7, except for a modification reflecting the definition of f^3.

Lemma 4.6. *Given a stochastic hyper-colouring f^3, there exists in the tree of histories $\mathbb{F}$ a well-structured subtree $\mathbb{A}$ under f^3.*

Proof. We first extract a subtree $\mathbb{A}$ of $\mathbb{F}$ by the same labelling process described in Step 1 in the proof of Lemma 3.7. Now we prove this subtree $\mathbb{A}$ is a well-structured subtree of $\mathbb{F}$ under f^3.

We keep the index order of $V_{\mathbf{A}} = \{q_i\}_{i\in\mathbf{Z}^+}$ as in the labelling process described in the proof of Lemma 3.7, and denote two collections of vertices labelled with r and b by S_r and S_b, respectively. That is, $S_r = \bigcup_{i\in\mathbf{Z}^+} E_i$ and $S_b = V_{\mathbf{A}} \setminus S_r$. We show that the two partitioning sets S_r and S_b of $V_{\mathbf{A}}$ can be regarded as $R(\mathbb{A})$ and $B(\mathbb{A})$, respectively, in Definition 4.4 and hence the covering subtree $\mathbb{A}$ is a well-structured subtree under f^3.

For any covering subforest $\mathbb{G}$ of $\mathbb{A}$ and any finite set $S \subset V_{\mathbf{A}}$, we define

$$d(S) := \min\{n : \bigcup_{0\le i\le n} L_i(\mathbb{A}) \supseteq S\}.$$

If no ambiguity, we abbreviate $d(S)$ to d. By the definition of $z(k)$ in the labelling process, we see that in fact

$$\bigcup_{0\le i\le d} L_i(\mathbb{A}) = \{q_i : 0 \le i < z(d)\}.$$

Hence $S \subseteq \bigcup_{0\le i<z(d)} q_i$. Recall that

$$E_i = \{q_l : 0 \le l < i,\ q_l \text{ is labelled with r}\}.$$

It follows that $S_r \cap S \subseteq E_{z(d)}$. On the other hand, if $E_{z(d)} \neq \emptyset$, then for any covering subforest of $\mathbb{G}_{z(d)}$, and so in particular for $\mathbb{G}$ (since $V_{\mathbf{G}} \subseteq V_{\mathbf{A}} \subseteq V_{\mathbf{G}_{z(d)}}$), there exists an infinite increasing sequence $\{\delta_i\}$ of bounded stopping times adapted to $\mathbb{G}$ such that

$$f^3(q, \delta_{2i-1}, \delta_{2i}) = \text{Red}\ \forall q \in E_{z(d)}\ \forall i > 0,$$

by the definition of r. Therefore,

$$f^3(q, \delta_{2i-1}, \delta_{2i}) = \text{Red } \forall q \in S_r \cap S \ \forall i > 0.$$

If $E_{z(d)} = \emptyset$, pick any infinite increasing sequence of bounded stopping times adapted to $\mathbb{G}$ to be $\{\delta_i\}_{i>0}$.

Suppose that one vertex q_m is in $S_b \cap S$. By Lemma 3.6 and the definition of $\mathbb{G}_m(q_m)(E_m)$, for any infinite increasing sequence $\{\tau_i\}$ of bounded stopping times adapted to $\mathbb{G}_m(q_m)(E_m)$ with

$$f^3(q, \tau_{2i-1}, \tau_{2i}) = \text{Red } \forall i > 0 \ \forall q \in E_m,$$

there exists $N_m \in \mathbb{N}$ such that $f^3(q_m, \tau_{2i-1}, \tau_{2i}) = \text{Blue } \forall i \geq N_m$. Since each bounded stopping time in $\{\delta_i\}_{i\in\mathbb{N}}$ is adapted to $\mathbb{G}$ and $\mathbb{G}$ is a covering subforest of $\mathbb{G}_m(q_m)(E_m)$, each bounded stopping time in $\{\delta_i\}_{i\in\mathbb{N}}$ is also adapted to $\mathbb{G}_m(q_m)(E_m)$. Because $E_{z(d)} \supseteq E_m$, we see

$$f^3(q, \delta_{2i-1}, \delta_{2i}) = \text{Red } \forall i > 0 \ \forall q \in E_m.$$

Therefore, we can take the sequence $\{\delta_i\}$ for the sequence $\{\tau_i\}$ above. Furthermore, for every vertex q_m in $S_b \cap S$, we can likewise take $\{\delta_i\}$ again for the sequence $\{\tau_i\}$ above. Hence there exists a corresponding $N_m \in \mathbb{N}$ such that

$$f^3(q_m, \delta_{2i-1}, \delta_{2i}) = \text{Blue } \forall i \geq N_m.$$

Because $S_b \cap S$ is finite, we can define N to be the maximum of those N_m. Hence,

$$f^3(q, \delta_{2i-1}, \delta_{2i}) = \text{Blue } \forall q \in S_b \cap S \ \forall i \geq N.$$

Define a new infinite sequence of bounded stopping time $\{\delta'_i\}_{i\in\mathbb{N}}$ such that $\delta'_i = \delta_{i+2N} \ \forall i \in \mathbb{N}$. This is the desired sequence of bounded stopping times as per Definition 4.4, for the finite set S in the covering subforest $\mathbb{G}$. □

We copy the definition of r_{θ_i} and b_{θ_i} from Section 3.4. That is, in any covering subforest $\mathbb{G}$ of $\mathbb{A}$, for any bounded stopping time θ adapted to $\mathbb{G}$, we have

$$r_\theta(\mathbb{G}) = \cup\{B(q) : q \in S_\theta(\mathbb{G}) \cap R(\mathbb{A})\}$$

and

$$b_\theta(\mathbb{G}) = \cup\{B(q) : q \in S_\theta(\mathbb{G}) \cap B(\mathbb{A})\}.$$

Lemma 3.8 is still true in the current circumstance with exactly the same proof.

We now prove Theorem 4.3 by the same approach as in the proof of the Finite-state Two-colour Stochastic Partition Theorem (Theorem 3.1).

Proof of Theorem 4.3: In a covering subforest $\bar{\mathbb{F}}$ of $\mathbb{A}$ obtained from Lemma 3.8, we define θ_1 to be $\sigma_0(\bar{\mathbb{F}})$. We construct the desired infinite increasing sequence of bounded stopping times inductively. Suppose that we have obtained $\langle\theta_1, \theta_2, ..., \theta_{2k-1}\rangle$. Denote $\bigcup_{1\leq i\leq 2k-1} S_{\theta_i}(\bar{\mathbb{F}})$ by S^k. Since $\bar{\mathbb{F}}$ is a covering subforest of $\mathbb{A}$ and S^k is finite, by Definition 4.4, there exists an infinite increasing sequence of bounded stopping times $\{\delta_i^k\}_{i\in\mathbb{N}}$ adapted to $\bar{\mathbb{F}}$ with the property

$$f^3(q, \delta_{2i-1}^k, \delta_{2i}^k) = \text{Red } \forall q \in S^k \cap R(\mathbb{A}) \ \forall i > 0$$

and

$$f^3(q, \delta_{2i-1}^k, \delta_{2i}^k) = \text{Blue } \forall q \in S^k \cap B(\mathbb{A}) \ \forall i > 0.$$

Let θ_{2k} be δ_1^k and θ_{2k+1} be δ_2^k. In this way, we find an infinite increasing sequence of bounded stopping times $\{\theta_n\}_{n\in\mathbb{N}}$. It follows from Lemma 3.8 that, for all of its initial subsequences $\{\theta_1, ..., \theta_{2i}, \theta_{2i+1}\}$ $(i \geq 1)$,

$$P(\hat{f}^3_{\theta_m,\theta_{2n},\theta_{2n+1}} = \text{Red } \forall 1 \leq m < 2n \leq 2i | R) > 1 - 5\epsilon/8, \quad \text{if } P(R) > 0,$$

and

$$P(\hat{f}^3_{\theta_m,\theta_{2n},\theta_{2n+1}} = \text{Blue } \forall 1 \leq m < 2n \leq 2i | B) > 1 - 5\epsilon/8, \quad \text{if } P(B) > 0,$$

which completes the proof. □

Comment 1: Analogously to Theorem 3.13 and Theorem 3.2, we can prove the finite versions of Theorem 4.2 which follows, by compactness arguments. In the first of those, the probability measure is fixed.

Theorem 4.7. *For a probability space $(\Omega, \mathcal{F}_\infty, P)$ constructed from a sequence of sets of countably many states, a set C of finitely many colours, an* *odd integer* $2m+1$ *with* $m \in \mathbb{N}$ *and any* $\epsilon > 0$*, there exists* $n = n(m, \epsilon, P) \in \mathbb{N}$ *such that for every stochastic hyper-colouring* f^3 *with values in* C*, there exists* $2m+1$ *bounded stopping times* $\theta_1 < \theta_2 < ... < \theta_{2m+1} \leq n$ *which satisfy*

$$P(\hat{f}^3_{\theta_1,\theta_2,\theta_3} = \hat{f}^3_{\theta_i,\theta_{2j},\theta_{2j+1}} \ \forall 1 \leq i < 2j; \ 1 \leq j \leq m) > 1 - \epsilon.$$

For the case when only finitely many states are allowed, we have a uniform result over the space of probability measures.

Theorem 4.8. *For a measurable space* $(\Omega, \mathcal{F}_\infty)$ *constructed from a sequence of sets each containing finitely many states, a set* C *of finitely many colours, an integer* $2m+1$ *with* $m \in \mathbb{N}$ *and any* $\epsilon > 0$*, there exists* $n = n(m, \epsilon) \in \mathbb{N}$ *such that, for every probability measure* P *defined in* $(\Omega, \mathcal{F}_\infty)$ *and every stochastic hyper-colouring* f^3 *with values in* C*, there exist* $2m+1$ *bounded stopping times* $\theta_1 < \theta_2 < ... < \theta_{2m+1} \leq n$ *with*

$$P(\hat{f}^3_{\theta_1,\theta_2,\theta_3} = \hat{f}^3_{\theta_i,\theta_{2j},\theta_{2j+1}} \ \forall 1 \leq i < 2j; \ 1 \leq j \leq m) > 1 - \epsilon.$$

Comment 2: One might hope to use in the proof a stronger marking scheme as follows, but this does not seem to be helpful.

Given a covering subforest $\mathbb{G}$ of $\mathbb{F}$ and a finite set M of vertices with $M \subseteq V_{\mathbb{G}}$, for every vertex q in $V_{\mathbb{G}} \setminus M$, mark q with symbol r relative to M in $\mathbb{G}$, if for any covering subforest $\mathbb{G}'$ of $\mathbb{G}$, there exists an infinite sequence of bounded stopping times $\delta_1 < \delta_2 < \delta_3 < ...$ adapted to $\mathbb{G}'$ such that

$$f^3(q, \delta_m, \delta_n) = f^3(\bar{q}, \delta_m, \delta_n) = \text{Red} \ \forall m < n \ \forall \bar{q} \in M.$$

Mark q with symbol b relative to M in $\mathbb{G}$ if q cannot be marked with r relative to M in $\mathbb{G}$.

Indeed, this stronger marking scheme looks more appropriate for application to the general claim (4.4) in regard to f^3, that is,

$$P(\hat{f}^3_{\theta_1,\theta_2,\theta_3} = \hat{f}^3_{\theta_{n_1},\theta_{n_2},\theta_{n_3}} \ \forall (n_1, n_2, n_3) \in \mathbb{N}^{(3)}) > 1 - \epsilon.$$

However, given a finite set $M \subset V_{\mathrm{G}}$ with $f^3(\bar{q}, \delta_m, \delta_n) = \text{Red } \forall m < n \; \forall \bar{q} \in M$, it is not easy to give a direct condition for marking a vertex with b. Furthermore, a probabilistic difficulty in this general stochastic claim above prevents me from progressing to a successful proof, and I have to adopt a simpler marking scheme for a weaker result (i.e., Theorem 4.3).

4.2 Further comments

Before explaining the difficulty in proving the general stochastic extension of Ramsey's theorem on f^k with $k > 2$, it is worth summarising first why the proof in Section 3.4 is valid for f^k with $k = 2$, but why the direct induction method fails. (The induction idea is indeed used at the end of the proof of Theorem 3.1 in Section 3.4, but the key of the proof, i.e., Lemma 3.7, is proved by other methods than an induction.)

Recall that in the proof of the Stochastic Infinitary Ramsey Theorem (Theorem 1.2), one key point is in dealing with the domain Z of the stochastic colouring f: it is not simply the collection of two partially ordered partial histories or bounded stopping times, but a collection of pairs each consisting of a partial history and a bounded stopping time consistent with this partial history. The induced random colouring $\hat{f}$ is from $\mathcal{T}^{(2)} \times \Omega$ to C. Given (σ, τ) in $\mathcal{T}^{(2)}$, it is essentially a stochastic colouring determined by the random selection of partial histories truncated at the first bounded stopping time σ. That is, $\hat{f}_{\sigma,\tau}$ is an $\mathcal{F}_\sigma$-measurable random variable. The asymmetry inside the domain Z of the stochastic colouring f inhibits any easy attempt on applying directly the induction method for 'Ramsey' problems.

We analyse here the impact of the 'asymmetry' on the proof of the Finite-state Finite-colour Stochastic Partition Theorem (Theorem 3.9). Suppose that we adopt the natural induction on bounded stopping times. That is, in a fixed set Q_i with

$$P(\hat{f}_{\theta_m,\theta_n} = c_i \; \forall 1 \leq m < n \leq t | Q_i) > 1 - \epsilon$$

where $C = \{c_1, ..., c_k\}$ and $1 \leq i \leq k$, we are considering the yet to be constructed bounded stopping time θ_{t+1}. Assume $\epsilon < 1/k$, $\theta_1 > 0$ and that P has a uniform distribution. We then cannot rule out the possible situation that in any such Q_i no bounded stopping time $\theta_{t+1} > \theta_t$ satisfies

$$P(\hat{f}_{\theta_m,\theta_n} = c_i \ \forall 1 \leq m < n \leq t+1 | Q_i) > 1 - \epsilon.$$

We can construct a counterexample by manipulating $\hat{f}_{\theta_1,\delta}$ for every bounded stopping time greater than θ_t, denoted here by δ. We firstly enumerate the elements in $S_{\theta_1}(\mathbb{F})$ as $q_1, q_2, \cdots$. Regarding the k colours as presented cyclically, for any bounded stopping time $\delta > \theta_t$, we can define $f(q_l, \delta) = c_{\text{ mod } (l/k)}$ so that $f(q_l, \delta) \neq f(q_{l+1}, \delta)$ for all successive pairs of (q_l, q_{l+1}) included in $S_{\theta_1}(\mathbb{F})$. So we cannot adopt the standard induction method on bounded stopping times in such 'Ramsey' problems.

Our approach in Section 3.4 is to focus on individual partial histories q in Ω_∞ in turn, and check whether there exists an infinite increasing sequence of bounded stopping times $\{\delta_l\}_{l \in \mathbb{N}}$ consistent with q and with a certain labelling property. The result can only be 'yes' or 'no'. Then in Lemma 3.8, we consider the probabilistic aspect of the problem, and take ϵ into account to find a nice subfiltration.

However, this method looks inadequate for the case of $k > 2$ on f^k. For example, assume $k = 3$ and Red $\in C$. We can of course check for any particular partial history $q \in \Omega_\infty$ whether there exists an increasing sequence of bounded stopping times $\{\delta_i\}_{i \in \mathbb{N}}$ such that

$$f^3(q, \delta_m, \delta_n) = \text{Red} \ \forall 1 \leq m < n.$$

Nevertheless, we have to note that, in a desired infinite increasing sequence $\{\theta_i\}_{i \in \mathbb{N}}$ as in (4.4) for $k = 3$, i.e.,

$$P(\hat{f}^3_{\theta_1,\theta_2,\theta_3} = \hat{f}^3_{\theta_{n_1},\theta_{n_2},\theta_{n_3}} \ \forall (n_1, n_2, n_3) \in \mathbb{N}^{(3)}) > 1 - \epsilon,$$

every θ_i with $i \geq 2$ can appear in two of the argument places in $\hat{f}^3$ and in f^3. Given $i \in \mathbb{N}$, we need to consider the random colour $\hat{f}^3_{\theta_m,\theta_i,\theta_n}$ for any $m < i < n$ and the random colour $\hat{f}^3_{\theta_m,\theta_n,\theta_i}$ for any $m < n < i$. That is, θ_i $(i \geq 2)$ can appear as the second or the third bounded stopping time in $\hat{f}^3$ and the corresponding formulation of f^3. When reconciling these 'double identities', we have to deal with the problem of compatibility of 'candidate' sequences $\{\delta_l\}_{l\in\mathbb{N}}$, given a set of partial histories at which we have stopped. To be specific, suppose that given $\{\theta_1, ..., \theta_i\}$, for a partial history q in S_{θ_i}, there exists an increasing sequence of bounded stopping times $\{\delta_l\}_{l\in\mathbb{N}}$ such that $f^3(q, \delta_m, \delta_n) = \text{Red}\ \forall 1 \leq m < n$. We can then pick one δ_l as θ_{i+1}. The important point is that θ_{i+2} should also be chosen from the same sequence $\{\delta_l\}_{l\in\mathbb{N}}$ to make $f^3(q, \theta_{i+1}, \theta_{i+2})$ the desired colour, Red. However, for any vertex $q' \in S_{\theta_{i+1}}$, any increasing sequence of bounded stopping times $\{\delta'_l\}_{l\in\mathbb{N}}$ with $f^3(q', \delta'_m, \delta'_n) = \text{Red}\ \forall 1 \leq m < n$ might have no common subsequence with $\{\delta_l\}_{l\in\mathbb{N}}$ above, i.e., $\{\delta_l\}_{l\in\mathbb{N}} \cap \{\delta'_l\}_{l\in\mathbb{N}} = \emptyset$. Furthermore, each $q' \in S_{\theta_{i+1}}$ might have different $\{\delta'_l\}_{l\in\mathbb{N}}$. The distinction between sequences $\{\delta_l\}_{l\in\mathbb{N}}$ associated with those $q \in S_{\theta_i}$ and $q' \in S_{\theta_{i+1}}$ invalidates our approach in Section 3.4.

Thus, any method to fix θ_{i+2} must take into account not only partial histories truncated at θ_i, but also the bounded stopping time θ_{i+1}. However, we have already shown that the standard induction method involving bounded stopping times is inappropriate for a proof of Ramsey's theorem in the stochastic context.

We may, of course, try other approaches. For example, under a partial history q, we may check whether there exists a covering subforest $\mathbb{G}$ such that any infinite increasing sequence of bounded stopping times $\{\delta_l\}_{l\in\mathbb{N}}$ adapted to $\mathbb{G}$ satisfies some conditions relative to q. That idea needs a closer look at the relationship among sequences of covering subforests, which is not straightforward either.

To sum up, for the analogous stochastic extension (4.4) to stochastic hypercolourings f^k with $k > 2$, we cannot prove it by either of the two natural methods: induction on bounded stopping times, or considering individual partial histories then fixing a common solution sequence of bounded stopping times. The difficulty in the former is the asymmetry inside the domain of f^k, and in the latter is the possible incompatibility of the sequences from different individual partial histories. I do not know whether the general stochastic extension (4.4) of Ramsey's theorem is true. If not, any counterexample would be sophisticated, since we have proved it true for the case of $k = 2$.

5

Applications in utility theory

We briefly recall some background in utility theory relevant for our application of the stochastic Ramsey theorem. (See [11] for further information.) In particular, we discuss the significance of monotonic utility functions defined on bounded stopping times.

We first construct a probability space by a coordinate process. For a set Ω, we define an infinite sequence of IID (independent and identically distributed) random variables $\{X_n : \Omega \to \{1, -1\},\ \forall n \in \mathbb{N}\}$. Regard the sequence $\{X_n\}_{n\in\mathbb{N}}$ as the coordinate process, and assign the probability measure P so that

$$P(X_n = 1) = P(X_n = -1) = 1/2 \quad \forall n \in \mathbb{N}.$$

We view this model in the sense that, at every stage in an infinitely long game, a fair coin is tossed and one player wins 1 pound if Head shows and loses 1 pound if Tail shows. Then X_n is the player's net winning at stage n in this betting game. We define Y_n to be the random variable representing the *accumulated fortune* at stage n, i.e., $Y_n = \sum_{1\leq k\leq n} X_k$. For completeness, we let $Y_0(\omega) = 0\ \forall\omega \in \Omega$. So for a bounded stopping time θ, Y_θ is the random cumulated fortune at θ. For a partial history $q = (q(1), ..., q(n))$, we denote $Y_n(\omega)$ by $Y(q)$ if $\omega|n = q$. By Doob's Optional-Stopping Theorem (cf. 10.10 in [17]),

$$E[Y_\theta|B(q)] = Y_q \ \forall (q,\theta) \in Z,$$

where as before $Z = \{(q,\tau) : q \in \Omega_{<\infty};\ \tau \in \mathcal{T}(q)\}$. In particular

$$E[Y_\theta] = 0 \quad \forall \theta \in \mathcal{T}. \tag{5.1}$$

From the perspective of utility theory, one can evaluate the expected utility of Y_θ. That is, the player is described by a utility function $u : \mathbb{Z} \to \mathbb{R}$ and his evaluation of Y_θ is

$$u[\theta] := \sum_{q \in S_\theta(\mathbf{F})} u\big(Y(q)\big)P\big(B(q)\big).$$

We assume that u is globally strictly increasing.

To show the difference between $E[Y_\theta]$ and $u[\theta]$, we compare two bounded stopping times $\theta_1 = 1$ and $\theta_2 = 100$. By (5.1), $E[Y_{\theta_1}] = E[Y_{\theta_2}] = 0$. However, $u[\theta]$ depends on the utility function u. If u is a strictly concave function, then $u[\theta_1] > u[\theta_2]$. The strictly concave utility function describes risk aversion. That is, a player with such u prefers a fixed monetary outcome to a lottery with expectation equal to that outcome. In this case $u(-n) + u(n) < u(-(n-1)) + u(n-1)\ \forall n \geq 1$, due to the concavity of u. If u is a strictly convex function, then $u[\theta_1] < u[\theta_2]$, which describes risk seeking. If u is a linear function, then $u[\theta_1] = u[\theta_2]$ and we say that player is risk neutral. Note that u can be any increasing function, with no constraint on global convexity or concavity.

One may view the expected utility as a 'generalisation' of expectation, as $u[\theta] = E[Y_\theta]$, if $u(x) = x$ for all x in $\mathbb{Z}$. For a pair (q,θ) in Z, the player can evaluate the expected utility of Y_θ conditional on q as

$$u[q,\theta] := \sum_{h \in S_\theta(\mathbf{F})} u\big(Y(h)\big)P\big(B(h)|B(q)\big).$$

Indeed, given $\theta = 100$, for $q_1 = \emptyset$ and $q_2 = (q(1), ..., q(99))$ where $q(k) = 1\ \forall 1 \leq k \leq 99$, the expected utility $u[q_1,\theta]$ and $u[q_2,\theta]$ is different. We can even generalise further to define a *history dependent utility function*: the player has a countably infinite sequence of utility functions $\{u_q\}_{q \in \Omega_{<\infty}}$ and for each pair (q, θ) in Z, his evaluation of Y_θ conditional on q is

$$\bar{u}[q, \theta] := \sum_{h \in S_\theta(\mathbf{F})} u_q(Y(h))P(B(h)|B(q)).$$

We of course assume each $u_q : \mathbb{Z} \to \mathbb{R}$ is an increasing function.

We apply the stochastic Ramsey theorem in the current model. Under the standard assumption that the utility is bounded (cf. e.g. [1], [2], [4] and [15]), we suppose that $u_q : \mathbb{Z} \to [0, N)$ where $N \in \mathbb{N}$ for each q.

Theorem 5.1. *For a history dependent utility function $\bar{u}$ generated from a sequence $\{u_q\}_{q \in \Omega_{<\infty}}$ and any $\epsilon > 0$, there exists an increasing sequence of bounded stopping times $\theta_1 < \theta_2 < \theta_3 < \ldots$ such that*

$$P(\omega \in \Omega : |\bar{u}[\omega|\theta_m, \theta_n] - \bar{u}[\omega|\theta_i, \theta_j]| < \epsilon \; \forall 1 \leq m < n \; \forall 1 \leq i < j) > 1 - \epsilon.$$

Proof. We give a partition of $[0, N)$ by $\{[i\epsilon, \min\{(i+1)\epsilon, N\}) : 0 \leq i < N/\epsilon;\ i \in \mathbb{Z}^+\}$. For every $(q, \theta) \in Z = \{(q, \tau) : q \in \Omega_{<\infty}, \tau \in \mathcal{T}(q)\}$, we define $f(q, \theta) := c_{i+1}$ if $\bar{u}[q, \theta] = u_q[q, \theta] \in [i\epsilon, \min\{(i+1)\epsilon, N\})$. Hence the function $f : Z \to \{c_1, \ldots, c_{\lceil N/\epsilon \rceil}\}$ is a stochastic colouring. Apply Theorem 1.2. □

Comment 1: If we define $u_q(Y(h)) = Y(h)$ for all pairs of partial histories q and h where h is an extension of q (i.e., $\exists \omega \in \Omega, n_1, n_2 \in \mathbb{N}$ with $n_1 < n_2$ such that $\omega|n_1 = q$ and $\omega|n_2 = h$), then we obtain the following conclusion.

For $\epsilon > 0$, $Y = (Y_n : n \geq 0)$ a bounded process adapted to the filtration $\{\mathcal{F}_n\}$ in $(\Omega, \mathcal{F}_\infty, P)$, there exists an infinite increasing sequence of bounded stopping times $\{\theta_n\}$ such that

$$P(|E[Y_{\theta_j}|\mathcal{F}_{\theta_i}] - Y_{\theta_i}| < \epsilon \; \forall 1 \leq i < j) > 1 - \epsilon.$$

Comment 2: The setup of Theorem 5.1 is in essence an example of a stochastic colouring with $\mathcal{F}$-consistency condition introduced in Section 1.1. It is unknown whether an application of stochastic colouring without $\mathcal{F}$-consistency but with practical meaning exists.

Bibliography

[1] K. Arrow, The Use of Unbounded Utility Functions in Expected-Utility Maximization: Response, Quarterly Journal of Economics. 88 (1974) 136-138.

[2] R. Aumann, The St. Petersburg Paradox: A Discussion of Some Recent Comments, Journal of Economic Theory. 14 (1977) 443-445.

[3] P. Billingsley, Probability and Measure, third ed., Wiley, 1995.

[4] D. Brito, Becker's Theory of the Allocation of Time and the St. Petersburg Paradox, Journal of Economic Theory. 10 (1975) 123-126.

[5] R. Engelking, General Topology (Sigma Series in Pure Mathematics), Heldermann Verlag, 1989.

[6] H. Furstenberg and B. Weiss, Markov Processes and Ramsey Theory for Trees, Combinatorics, Probability and Computing. 12 (2003) 547-563.

[7] R. L. Graham, B. L. Rothschild and J. H. Spencer, Ramsey Theory, second ed., Wiley, 1990.

[8] T. J. Jech, Set Theory, Springer, 2003.

[9] A. S. Kechris, Classical Descriptive Set Theory, Springer, 1995.

[10] R. Lyons with Y. Peres, Probability on Trees and Networks, Cambridge University Press, to appear.

[11] J. Von Neumann and O. Morgenstern, Theory of Games and Economic Behavior (Commemorative Edition), Princeton University Press, 2007.

[12] K. R. Parthasarathy, Probability Measures on Metric Spaces, Academic Press, 1967, p. 27 and p. 45.

[13] F. Ramsey, On a Problem of Formal Logic, Proceedings of the London Mathematical Society. 30 (1930) 264-286.

[14] W. Rudin, Real and Complex Analysis, third ed., McGraw-Hill, 1987.

[15] P. Samuelson, St. Petersburg Paradoxes: Defanged, Dissected and Historically Described, Journal of Economic Literature. 15 (1977) 24-55.

[16] E. Shmaya and E. Solan, Two Player Non Zerosum Stopping Games in Discrete Time, The Annals of Probability. 32 (2004) 2733-2764.

[17] D. Williams, Probability with Martingales, Cambridge University Press, 1991.