FLUCTUATIONS OF BALANCED URNS WITH INFINITELY MANY COLOURS

SVANTE JANSON, CÉCILE MAILLER, AND DENIS VILLEMONAIS

Abstract. In this paper, we prove convergence and fluctuation results for measure-valued Pólya processes (MVPPs, also known as Pólya urns with infinitely-many colours). Our convergence results hold almost surely and in L^2, under assumptions that are different from that of other convergence results in the literature. Our fluctuation results are the first second-order results in the literature on MVPPs; they generalise classical fluctuation results from the literature on finitely-many-colour Pólya urns. As in the finitely-many-colour case, the order and shape of the fluctuations depend on whether the "spectral gap is small or large".

To prove these results, we show that MVPPs are stochastic approximations taking values in the set of measures on a measurable space E (the colour space). We then use martingale methods and standard operator theory to prove convergence and fluctuation results for these stochastic approximations.

1. Introduction

1.1. A brief overview of the theory of Pólya urns.

A d-colour Pólya urn is a stochastic process that describes the evolution of an urn containing balls of d different colours. It is a Markov process that depends on two parameters: the initial composition of the urn $\mathfrak{u}_0 \in \mathbb{N}^d$ and a replacement matrix $\mathfrak{r} = (\mathfrak{r}_{x,y})_{1\leq x,y\leq d}$, which has integer entries. At time zero, the urn contains $\mathfrak{u}_{n,x}$ balls of colour x, for all $1 \leq x \leq d$. At every discrete time-step, we pick a ball uniformly at random in the urn, and if it is of colour x, we replace it in the urn together with an additional $\mathfrak{r}_{x,y}$ balls of colour y, for all $1 \leq y \leq d$. The quantity of interest is the process $(\mathfrak{u}_n)_{n\geq 0}$, where, for all $n \geq 0$, the vector $\mathfrak{u}_n = (\mathfrak{u}_{n,1}, \dots, \mathfrak{u}_{n,d})$ is the composition of the urn at time n.

As expected, the behaviour of the composition vector at large times depends on the replacement matrix. The case when the replacement matrix is the identity was studied by Markov [28] and then Pólya and Eggenberger [9]. It is well-known that, in this case, $\mathfrak{u}_n/n$ converges almost surely to a d-dimensional Dirichlet random variable of parameter $\mathfrak{u}_0$. The fluctuations around this limit are Gaussian, conditioned on the limit. (See [30, Section 2.3.1].)

Pólya urns whose replacement matrix is irreducible (the irreducibility assumption can be weakened, see Janson [18]) exhibit a drastically different behaviour, see e.g. Athreya and Karlin [1]: in that case, if for simplicity all replacements $\mathfrak{r}_{x,y}$ are non-negative (this too can be relaxed), the Perron–Frobenius theorem implies that the spectral radius $\mathfrak{s}$ of $\mathfrak{r}$ is also a simple eigenvalue of $\mathfrak{r}$, and that there exists a unit left-eigenvector v associated to $\mathfrak{s}$ whose coordinates are all non-negative. Then, as n goes to infinity, $\mathfrak{u}_n/n$ converges almost surely to $\mathfrak{s}v$. Interestingly, the fluctuations around this limit are either Gaussian and of order $\sqrt{n}$, or non-Gaussian and of higher order, depending on the spectral gap of $\mathfrak{r}$ (see, e.g. Janson [18] or Pouyanne [31]).

The main differences between the identity and the irreducible cases are that (1) the limit of $\mathfrak{u}_n/n$ is random in the identity case, and deterministic in the irreducible case, (2) it

Date: 25 November, 2021.
2020 *Mathematics Subject Classification.* 60F05, 60F25, 60J80, 62L20.
SJ is supported by the Knut and Alice Wallenberg Foundation.
CM is grateful to EPSRC for support through the fellowship EP/R022186/1.

depends on the initial composition in the identity case, while it does not in the irreducible case, and (3) the irreducible case sometimes exhibits non-Gaussian fluctuations.

Since these seminal results, the model of Pólya urns has been extended and more precise asymptotic results have been proved. The most natural extension is to allow balls to be removed from the urn: It is standard to allow the diagonal coefficients of the replacement matrix to equal -1, meaning that the ball that was drawn is removed from the urn. One can also allow other coefficients of the replacement matrix to be negative and work conditionally on "tenability", which is the event that all coefficients of the composition vector stay non-negative at all times. The model can also be extended by allowing the replacement matrix to be random (at each time step, we use a new realisation of this matrix), different colours to have different weights or activities (a ball is drawn with probability proportional to its weight). These three generalisations are for example considered in [18] (see Remark 4.2 therein for ball substractions).

1.2. **Measure-valued Pólya processes.** Measure-valued Pólya processes were introduced by Bandyopadhyay and Thacker [3], and shortly after by Mailler and Marckert [26], as a generalisation of Pólya urns to infinitely many colours. They both considered cases that can be seen as corresponding to the irreducible case in Section 1.1. In fact, the generalisation to infinitely many colours in the diagonal case is much older and dates back to Blackwell and MacQueen [4].

In the analogue of the irreducible case, the theory is very recent and, as far as we know, there are only five papers on the subject: Bandyopadhyay and Thacker [3], Mailler and Marckert [26], Janson [21], Mailler and Villemonais [27], and Bandyopadhyay, Janson and Thacker [2]. The main difficulty is that the linear algebra tools used in the study of Pólya urns are replaced by operator theory in an infinite dimensional space.

In the model introduced by [3] and [26], a measure-valued Pólya process (MVPP) is defined as a Markov process $(\mathfrak{m}_n)_{n\geq 0}$ taking values in the set of positive measures on a measurable space E of colours. The process depends on two parameters again: the initial composition measure $\mathfrak{m}_0$ and the replacement kernel $(R_x)_{x\in E}$ (a family of positive measures on E; see Appendix A for measurability issues).

At every discrete time-step $n \geq 1$, a random colour Y_n is drawn at random in E with probability distribution $\mathfrak{m}_{n-1}/\mathfrak{m}_{n-1}(E)$, and then $\mathfrak{m}_n$ is defined as $\mathfrak{m}_{n-1} + R_{Y_n}$ (see Section 2 for details).

The authors of [3] and [26] see the MVPP as a branching version of the E-valued Markov chain $(\mathfrak{w}_n)_{n\geq 0}$ with transition kernel $(R_x)_{x\in E}$. They assume that the MVPP is "balanced", i.e., that the R_x's are all probability measures, which makes the Markov chain well defined. They use this representation to prove that, if $(\mathfrak{w}_n)_{n\geq 0}$ is "ergodic" (in a general sense that allows renormalisations), then a renormalised version of $\mathfrak{m}_n/\mathfrak{m}_n(E)$ converges in probability to the limiting distribution of $(\mathfrak{w}_n)_{n\geq 0}$. The "ergodicity" assumption in this MVPP case can be seen as the equivalent of the "irreducibility" assumption in the finitely-many-colour case. This result is improved by Janson [21], who allows the replacement kernel to be random.

Bandyopadhyay, Janson and Thacker [2] later built on these methods to prove that the convergence results of [3] and [26] hold almost surely, under a condition that they call "uniform ergodicity" on the underlying Markov chain $(\mathfrak{w}_n)$, and if the set of colours is countable.

Using a different approach, Mailler and Villemonais [27] were able to consider non-balanced, weighted MVPPs, also with random replacements; these are three generalisations that are classical in the finitely-many-colour case and that extend the range of applications. In the non-balanced case, R_x may be a defective measure, so the underlying Markov chain $(\mathfrak{w}_n)_{n\geq 0}$ has an absorbing "cemetery" state. The authors show that, if the continuous-time version of the underlying Markov chain admits a quasi-stationary distribution (and under other important assumptions), then $\mathfrak{m}_n/n$ converges almost surely to this quasi-stationary distribution. They use stochastic approximation methods, which is difficult since the stochastic approximation takes values in a non-compact space as soon as the space of colours is non-compact (which is desirable for many applications), but which gives almost sure convergence instead of the convergence in probability of [3] and [26]. The difficulty coming from the fact that the stochastic approximation takes values in a non-compact space is overcome by a Lyapunov-type assumption. The main drawback of this method is that the Markov chain needs to be "quasi-ergodic" without any renormalisations, whereas renormalisations were allowed in [3] and [26].

1.3. **Our contribution.** In this paper, we prove limit theorems for the fluctuations of an MVPP around its almost sure limit: we are able to generalise the fluctuations results of [18] to the infinitely-many-colour case. Our framework is close to that of [27], although we restrict ourselves to the balanced case; we expect the non-balanced case to be more challenging and leave it open for now.

Interestingly, our results do not use the results of [27]: they are totally self-contained, and our methods also give almost sure convergence of $\mathfrak{m}_n/\mathfrak{m}_n(E)$ to its limit, under a set of assumptions that are different from those of [27]. Similarly to [27], we use a Lyapunov-type assumption to deal with the fact that, in general, $\mathfrak{m}_n/\mathfrak{m}_n(E)$ takes values in a non-compact space.

To prove these results, we use stochastic approximation and thus martingale methods, together with standard operator theory (in particular, we refer several times to the book of Conway [5] on the subject).

1.4. **Some notation and conventions.** "Positive" is used in the weak sense, i.e., non-negative.

The notation 1 stands for the usual number, and also for the function that is constant equal to 1 on E. Indicator functions are denoted by $\mathbf{1}$.

$\mathbf{I}$ stands for the identity operator. As usual, for any complex number $z \in \mathbb{C}$ and for any operator T, the operator $T + z$ stands for $T + z\mathbf{I}$.

If T is a bounded operator in a Banach space $\mathcal{X}$, and Δ is a clopen (closed and open) subset of its spectrum $\sigma(T)$, let $\Pi_\Delta = \Pi_\Delta(T)$ denote the corresponding spectral projection in $\mathcal{X}$. (See e.g. [7, VII.3.17–20] or [5, Exercise VII.4.9 and VII.(6.9)].) In particular, if λ is an isolated point in $\sigma(T)$, $\Pi_\lambda := \Pi_{\{\lambda\}}$ is a projection onto the corresponding generalized eigenspace. Note that T commutes with Π_Δ, and thus T maps the range $\Pi_\Delta(\mathcal{X})$ into itself (i.e., $\Pi_\Delta(\mathcal{X})$ is an invariant subspace); moreover the spectrum of the restriction of T to $\Pi_\Delta\mathcal{X}$ equals Δ [5, after Equation VII.6.9].

For any non-negative integer $n \geq 1$, $\mathbb{E}_n$ is the conditional expectation with respect to $\mathcal{F}_n$, the σ-field generated by all events up to time n, i.e., by Y_i and $R^{(i)}_{Y_i}$ for $1 \leq i \leq n$.

Let $\mathcal{M}(E)$ be the space of complex measures on E (recall that these are finite by definition), and let $\mathcal{M}_\mathbb{R}(E), \mathcal{M}_+(E), \mathcal{M}_{>0}(E), \mathcal{P}(E)$ denote the subsets of finite signed (i.e., real-valued) measures, finite positive measures, finite positive non-zero measures, and probability measures, respectively. These sets can all be regarded as measurable spaces, with the σ-fields generated by the mappings $\mu \mapsto \mu(A)$, $A \in \mathcal{E}$.

If μ is a (possibly signed or complex) measure on E and f is a measurable function, then $\mu f := \int f \, \mathrm{d}\mu$ (whenever this is defined).

For a complex measure μ on E, let $|\mu|$ denotes its total variation measure, and $\|\mu\| = |\mu|(E)$ its total variation. If w is a positive function on E, then $\mathcal{M}(w)$ is the Banach space of complex measures μ on E, such that the norm $\|\mu\|_w := |\mu|w$ is finite. $\mathcal{P}(w) := \mathcal{P}(E) \cap \mathcal{M}(w)$ is the subset of probability measures in $\mathcal{M}(w)$.

For any positive function w on E, we define the complex Banach space

$$B(w) := \Big\{ g : E \to \mathbb{C} \;\Big|\; g \text{ is measurable and } \|g\|_{B(w)} := \sup_{x \in E} \frac{|g(x)|}{w(x)} < +\infty \Big\}. \tag{1.1}$$

In the special case $w = 1$ we write $B(E)$, the space of bounded measurable functions on E. Note that $\mathcal{M}(w)$ can be regarded (isometrically) as a subspace of the dual space $B(w)^*$ in the obvious way.

If $\beta = (\beta_x)_x$ is a kernel from E to a measurable space F (see Appendix A for definition), then $\mu\beta$ denotes the measure on F given by

$$\mu\beta(A) := \int_E \beta_x(A)\,\mathrm{d}\mu(x). \tag{1.2}$$

(This is the projection onto F of the measure $\mu \otimes \beta$ defined in (A.1).) We extend (1.2) to complex measures μ and signed kernels β such that $\int_E \|\beta_s\|\,\mathrm{d}|\mu|(s) < \infty$.

If T is a bounded operator on $B(w)$ such that its adjoint maps $\mathcal{M}(w)$ into itself, then we write the adjoint as T acting on the right on measures; we then have the associativity

$$(\mu T)f = \mu(Tf) \tag{1.3}$$

for (suitable) measures μ and functions f on E.

For a Banach space D, we use $\|\cdot\|_D$ both for the norm of elements of D, and for the operator norm of operators $D \to D$.

We also make use of the following usual notations and conventions: $x \vee y := \max\{x, y\}$; $x \wedge y := \min\{x, y\}$; $xy \wedge z = (xy) \wedge z$; empty sums are 0 and empty products are 1; $\inf \emptyset := +\infty$ and $\sup \emptyset := -\infty$.

We let C and $\mathfrak{C}$ denote unspecified constants whose meaning may change from one occurrence to the next. We use $\mathfrak{C}$ for constants that may depend on $\mathfrak{m}_0$ while C denotes constants that do not depend on $\mathfrak{m}_0$. Subscripts may be used to identify specific constants.

1.5. Plan of the paper.

In Section 2, we define our model, state and discuss our assumptions and our main results. Our main results are two main theorems: Theorem 2.10 states convergence of the MVPP, Theorem 2.13 gives the fluctuations of the MVPP around its limit. In Section 3, we prove Theorem 2.10, and in Section 4, we prove Theorem 2.13. In Section 5, we prove Theorems 2.25-2.27, which give conditions for the limits in Theorem 2.13 to be non-degenerate. In Section 6, we apply our main result to four examples: the out-degree profile of the random recursive tree, the heat kernel on the square, a branching random walk, and reinforced processes on a countable state-space.

Finally, we have three appendices. In Appendix A, we discuss the construction of the MVPP and measurability issues. In Appendix B, we state some general results on the spectra of operators on Banach spaces, which are useful for our proofs. Appendix C, we prove a technical lemma that is used in the proof of Theorem 2.13.

2. Model and main results

Let $(E, \mathcal{E})$ be a measurable space, $R^{(1)} = (R^{(1)}_x)_{x\in E}$ be a set of finite (possibly signed) random measures on E indexed by $x \in E$, and let $\mathfrak{m}_0$ be a (non-random) finite measure on E. (E may be called the colour space.) We define the measure-valued Pólya process (MVPP) $(\mathfrak{m}_n)_{n\geq 0}$ of initial composition $\mathfrak{m}_0$ and random replacement kernel $R^{(1)}$ as the Markov process given by the following recursion. See Appendix A for some technical details, including measurability assumptions.

Given $\mathfrak{m}_n$ with $n \geq 0$, first sample $Y_{n+1} \in E$ such that Y_{n+1} is a random variable whose conditional distribution on E, given $\mathfrak{m}_n$ and the previous history, is

$$\widetilde{\mathfrak{m}}_n := \mathfrak{m}_n/\mathfrak{m}_n(E). \tag{2.1}$$

Then, let

$$\mathfrak{m}_{n+1} := \mathfrak{m}_n + R^{(n+1)}_{Y_{n+1}}, \tag{2.2}$$

where $R^{(n+1)}_{Y_{n+1}}$, conditioned on $\mathfrak{m}_n$, $Y_{n+1} = y$ and the previous history, has the distribution $\mathcal{R}_y := \mathcal{L}(R^{(1)}_y)$.

We assume that $R^{(1)}_x$ is positive on $E \setminus \{x\}$ but allow $R^{(1)}_x(\{x\}) \in (-\infty, \infty)$. We assume that the urn is *tenable*, i.e. that almost surely, $\mathfrak{m}_n$ is a non-zero positive measure for all $n \geq 0$, so $\widetilde{\mathfrak{m}}_n$ and Y_{n+1} are well defined. This is the case if, for example, $\mathfrak{m}_0$ is a non-zero positive measure and each $R^{(1)}_x$ a.s. is a positive measure.

Remark 2.1. We assume for convenience that $\mathfrak{m}_0$ is non-random (except when we explicitly say otherwise). Extensions to random $\mathfrak{m}_0$ follow by conditioning on $\mathfrak{m}_0$, see Remark 2.14 for details. To enable such extensions, some of the results are stated with constants that do not depend on $\mathfrak{m}_0$ (they may depend on the distribution of the replacement kernel $R^{(1)}$), so that the dependence on $\mathfrak{m}_0$ is explicit. Recall that, by convention, C does not depend on $\mathfrak{m}_0$ while $\mathfrak{C}$ may depend on $\mathfrak{m}_0$. The reader who is interested only in a non-random $\mathfrak{m}_0$ may simplify some expressions and arguments by allowing all constants to depend on $\mathfrak{m}_0$. □

Throughout the paper, we also make the following assumptions (B), (H), and (N).

We assume that the urn is *balanced*:

(B) For all $x \in E$, $R^{(1)}_x(E) = 1$ almost surely.

Note that (B) implies that the total mass is deterministic a.s.:

$$\mathfrak{m}_n(E) = \mathfrak{m}_0(E) + n. \tag{2.3}$$

As said above, $R^{(1)}_x$ does not have to be a positive measure. Nevertheless, we will see that (B) and our assumption (H)(ii) below imply that, for every $x \in E$,

$$\mathbb{E}\,\|R^{(1)}_x\| < +\infty \tag{2.4}$$

Hence, we can define the expectation $\mathbb{E}\,R^{(1)}_x$ of the random signed measure $R^{(1)}_x$, which we denote by $\overline{R}_x$, i.e.,

$$\overline{R}_x(A) := \mathbb{E}\big[R^{(1)}_x(A)\big], \qquad A \in \mathcal{E}. \tag{2.5}$$

It follows from (2.4) that $\overline{R}_x$ is a finite signed measure on E and from (B) that $\overline{R}_x(E) = 1$. Moreover, $\overline{R}_x$ is positive on $E \setminus \{x\}$, and it will follow from Assumption (H) below that

$$\sup_{x\in E} |\overline{R}_x(\{x\})| < +\infty. \tag{2.6}$$

In particular, $\overline{R}_x f$ is well defined for all non-negative measurable functions $f : E \to [0, +\infty)$. Note also that $\overline{R}$ is a signed kernel from E to E (see Remark A.1),

Let $W : E \to [1, +\infty)$ be a fixed function and let $V := W^q : E \to [1, +\infty)$ for some fixed $q > 2$. We assume that V and W satisfy the following.

(H) (i) There exists $\vartheta \in (0, 1)$ and $C_1 \geq 0$ such that, for all $x \in E$,

$$\overline{R}_x V \leq \vartheta V(x) + C_1. \tag{2.7}$$

(ii) There exists $C_2 > 0$ such that, for every $x \in E$,

$$\mathbb{E}\left[\Big(|R^{(1)}_x|W\Big)^q\right] \leq C_2 W(x)^q = C_2\,V(x). \tag{2.8}$$

(iii) In addition, $\mathfrak{m}_0 V < +\infty$.

Remark 2.2. An important case is simply to choose $W = 1$, and thus $V = 1$. Note that for $W = 1$, (H) is equivalent to assuming that there exists a constant $C_2 > 0$ such that $\mathbb{E}[\|R^{(1)}_x\|^q] \leq C_2$. In particular, if $W = 1$ and $R^{(1)}_x$ is positive (a.s.) for every $x \in E$, and thus $\|R^{(1)}_x\| = 1$ by (B), then (H) holds automatically. □

Remark 2.3. If $R^{(1)}_x$ a.s. is a positive measure, and thus a probability measure by (B), then Jensen's inequality yields

$$\big(|R^{(1)}_x|W\big)^q = \big(R^{(1)}_x W\big)^q \leq R^{(1)}_x W^q = R^{(1)}_x V. \tag{2.9}$$

Hence, if also (2.7) holds, then

$$\mathbb{E}\big[\big(|R^{(1)}_x|W\big)^q\big] \leq \mathbb{E}\big[R^{(1)}_x V\big] = \overline{R}_x V \leq \vartheta V(x) + C_1 \leq (\vartheta + C_1)V(x), \tag{2.10}$$

i.e., (2.8) holds with $C_2 = \vartheta + C_1$. Consequently, if $R_x^{(1)}$ a.s. is a positive measure (for every $x \in E$), then (H)(ii) follows from (H)(i) and (B).

More generally, if we assume that for some constant C and every $x \in E$,

$$|R_x^{(1)}\{x\}| \le C \qquad \text{a.s.} \tag{2.11}$$

(in other words, subtractions are uniformly bounded), then (B) implies

$$\|R_x^{(1)}\| = |R_x^{(1)}|(E) \le R_x^{(1)}(E) + 2|R_x^{(1)}\{x\}| \le C \qquad \text{a.s.} \tag{2.12}$$

and Hölder's inequality yields

$$|R_x^{(1)}|W \le \bigl(|R_x^{(1)}|W^q\bigr)^{1/q}\bigl(|R_x^{(1)}|1\bigr)^{1-1/q} \le C\bigl(|R_x^{(1)}|V\bigr)^{1/q}. \tag{2.13}$$

Hence, using (2.11) again,

$$\bigl(|R_x^{(1)}|W\bigr)^q \le C|R_x^{(1)}|V \le CR_x^{(1)}V + 2C|R_x^{(1)}\{x\}|V(x) \le CR_x^{(1)}V + CV(x), \tag{2.14}$$

and taking expectations, we obtain

$$\mathbb{E}\bigl[\bigl(|R_x^{(1)}|W\bigr)^q\bigr] \le C\overline{R}_xV + CV(x). \tag{2.15}$$

Consequently, if (2.11) holds, then (H)(ii) follows from (H)(i) and (B). ☐

Remark 2.4. The example in Section 6.1 shows that the assumption (H) is important for our results and cannot be weakened much. In particular, it is not enough to take $q < 2$ above, see Remark 6.6. We do not know whether our results hold with $q = 2$, and leave that as an open problem. ☐

Remark 2.5. By Jensen's inequality, it follows from (H)(ii) that

$$\mathbb{E}[|R_x^{(1)}|W] \le \mathbb{E}\left[\bigl(|R_x^{(1)}|W\bigr)^q\right]^{1/q} \le C_2^{1/q}\,V(x)^{1/q} = CW(x). \tag{2.16}$$

In particular, this implies (2.4) above. Moreover, it also implies that

$$|\overline{R}_x(\{x\})|W(x) \le |\overline{R}_x|W \le \mathbb{E}[|R_x^{(1)}|W] \le CW(x), \tag{2.17}$$

so that $|\overline{R}_x(\{x\})| \le C$, which entails (2.6). ☐

Finally, we assume that, with notation as in Section 1.4,

(N) There exists a probability measure ν such that $\nu\overline{R} = \nu$ and $\nu V < +\infty$.

Let

$$\mathbf{R} : f \mapsto (x \in E \mapsto \overline{R}_x f) \tag{2.18}$$

be the operator corresponding to $\overline{R}$. Since $\overline{R}$ is a signed kernel from E to E, $\mathbf{R}$ maps suitable (e.g. bounded) measurable functions on E to measurable functions on E. As remarked above, the balance assumption (B) implies that $\overline{R}_x(E) = 1$ for every $x \in E$, i.e.,

$$\mathbf{R}1 = 1, \tag{2.19}$$

and Assumption (N) yields

$$\nu\mathbf{R} = \nu. \tag{2.20}$$

We also see that (2.7) can be written $\mathbf{R}V \le \vartheta V + C_1$.

It follows from (2.16) that $\mathbf{R}$ defines a bounded operator on $B(W)$; by default, we regard $\mathbf{R}$ as an operator on $B(W)$ unless we say otherwise. In particular, we let $\sigma(\mathbf{R})$ denote the spectrum of $\mathbf{R}$ on $B(W)$, i.e. the set of all $\lambda \in \mathbb{C}$ such that $\mathbf{R} - \lambda\mathbf{I}$ is not invertible.

In the following theorems, which are our main results, we increase the generality by considering $\mathbf{R}$ as a bounded operator on a closed subspace D of the Banach space $B(W)$ such that $\mathbf{R}(D) \subseteq D$ (i.e., D is *stable*, or *invariant*); the most important case is simply $D = B(W)$. We denote by $\mathbf{R}_D$ the restriction of $\mathbf{R}$ to D, and denote its spectrum by $\sigma(\mathbf{R}_D)$.

To state our main results, we use the following definitions.

Definition 2.6. We say that a bounded operator T on a Banach space $\mathcal{X}$ is *simply logarithmically quasi-compact* (*slqc*) if

(QC1) 1 is an isolated point in $\sigma(T)$, and the corresponding spectral projection $\Pi_1 = \Pi_1(T)$ has rank 1.

(QC2) We have $\sigma(T) \setminus \{1\} \subset \{\lambda : \operatorname{Re} \lambda < 1\}$.

The reason for our name is that the conditions say that the operator e^T is quasi-compact (see Remark B.8) with a single dominating eigenvalue that has a one-dimensional generalized eigenspace; for convenience, we assume that T is normalised such that its dominating eigenvalue is 1.

By (QC1), T maps the one-dimensional subspace $\Pi_1(\mathcal{X})$ into itself. Since the restriction of T to this subspace has spectrum $\{1\}$, it follows that 1 is an eigenvalue of T; moreover, the corresponding eigenvectors are precisely the non-zero elements of $\Pi_1\mathcal{X}$; thus the eigenvector is unique up to a scalar factor. (We can regard (QC1) as a generalisation of the finite-dimensional condition that the eigenvalue 1 has algebraic multiplicity 1.)

Definition 2.7. We say that an operator T is *small* if it is slqc and in addition

(S) $\quad \sigma(T) \setminus \{1\} \subset \{\lambda : \operatorname{Re} \lambda < \frac{1}{2}\}$.

Remark 2.8. This definition of *small* operator is analogous to the terminology used in the context of finitely-many-colour urns: a Pólya urn whose spectral gap is at least half of its spectral radius is called a *small* Pólya urn (see, e.g., [31] where this vocabulary is first used). We comment later on the similarities and differences between our results and the fluctuation results of [18] for Pólya urns. □

We define, for a closed invariant subspace $D \subseteq B(W)$,

$$\theta_D := \sup \operatorname{Re}\big(\sigma(\mathbf{R}_D) \setminus \{1\}\big), \tag{2.21}$$

and, in particular,

$$\theta := \theta_{B(W)} = \sup \operatorname{Re}\big(\sigma(\mathbf{R}) \setminus \{1\}\big). \tag{2.22}$$

Note that if T is a bounded operator on a complex Banach space, then its spectrum $\sigma(T)$ is compact [5, Theorem VII.3.6]. This gives immediately:

Lemma 2.9. (i) *If the operator $\mathbf{R}_D$ is slqc, then $\theta_D < 1$.*
(ii) *If $\mathbf{R}_D$ is slqc, then $\mathbf{R}_D$ is small if and only if $\theta_D < \frac{1}{2}$.* □

The first theorem gives several versions of a law of large numbers for $\widetilde{\mathfrak{m}}_n$.

Theorem 2.10. *Let $(\mathfrak{m}_n)_{n\geq 0}$ be a* MVPP *with initial composition $\mathfrak{m}_0$ and random replacement kernel $R^{(1)}$. Suppose that $R^{(1)}$ satisfies Assumptions* (B), (H), *and* (N). *Let D be a closed invariant subspace of $B(W)$ such that $1 \in D$ and the restriction $\mathbf{R}_D$ of $\mathbf{R}$ to D is slqc.*

(i) *Then $\theta_D < 1$ and, for every $\delta \in (0, 1-\theta_D)$, there exists a constant C_δ such that, for any $f \in D$,*

$$\mathbb{E}\,|\widetilde{\mathfrak{m}}_n f - \nu f|^2 \leq C_\delta\, \widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+n}\right)^{2\delta\wedge 1} \|f\|^2_{B(W)}, \qquad \forall n \geq 1. \tag{2.23}$$

If, in addition, $\delta < 1/2$, then

$$n^\delta |\widetilde{\mathfrak{m}}_n f - \nu f| \xrightarrow[n\to+\infty]{a.s.} 0. \tag{2.24}$$

(ii) *If in addition $\mathbf{R}$ is an slqc operator on $B(W)$, then $\theta < 1$ and, for all $\delta \in (0, 1-\theta)$, for all $f \in B(W^2)$,*

$$\mathbb{E}\,|\widetilde{\mathfrak{m}}_n f - \nu f| \leq C_\delta\, \widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+n}\right)^{(2\delta\wedge 1)\frac{q/2-1}{q-1}} \|f\|_{B(W^2)}, \qquad \forall n \geq 1. \tag{2.25}$$

Remark 2.11. In the case of a metric space E and $D = B(W)$, (2.24) implies $\widetilde{\mathfrak{m}}_n \xrightarrow{\text{a.s.}} \nu$ in the usual weak topology, but it is stronger since it also implies $\widetilde{\mathfrak{m}}_n(A) \xrightarrow{\text{a.s.}} \nu(A)$ for every measurable set A. In particular, we recover and improve on the results obtained in [27] in the balanced non-weighted case. □

Remark 2.12. In the following theorem, we consider the asymptotic distribution of $\widetilde{\mathfrak{m}}_n f$ for a general complex-valued function $f \in D$. In parts (1) and (2) below, the limit is a complex normal distribution, which we describe by identifying $\mathbb{C}$ with $\mathbb{R}^2$; we thus give the covariance matrix of its real and imaginary parts in (2.27) and (2.33). Note that this complex normal distribution in (2.27) and (2.33) can equivalently be characterised as the distribution of a complex normal variable ζ with

$$\mathbb{E}\,\zeta = 0, \qquad \mathbb{E}\,\zeta^2 = \chi(f), \qquad \mathbb{E}\,|\zeta|^2 = \sigma^2(f). \tag{2.26}$$

In applications, we usually consider real f, and then the results simplify since the imaginary parts disappear; in fact, in this case, $\chi(f) = \sigma^2(f)$ is always real, and the limit distributions in (2.27) and (2.33) are just $\mathcal{N}\big(0, \sigma^2(f)\big)$.

If D is closed under complex conjugation, for example if $D = B(W)$, then the results for complex f follow easily from the results for real f by considering real and imaginary parts (and the Cramér–Wold device). Our formulation allows for other interesting domains D, for example $D = \Pi_\lambda B(W) + \mathbb{C}1$ where λ is a non-real isolated point in the spectrum $\sigma(\mathbf{R})$. (See also Example 2.20.) □

The second theorem treats the fluctuations around the limit. As in the finite colour case (see e.g. [18; 31]), there are (under some additional hypotheses) three cases depending on the size of the spectral gap (or, equivalently, on θ_D); in the theorem below we indicate the range of θ_D for each case. Recall that we regard $\mathbf{R}$ as an operator on $B(W)$.

Theorem 2.13. *Let $(\mathfrak{m}_n)_{n\geq 0}$ be a* MVPP *with initial composition $\mathfrak{m}_0$ and random replacement kernel $R^{(1)}$. We assume that $R^{(1)}$ satisfies Assumptions* (B), (H), *and* (N). *Let D be a closed invariant subspace of $B(W)$ such that $1 \in D$ and the restriction $\mathbf{R}_D$ of $\mathbf{R}$ to D is slqc. Then, the following hold.*

(1) *(The case $\theta_D < 1/2$.) If $\mathbf{R}_D$ is small and $\mathbf{R}$ is slqc, then for any $f \in D$,*

$$n^{1/2}(\widetilde{\mathfrak{m}}_n f - \nu f) \xrightarrow{\text{d}} \mathcal{N}\left(0, \frac{1}{2}\begin{pmatrix} \sigma^2(f) + \operatorname{Re}(\chi(f)) & \operatorname{Im}(\chi(f)) \\ \operatorname{Im}(\chi(f)) & \sigma^2(f) - \operatorname{Re}(\chi(f)) \end{pmatrix}\right), \tag{2.27}$$

where

$$\chi(f) := \int_0^\infty \nu\mathbf{B}\big(\mathrm{e}^{s\mathbf{R}}(f - \nu f)\big)\mathrm{e}^{-s}\,\mathrm{d}s = \int_E \int_0^\infty \mathbf{B}_x\big(\mathrm{e}^{s\mathbf{R}}(f - \nu f)\big)\mathrm{e}^{-s}\,\mathrm{d}s\,\mathrm{d}\nu(x) \tag{2.28}$$

and

$$\sigma^2(f) := \int_0^\infty \nu\mathbf{C}\big(\mathrm{e}^{s\mathbf{R}}(f - \nu f)\big)\mathrm{e}^{-s}\,\mathrm{d}s = \int_E \int_0^\infty \mathbf{C}_x\big(\mathrm{e}^{s\mathbf{R}}(f - \nu f)\big)\mathrm{e}^{-s}\,\mathrm{d}s\,\mathrm{d}\nu(x) \tag{2.29}$$

with

$$\mathbf{B}(f) : x \mapsto \mathbf{B}_x(f) := \mathbb{E}\left[\big(R_x^{(1)} f\big)^2\right] \quad \textit{and} \quad \mathbf{C}(f) : x \mapsto \mathbf{C}_x(f) := \mathbb{E}\left[\big|R_x^{(1)} f\big|^2\right] \tag{2.30}$$

and with absolutely convergent integrals.

(2) *(The case $\theta_D = 1/2$.) If $\mathbf{R}_D$ and $\mathbf{R}$ are slqc and the spectrum of $\mathbf{R}_D$ is given by*

$$\sigma(\mathbf{R}_D) = \{1, \lambda_1, \dots, \lambda_p\} \cup \Delta \tag{2.31}$$

for some $p \geq 1$, where $\operatorname{Re}(\lambda_1) = \dots = \operatorname{Re}(\lambda_p) = 1/2$ and $\sup \operatorname{Re}(\Delta) < 1/2$, let

$$\kappa_j := \min\big\{k \geq 1 : (\mathbf{R}_D - \lambda_j \mathbf{I})^k = 0 \text{ on } \Pi_{\lambda_j} D\big\}, \qquad 1 \leq j \leq p, \tag{2.32}$$

and $\kappa := \max_{j \leq p} \kappa_j$. Assume that $\kappa < \infty$. Then, for any $f \in D$,

$$\frac{n^{1/2}}{(\log n)^{\kappa - 1/2}}\big(\widetilde{\mathfrak{m}}_n f - \nu f\big) \xrightarrow{\text{d}} \mathcal{N}\left(0, \frac{1}{2}\begin{pmatrix} \sigma^2(f) + \operatorname{Re}(\chi(f)) & \operatorname{Im}(\chi(f)) \\ \operatorname{Im}(\chi(f)) & \sigma^2(f) - \operatorname{Re}(\chi(f)) \end{pmatrix}\right), \tag{2.33}$$

where

$$\chi(f) := \sum_{j,j'=1}^{p} \frac{\mathbf{1}_{\kappa_j=\kappa_{j'}=\kappa,\ \overline{\lambda_j}=\lambda_{j'}}}{(2\kappa-1)((\kappa-1)!)^2} \nu\widetilde{\mathbf{B}}\left((\mathbf{R}-\lambda_j\mathbf{I})^{\kappa-1}\Pi_{\lambda_j} f, (\mathbf{R}-\lambda_{j'}\mathbf{I})^{\kappa-1}\Pi_{\lambda_{j'}} f\right) \tag{2.34}$$

and

$$\sigma^2(f) := \sum_{j=1}^{p} \frac{\mathbf{1}_{\kappa_j=\kappa}}{(2\kappa-1)((\kappa-1)!)^2} \nu\mathbf{C}\left((\mathbf{R}-\lambda_j\mathbf{I})^{\kappa-1}\Pi_{\lambda_j} f\right) \tag{2.35}$$

with $\mathbf{C}$ *as in* (2.30) *and*

$$\widetilde{\mathbf{B}}(f,g) : x \mapsto \widetilde{\mathbf{B}}_x(f,g) := \mathbb{E}[R_x^{(1)} f \cdot R_x^{(1)} g], \tag{2.36}$$

(3) *(The case $\theta_D > 1/2$.) If $\mathbf{R}_D$ is slqc and the spectrum of $\mathbf{R}_D$ is given by*

$$\sigma(\mathbf{R}_D) = \{1, \lambda_1, \dots, \lambda_p\} \cup \Delta \tag{2.37}$$

for some $p \geq 1$, where $\mathrm{Re}(\lambda_1) = \cdots = \mathrm{Re}(\lambda_p) \in (1/2, 1)$, and $\sup \mathrm{Re}(\Delta) < \mathrm{Re}(\lambda_1)$, let κ_j $(1 \leq j \leq p)$ be defined by (2.32) *and let $\kappa := \max_{j\leq p} \kappa_j$. Assume that $\kappa < \infty$. Then, for any $f \in D$, there exist complex random variables $\Lambda_1, \dots, \Lambda_p \in L^2$ such that*

$$\frac{n^{1-\mathrm{Re}\,\lambda_1}}{(\log n)^{\kappa-1}} \left(\widetilde{\mathfrak{m}}_n f - \nu f\right) - \sum_{j=1}^{p} n^{\mathrm{i}\,\mathrm{Im}\,\lambda_j} \Lambda_j \to 0 \tag{2.38}$$

a.s. and in L^2. Furthermore,

$$\mathbb{E}\,\Lambda_j = \frac{\Gamma(\mathfrak{m}_0(E)+1)}{(\kappa-1)!\,\Gamma(\mathfrak{m}_0(E)+\lambda_j)} \widetilde{\mathfrak{m}}_0(\mathbf{R}-\lambda_j)^{\kappa-1}\Pi_{\lambda_j} f. \tag{2.39}$$

Remark 2.14. To adapt our results to a random $\mathfrak{m}_0$, we make the same assumptions as in Theorems 2.10 and 2.13, and we assume that (H)(iii) holds almost surely. Under these assumptions, Theorems 2.10 and 2.13 apply conditionally on $\mathfrak{m}_0$. This implies that, under these assumptions, the almost-sure convergences in Theorems 2.10 and 2.13 still hold for random $\mathfrak{m}_0$. Furthermore, since the limiting distributions in Theorem 2.13 (1) and (2) do not depend on $\mathfrak{m}_0$, it also implies that the convergences in distribution in Theorem 2.13(1) and (2) hold if $\mathfrak{m}_0$ is random. If in addition $\mathbb{E}\,\widetilde{\mathfrak{m}}_0 V < +\infty$, then, by dominated convergence, the left-hand-side terms of (2.23) and (2.25) also converge to 0 when $n \to +\infty$.

Under the assumptions of Theorem 2.13(3), conditioning on $\mathfrak{m}_0$ shows that (2.38) still holds a.s. for a random $\mathfrak{m}_0$, with (2.39) replaced by

$$\mathbb{E}\,\Lambda_j = \mathbb{E}\left[\frac{\Gamma(\mathfrak{m}_0(E)+1)}{(\kappa-1)!\,\Gamma(\mathfrak{m}_0(E)+\lambda_j)} \widetilde{\mathfrak{m}}_0(\mathbf{R}-\lambda_j)^{\kappa-1}\Pi_{\lambda_j} f\right]. \tag{2.40}$$

Moreover, it follows from the proof that under the additional assumption that

$$\mathbb{E}\left[(\mathfrak{m}_0(E)+1)^{2(1-\mathrm{Re}\,\lambda_1)} \widetilde{\mathfrak{m}}_0 V\right] < \infty, \tag{2.41}$$

(2.38) holds also in L^2, see Remark 4.6. (Note that $\mathbb{E}\left[(\mathfrak{m}_0(E)+1)\widetilde{\mathfrak{m}}_0 V\right] < \infty$ suffices for (2.41).) □

Remark 2.15. The operator $\mathbf{B}$ defined in (2.30) is the quadratic operator corresponding to the bilinear operator $\widetilde{\mathbf{B}}$ in (2.36), i.e., $\mathbf{B}(f) = \widetilde{\mathbf{B}}(f,f)$. Similarly, $\mathbf{C}(f) = \widetilde{\mathbf{B}}(f,\overline{f})$. It follows from (2.8) that the bilinear map $\widetilde{\mathbf{B}}$ is bounded, and thus continuous, as a mapping $B(W) \times B(W) \to B(W^2)$. Indeed, for all $f, g \in B(W)$ such that $\|f\|_{B(W)} = \|g\|_{B(W)} = 1$, we have, for all $x \in E$,

$$\mathbb{E}[|R_x^{(1)} f \cdot R_x^{(1)} g|] \leq \mathbb{E}\left[(|R_x^{(1)}|W)^2\right] \leq \mathbb{E}\left[(|R_x^{(1)}|W)^q\right]^{2/q} \leq C\,W(x)^2, \tag{2.42}$$

where we used Jensen's inequality and (2.8). Hence, $\mathbf{B}$ and $\mathbf{C}$ are continuous maps $B(W) \to B(W^2)$, and

$$\|\mathbf{B}(f)\|_{B(W^2)} \le C\|f\|^2_{B(W)}, \qquad \|\mathbf{C}(f)\|_{B(W^2)} \le C\|f\|^2_{B(W)}. \tag{2.43}$$

□

Remark 2.16. Just as in the finitely-many-colour case, the limit result (2.38) implies convergence in distribution of $\frac{n^{1-\operatorname{Re}\lambda_1}}{(\log n)^{\kappa-1}}\big(\tilde{\mathfrak{m}}_n f - \nu f\big)$ for suitable subsequences, but in general not for the full sequence. □

Remark 2.17. The asymptotic normality in parts (1) and (2) extends immediately to joint convergence for any number of $f \in D$, by the Cramér–Wold device [12, Theorem 5.10.5]; the asymptotic covariances are given by obvious bilinear analogues of the variance formulas in the theorem (cf. Remark 2.15).

In part (3), joint (subsequence) convergence in distribution for several $f \in D$ is immediate from the a.s. convergence in (2.38). □

Remark 2.18. The assumption $\kappa < \infty$ in parts (2) and (3) holds, in particular, if we have $\dim(\Pi_{\lambda_j} D) < \infty$ for each $j \le p$. To see this, let $D_j := \Pi_{\lambda_j} D$, and note that if $\dim(D_j) < \infty$, then the restriction R_{D_j} is an operator in the finite-dimensional vector space D_j with spectrum $\sigma(R_{D_j}) = \{\lambda_j\}$; hence the operator $R_{D_j} - \lambda_j \mathbf{I}$ is nilpotent (as is shown by the Jordan decomposition), and thus κ_j in (2.32) is finite; in fact,

$$\kappa_j \le \dim(D_j). \tag{2.44}$$

Hence, $\kappa \le \max_j \dim(D_j) < \infty$ if all $D_j = \Pi_{\lambda_j} D$ have finite dimensions. □

Remark 2.19. Note that allowing a domain $D \subset B(W)$ leads to a more complete result. For instance, if Δ is a clopen subset of $\sigma(\mathbf{R})$, then one can consider the operator $\mathbf{R}_D$ acting on $D = \Pi_\Delta B(W) + \mathbb{C}1$, whose spectrum is $\{1\} \cup \{\Delta\}$ which may be strictly included in $\sigma(\mathbf{R})$. In that case, the assumptions in Theorem 2.10(1)–(3) on $\mathbf{R}_D$ become assumptions on Δ, and then the theorem yields results for $f \in D$, even if the assumptions are not satisfied for $\sigma(\mathbf{R})$.

For another example where subspaces are useful, see Remark 6.13. □

Example 2.20. We give a simple example; further examples are given in Section 6. Suppose that $\mathbf{R}$ is slqc in $B(W)$, and that $f \in B(W)$ is an eigenfunction: $\mathbf{R}f = \lambda f$ with $\lambda \neq 1$. Then Theorem 2.10 applies to the two-dimensional space D spanned by f and 1.

If $\operatorname{Re}\lambda < 1/2$, then (1) yields the asymptotic normality (2.27). We have $\nu f = 0$ and $e^{s\mathbf{R}} f = e^{s\lambda} f$, and thus (2.28) and (2.29) yield $\chi(f) = (1-2\lambda)^{-1}\nu\mathbf{B}(f)$ and $\sigma^2(f) = (1 - 2\operatorname{Re}\lambda)^{-1}\nu\mathbf{C}(f)$.

If $\operatorname{Re}\lambda = 1/2$, then (2) applies instead, with $p = 1$ and $\kappa = 1$; (2.34) and (2.35) yield $\chi(f) = \nu\mathbf{B}(f)$ and $\sigma^2(f) = \nu\mathbf{C}(f)$.

Finally, if $\operatorname{Re}\lambda > 1/2$, then (3) applies, with $\sigma(\mathbf{R}_D) = \{1, \lambda\}$ and $\kappa = 1$; (2.38) shows that there exists a complex random variable Λ such that $n^{1-\operatorname{Re}(\lambda)}(\tilde{\mathfrak{m}}_n f - \nu f) - n^{\mathrm{i}\operatorname{Im}(\lambda)}\Lambda \to 0$, and hence $n^{1-\lambda}(\tilde{\mathfrak{m}}_n f - \nu f) \to \Lambda$, almost surely and in L^2 when $n \to +\infty$. □

Remark 2.21. In Theorem 2.13, the assumption that $1 \in D$ is in fact not necessary. We make this assumption for convenience and because, in practice, as one can see in Example 2.20, 1 can always be added to D to enter the setting of our results. □

Example 2.22. The classical generalised Pólya urn model with finitely-many colours is given by $E = \{1, \dots, d\}$ and $R^{(1)}_x = (\mathfrak{r}^{(1)}_{x,1}\delta_1 + \dots + \mathfrak{r}^{(1)}_{x,d}\delta_d)/S$, where $\mathfrak{r}^{(1)} = (\mathfrak{r}^{(1)}_{x,y})_{1\le x,y\le d}$ is a (possibly random) matrix of integers, with $\mathfrak{r}^{(1)}_{x,y} \ge 0$ when $x \neq y$, $\mathfrak{r}^{(1)}_{x,x} \ge -1$ for all $1 \le x \le d$, δ_x is a point mass (Dirac measure) at x, and S is a scaling factor (for convenience). We apply our results to that case and compare the outcome to results from the literature. This model satisfies

(B) if and only if the replacement matrix $\mathfrak{r}^{(1)}$ is "balanced", i.e. if all row sums are equal to S (a.s.);

(H) always when (B) holds, since then $-1 \le \mathfrak{r}^{(1)}_{x,y} \le S+1$ a.s. (We take $V \equiv 1$.)

(N) always when (B) holds, since then the non-negative matrix $\mathbb{E}\,\mathfrak{r}^{(1)} + \mathbf{I} = (\mathbb{E}[\mathfrak{r}^{(1)}_{x,y} + \delta_{x,y}])_{x,y}$ has a positive right eigenvector with eigenvalue $S+1$ (viz. 1), and therefore it follows from the the Perron–Frobenius theorem that it also has a non-negative left eigenvector $u = (u_x)_1^d$ with this eigenvalue; we may assume that $\sum_x u_x = 1$ and then take $\nu = \sum_x u_x \delta_x$.

Furthermore, $B(1) = B(E)$ is the space of all functions from $\{1, \dots, d\}$ to $\mathbb{C}$, i.e. $\mathbb{C}^d$. Under (B), the operator $\mathbf{R}$ defined by $\bar{\mathfrak{r}}/S = \mathbb{E}[\mathfrak{r}^{(1)}]/S$ on $\mathbb{C}^d$ is slqc if and only if the eigenvalue 1 has (algebraic) multiplicity 1. Under these assumptions, (2.23)–(2.24) imply that, if $\mathfrak{u}_n$ is the composition vector of the urn at time n, i.e. the vector whose i-th coordinate is the number of balls of colour i in the urn at time n, then

$$\|\mathfrak{u}_n - v\| = o(n^{-\delta}) \quad \text{a.s. and in } L^2 \text{ as } n \to \infty, \tag{2.45}$$

for all $\delta \in (0, (1-\theta) \wedge 1/2)$, where θ is the maximum of the real parts of the eigenvalues of $\mathfrak{r}$ excluding 1. Furthermore, Theorem 2.13(1) and (2) allow us to recover versions of the limit theorems [18, Theorems 3.22, 3.23 and 3.24]: the only caveat is that we make the additional assumption that the replacement matrix is balanced. □

Remark 2.23. As mentioned in the introduction, it is standard in the theory of Pólya urns to associate different weights (or activities) to the balls of different colours. In this generalisation, when picking a ball at random at the n-th step, one pick each of the balls with probability proportional to its weight, and then applies the replacement rule as normal depending on the colour of the drawn ball.

In [27], the authors generalise this concept of weight in the infinitely-many-colour case: for all positive kernel $\mathbf{P} = (P_x)_{x \in E}$, they define the MVPP $\mathfrak{m}_n$ as in (2.2), except that, conditionally on $\mathfrak{m}_n$, Y_{n+1} is drawn according to the distribution $\mathfrak{m}_n\mathbf{P}/\mathfrak{m}_n\mathbf{P}(E)$.

One can apply our main results (Theorems 2.10 and 2.13) to $\mathfrak{m}'_n := \mathfrak{m}_n\mathbf{P}$, which is an MVPP of replacement kernel $\mathbf{RP}$. Our assumptions require in particular that $\mathbf{RP}$ satisfies the balance assumption (B). From our main results applied to $\mathfrak{m}'_n$, if the operator induced by $\mathbf{P}$ is invertible, one can deduce a fluctuation result for the original weighted MVPP $\mathfrak{m}_n$.

Even if $\mathbf{P}$ were non-invertible, it would be straightforward to generalise our proofs to the weighted case under the assumption that $\mathbf{RP}$ is balanced; since our proofs are already technical, and since the balance assumption restricts greatly the set of weighted kernels one could use, we do not extend our framework to include this case. □

Remark 2.24. In the theorems above, we regard $\mathbf{R}$ as an operator on $B(W)$, where the possibility to choose a suitable W gives additional flexibility. (Warning: the spectrum $\sigma(\mathbf{R})$, and thus e.g. θ, may change if we change W, see the example in Section 6.1.) The space $B(W)$ seems natural and convenient for applications, but it is not the only reasonable choice of a function space.

First, in typical cases, we may ignore functions that are 0 ν-a.e. and it is then equivalent to consider $\mathbf{R}$ as an operator on the quotient space of $B(W)$ modulo functions that are 0 ν-a.e., which we denote by $L^\infty(W;\nu)$; see Lemma 5.1 which implies that $\mathbf{R}$ always is well defined on $L^\infty(W;\nu)$. However, Example 5.2 shows that there are (exceptional) cases when null sets and functions cannot be completely ignored.

More importantly, the examples in Sections 6.2 and 6.3 use Fourier analysis and it is then convenient to consider $\mathbf{R}$ as an operator on $L^2(E,\nu)$. In these examples we transfer spectral properties of $\mathbf{R}$ from $L^2(E,\nu)$ to $B(E)$ and then apply the theorems above. However, for these and other similar examples, it would be desirable to have extensions of the theorems above where $B(W)$ is replaced by a more general function space on E, including $L^2(E,\nu)$ as a possible choice. (Other choices might also be useful in other applications.) In the present paper, however, we consider only the theorems as stated above, with $\mathbf{R}$ acting on $B(W)$ (and invariant subspaces thereof). □

2.1. **Degenerate limits?** The limit results in Theorem 2.13 are less interesting when the limit distribution is identically 0. We characterize here these degenerate cases, and begin by showing that in part (1) of Theorem 2.13, the limit is non-degenerate except in trivial cases. Proofs are given in Section 5.

Theorem 2.25. *Suppose that the conditions of Theorem 2.13(1) hold, and let $f \in D$. Let $\Sigma(f)$ be the covariance matrix in* (2.27)*. Then the following are equivalent:*

(i) $\Sigma(f) = 0$.
(ii) $\sigma^2(f) = 0$.
(iii) $\nu\mathbf{C}(f - \nu f) = 0$.
(iv) $R_x^{(1)} f = \nu f$ *a.s., for ν-a.e. x.*

On the contrary, in Theorem 2.13(2), the asymptotic distribution depends only on $\Pi_{\lambda_1} f, \dots, \Pi_{\lambda_p} f$, and thus it degenerates to 0 for many f. (For such f, it might be possible to apply the theorem with a smaller space D.) In fact, in typical applications, the projections Π_{λ_j} project onto finite-dimensional subspaces, and thus their kernels are very large. We have the following characterization.

Theorem 2.26. *Suppose that the conditions of Theorem 2.13(2) hold, and let $f \in D$. Let $\Sigma(f)$ be the covariance matrix in* (2.33)*. Then the following are equivalent:*

(i) $\Sigma(f) = 0$.
(ii) $\sigma^2(f) = 0$.
(iii) $(\mathbf{R} - \lambda_j \mathbf{I})^{\kappa-1} \Pi_{\lambda_j} f = 0$ *ν-a.e., for every $j = 1, \dots, p$.*
(iv) $(\mathbf{R} - \lambda_j \mathbf{I})^{\kappa-1} \Pi_{\lambda_j} f = 0$ *ν-a.e., for every $j = 1, \dots, p$ such that $\kappa_j = \kappa$.*

In Theorem 2.13(3), the situation is similar to Theorem 2.13(2). The characterization is more technical, partly because the limit distribution now also depends on the initial values $\mathfrak{m}_0$; we give several equivalent conditions. Note that the sum $\sum_j n^{\mathrm{i}\,\mathrm{Im}\,\lambda_j} \Lambda_j$ in (2.38) vanishes for all n if and only if $\Lambda_j = 0$ for each $j = 1, \dots, p$. In typical cases, the conditions below are satisfied only for $g_j = 0$, but Example 5.2 gives an example where g_j is non-zero and the conditions are satisfied for some, but not all, $\mathfrak{m}_0$.

Theorem 2.27. *Suppose that the conditions of Theorem 2.13(3) hold, and let $f \in D$. Let Λ_j be as in* (2.38)*, and let $g_j := (\mathbf{R} - \lambda_j)^{\kappa-1} \Pi_{\lambda_j} f$. Then the following are equivalent, for each $j = 1, \dots, p$:*

(i) Λ_j *is (a.s.) non-random.*
(ii) $\Lambda_j = \mathbb{E}\,\Lambda_j$ *a.s.*
(iii) $\Lambda_j = 0$ *a.s.*
(iv) $\mathfrak{m}_n g_j = 0$ *a.s., for every $n \geq 0$.*
(v) $\mathfrak{m}_0 g_j = 0$ *and* $R_{Y_{n+1}}^{(n+1)} g_j = 0$ *a.s., for every $n \geq 0$.*
(vi) $\mathfrak{m}_n |g_j| = 0$ *a.s., for every $n \geq 0$.*
(vii) $\mathfrak{m}_n \{x : |g_j(x)| \neq 0\} = 0$ *a.s., for every $n \geq 0$.*
(viii) $\mathfrak{m}_0 \mathbf{R}^n |g_j| = 0$*, for every $n \geq 0$.*

Moreover, if $\mathbf{R}$ is slqc on $B(W)$, then (i)–(viii) *imply*

(ix) $g_j = 0$ *ν-a.e.*

Conversely, if $\mathfrak{m}_0$ is absolutely continuous w.r.t. ν, then (ix) *implies* (i)–(viii).

Remark 2.28. It follows from Theorems 2.26 and 2.27 that when considering joint limits for several $f \in D$ in parts (2) and (3) of Theorem 2.13 (see Remark 2.17), the limit distribution is supported on a subspace of dimension at most

$$\sum_{j=1}^{p} \dim\bigl[(\mathbf{R} - \lambda_j)^{\kappa-1} \Pi_{\lambda_j}(D)\bigr] \tag{2.46}$$

with equality in typical cases (we leave the details to the reader). Note that this is always at most $\sum_{j=1}^{p} \dim\bigl[\Pi_{\lambda_j}(D)\bigr]$. □

3. Proof of Theorem 2.10

We assume throughout this section that Assumptions (B), (H), and (N) hold. Recall that $\mathfrak{m}_0$ is non-random (unless we explicitly say otherwise), and that constants C do not depend on $\mathfrak{m}_0$. The claims about θ and θ_D in Theorem 2.10 follow by Lemma 2.9.

3.1. **Preliminary results.** Define, for $n \geq 0$, the random signed measure

$$\mathfrak{v}_n := \widetilde{\mathfrak{m}}_n - \nu. \tag{3.1}$$

Lemma 3.1. *For all $n \geq 0$,*

$$\mathfrak{v}_n = \mathfrak{v}_0 B_{0,n} + \sum_{i=1}^{n} \gamma_{i-1} \Delta M_i B_{i,n}, \tag{3.2}$$

where, for all $n \geq 0$ and $0 \leq i \leq n$,

$$B_{i,n} := \prod_{j=i}^{n-1} (\mathbf{I} + \gamma_j(\mathbf{R} - \mathbf{I})), \tag{3.3}$$

$$\Delta M_{n+1} := R^{(n+1)}_{Y_{n+1}} - \mathbb{E}_n R^{(n+1)}_{Y_{n+1}} = R^{(n+1)}_{Y_{n+1}} - \mathbb{E}_n \overline{R}_{Y_{n+1}} = R^{(n+1)}_{Y_{n+1}} - \widetilde{\mathfrak{m}}_n \mathbf{R}, \tag{3.4}$$

$$\gamma_n := \frac{1}{n+1+\mathfrak{m}_0(E)}. \tag{3.5}$$

Proof. By definition, for any $n \geq 0$, we have (2.2), where the conditional distribution of Y_{n+1} given $\mathfrak{m}_n$ is $\widetilde{\mathfrak{m}}_n$. Furthermore, (B) implies, see (2.3),

$$\mathfrak{m}_{n+1}(E) = \mathfrak{m}_0(E) + n + 1 = 1/\gamma_n. \tag{3.6}$$

Together with (3.4), this implies

$$\begin{aligned}
\widetilde{\mathfrak{m}}_{n+1} &= \frac{\mathfrak{m}_{n+1}}{\mathfrak{m}_{n+1}(E)} = \frac{\mathfrak{m}_n}{\mathfrak{m}_{n+1}(E)} + \gamma_n R^{(n+1)}_{Y_{n+1}} \\
&= \frac{\mathfrak{m}_n(E)}{\mathfrak{m}_{n+1}(E)} \cdot \widetilde{\mathfrak{m}}_n + \gamma_n R^{(n+1)}_{Y_{n+1}} = (1-\gamma_n)\widetilde{\mathfrak{m}}_n + \gamma_n R^{(n+1)}_{Y_{n+1}} \\
&= \widetilde{\mathfrak{m}}_n + \gamma_n \widetilde{\mathfrak{m}}_n (\mathbf{R} - \mathbf{I}) + \gamma_n \Delta M_{n+1},
\end{aligned} \tag{3.7}$$

By definition, $\mathfrak{v}_n = \widetilde{\mathfrak{m}}_n - \nu$, and by (2.20), we have $\nu(\mathbf{R} - \mathbf{I}) = \nu\mathbf{R} - \nu = 0$; therefore, (3.7) implies

$$\mathfrak{v}_{n+1} = \mathfrak{v}_n + \gamma_n \mathfrak{v}_n(\mathbf{R} - \mathbf{I}) + \gamma_n \Delta M_{n+1} = \mathfrak{v}_n\big(\mathbf{I} + \gamma_n(\mathbf{R} - \mathbf{I})\big) + \gamma_n \Delta M_{n+1}, \tag{3.8}$$

and (3.2) follows by induction. □

As noted above, it follows from (2.16) that $\mathbf{R}$ is a bounded operator on $B(W)$; hence every $B_{i,n}$ is too. Dually, $\mathbf{R}$ and $B_{n,i}$ are bounded operators on $\mathcal{M}(W)$ (acting on the right). Moreover:

Lemma 3.2. *A.s., for every $n \geq 0$, we have $\mathfrak{m}_n, \widetilde{\mathfrak{m}}_n, \mathfrak{v}_n, R^{(n+1)}_{Y_{n+1}}, \Delta M_{n+1} \in \mathcal{M}(W)$. Moreover,*

$$\mathbb{E}\big[\mathfrak{m}_n W\big] < \infty \qquad \text{and} \qquad \mathbb{E}\big[|\mathfrak{v}_n| W\big] < \infty, \tag{3.9}$$

and

$$\sup_{n\geq 0} \mathbb{E}\, \widetilde{\mathfrak{m}}_n V \leq C \widetilde{\mathfrak{m}}_0 V \qquad \text{and} \qquad \sup_{n \geq 1} \mathbb{E}\, V(Y_n) \leq C \widetilde{\mathfrak{m}}_0 V. \tag{3.10}$$

Finally, there exists a constant C such that for every $g \in B(W)$

$$|\mathfrak{v}_0 g|^2 \leq C \|g\|^2_{B(W)} (\widetilde{\mathfrak{m}}_0 W)^2 \leq C \|g\|^2_{B(W)} \widetilde{\mathfrak{m}}_0 V, \tag{3.11}$$

$$\mathbb{E}\big|\Delta M_i g\big|^2 \leq C \|g\|^2_{B(W)} \widetilde{\mathfrak{m}}_0 V, \qquad i \geq 1, \tag{3.12}$$

$$\mathbb{E}_{i-1}\big|\Delta M_i g\big|^q \leq C \|g\|^q_{B(W)} \widetilde{\mathfrak{m}}_{i-1}(V), \qquad i \geq 1. \tag{3.13}$$

Proof. **We start by proving $\mathfrak{m}_n, \widetilde{\mathfrak{m}}_n, \mathfrak{v}_n, R^{(n+1)}_{Y_{n+1}}, \Delta M_{n+1} \in \mathcal{M}(W)$ and (3.9).** By construction, the conditional distribution of $R^{(n+1)}_{Y_{n+1}}$ given $\mathcal{F}_n$ and $Y_{n+1} = y$, for some $y \in E$, equals the distribution of $R^{(1)}_y$. Hence, using (H)(ii) through its consequence (2.16),

$$\mathbb{E}\big[|R^{(n+1)}_{Y_{n+1}}W| \bigm| \mathcal{F}_n, Y_{n+1} = y\big] = \mathbb{E}\big|R^{(1)}_y W\big| \le \mathbb{E}\big[|R^{(1)}_y|W\big] \le CW(y). \tag{3.14}$$

In other words, a.s.,

$$\mathbb{E}\big[|R^{(n+1)}_{Y_{n+1}}W| \bigm| \mathcal{F}_n, Y_{n+1}\big] \le CW(Y_{n+1}). \tag{3.15}$$

Furthermore, Y_{n+1} has the conditional distribution $\widetilde{\mathfrak{m}}_n$ given $\mathcal{F}_n$, and thus, by taking the conditional expectation $\mathbb{E}_n$ in (3.15),

$$\begin{aligned}\mathbb{E}_n \big|R^{(n+1)}_{Y_{n+1}}W\big| = \mathbb{E}\big[|R^{(n+1)}_{Y_{n+1}}W| \bigm| \mathcal{F}_n\big] \le C\,\mathbb{E}\big[W(Y_{n+1}) \bigm| \mathcal{F}_n\big] = C\int_E W(y)\,\mathrm{d}\widetilde{\mathfrak{m}}_n(y)\\ = C\widetilde{\mathfrak{m}}_n W. \end{aligned}\tag{3.16}$$

Hence, by (2.2), (2.1) and (2.3),

$$\begin{aligned}\mathbb{E}\,\mathfrak{m}_{n+1}W &= \mathbb{E}\,\mathfrak{m}_n W + \mathbb{E}\,R^{(n+1)}_{Y_{n+1}}W \le \mathbb{E}\,\mathfrak{m}_n W + C\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n W\\ &= \Big(1 + \frac{C}{\mathfrak{m}_0(E)+n}\Big)\,\mathbb{E}\,\mathfrak{m}_n W.\end{aligned}\tag{3.17}$$

Hence, the first part of (3.9) follows by induction, since $\mathfrak{m}_0 W \le \mathfrak{m}_0 V < \infty$ by (H)(iii) (recall that $W = V^{1/q}$, $q > 2$, and $V \ge 1$).

Consequently, for every n, a.s., $\mathfrak{m}_n W < \infty$ and thus $\mathfrak{m}_n \in \mathcal{M}(W)$, recalling that $\mathfrak{m}_n$ is a positive measure. Hence, also $\widetilde{\mathfrak{m}}_n \in \mathcal{M}(W)$ and, by (2.2), $R^{(n+1)}_{Y_{n+1}} \in \mathcal{M}(W)$.

Since $\mathbf{R}$ acts on $\mathcal{M}(W)$ as noted above, we further obtain $\widetilde{\mathfrak{m}}_n\mathbf{R} \in \mathcal{M}(W)$ and thus (3.4) yields $\Delta M_{n+1} \in \mathcal{M}(W)$ a.s.

Finally, (3.1) implies that $|\mathfrak{v}_n| \le \widetilde{\mathfrak{m}}_n + \nu$ and (N) shows that $\nu \in \mathcal{M}(V) \subseteq \mathcal{M}(W)$; hence, $\mathfrak{v}_n W \in \mathcal{M}(W)$ a.s. and $\mathbb{E}\big[|\mathfrak{v}_n|W\big] < \infty$ follow from the results for $\mathfrak{m}_n$ and $\widetilde{\mathfrak{m}}_n$.

We now prove (3.10). Recall that, by Assumption (H), $\mathbf{R}V \le \vartheta V + C_1$. Taking expectations in (3.7), since $\mathbb{E}\,\Delta M_{n+1} = 0$, we obtain

$$\mathbb{E}\,\widetilde{\mathfrak{m}}_{n+1} = \mathbb{E}\,\widetilde{\mathfrak{m}}_n + \gamma_n\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n(\mathbf{R}-\mathbf{I}) = (1-\gamma_n)\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n + \gamma_n\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n\mathbf{R} \tag{3.18}$$

and thus

$$\begin{aligned}\mathbb{E}\,\widetilde{\mathfrak{m}}_{n+1}V &= (1-\gamma_n)\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n V + \gamma_n\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n\mathbf{R}V \le (1-\gamma_n)\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n V + \gamma_n\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n(\vartheta V + C_1)\\ &= (1-\gamma_n+\gamma_n\vartheta)\,\mathbb{E}\,\widetilde{\mathfrak{m}}_n V + \gamma_n C_1.\end{aligned}\tag{3.19}$$

Recall that $\widetilde{\mathfrak{m}}_0 V < \infty$ by (H)(iii). Let $C_0 := \widetilde{\mathfrak{m}}_0 V \vee \frac{C_1}{1-\vartheta}$. Then $\mathbb{E}\,\widetilde{\mathfrak{m}}_n V \le C_0$ by induction; indeed, the induction hypothesis and (3.19) yield

$$\mathbb{E}\,\widetilde{\mathfrak{m}}_{n+1}V \le (1-(1-\vartheta)\gamma_n)C_0 + \gamma_n C_1 = C_0 + \gamma_n(C_1 - (1-\vartheta)C_0) \le C_0. \tag{3.20}$$

Because $\widetilde{\mathfrak{m}}_0 V \ge 1$, we have that $C_0 \le (1 + \frac{C_1}{1-\vartheta})\widetilde{\mathfrak{m}}_0 V$, which proves the first part of (3.10). Finally, since Y_n has the conditional distribution $\widetilde{\mathfrak{m}}_{n-1}$ given $\mathcal{F}_{n-1}$, this implies

$$\mathbb{E}\,V(Y_n) = \mathbb{E}[\mathbb{E}_{n-1}\,V(Y_n)] = \mathbb{E}[\widetilde{\mathfrak{m}}_{n-1}V] \le C\widetilde{\mathfrak{m}}_0 V. \tag{3.21}$$

It only remains to prove (3.11), (3.12) and (3.13). For (3.11) note that $\mathfrak{v}_0 = \widetilde{\mathfrak{m}}_0 - \nu$ and thus

$$\begin{aligned}|\mathfrak{v}_0 g|^2 &\le 2\,(\widetilde{\mathfrak{m}}_0|g|)^2 + 2(\nu|g|)^2 \le 2\,\big((\widetilde{\mathfrak{m}}_0 W)^2 + (\nu W)^2\big)\,\|g\|^2_{B(W)}\\ &\le C(\widetilde{\mathfrak{m}}_0 W)^2\|g\|^2_{B(W)} \le C\widetilde{\mathfrak{m}}_0 V\,\|g\|^2_{B(W)},\end{aligned}\tag{3.22}$$

where we used the fact that $(\widetilde{\mathfrak{m}}_0 W)^2 \ge 1$ and $(\nu W)^2 \le \nu(W^2) \le \nu V < +\infty$ by (N), and similarly $(\widetilde{\mathfrak{m}}_0 W)^2 \le \widetilde{\mathfrak{m}}_0(W^2) \le \widetilde{\mathfrak{m}}_0 V$.

For $i \geq 1$, by the definition (3.4),

$$\mathbb{E}_{i-1}\bigl|\Delta M_i g\bigr|^q = \mathbb{E}_{i-1}\bigl|R^{(i)}_{Y_i} g - \mathbb{E}_{i-1} R^{(i)}_{Y_i} g\bigr|^q \leq C\ \mathbb{E}_{i-1}\bigl|R^{(i)}_{Y_i} g\bigr|^q \leq C\|g\|^q_{B(W)}\, \mathbb{E}_{i-1}\bigl(|R^{(i)}_{Y_i}|W\bigr)^q. \tag{3.23}$$

Furthermore, arguing as in (3.14)–(3.16), now using (2.8), gives

$$\mathbb{E}_{i-1}\bigl(|R^{(i)}_{Y_i}|W\bigr)^q \leq C_2 \int_E V(x)\,\mathrm{d}\widetilde{\mathfrak{m}}_{i-1}(x) = C_2 \widetilde{\mathfrak{m}}_{i-1} V. \tag{3.24}$$

Therefore, (3.13) follows by (3.23) and (3.24). By taking the expectation in (3.13) and using (3.10), we obtain

$$\mathbb{E}\bigl|\Delta M_i g\bigr|^q \leq C\|g\|^q_{B(W)}\, \widetilde{\mathfrak{m}}_0 V. \tag{3.25}$$

Since $q > 2$, (3.12) follows from (3.25) by Jensen's inequality and since $\widetilde{\mathfrak{m}}_0 V \geq 1$. □

Lemma 3.1 implies that if $f \in B(W)$, then (with all terms a.s. finite and integrable by Lemma 3.2)

$$\mathfrak{v}_n f = \mathfrak{v}_0 B_{0,n} f + \sum_{i=1}^{n} \gamma_{i-1} \Delta M_i B_{i,n} f. \tag{3.26}$$

Note that the sequence of partial sums of (3.26) is a martingale, since $\mathbb{E}_{i-1}\,\Delta M_i = 0$. For later use, note that (2.19) and (2.20) imply

$$B_{i,n} 1 = 1, \tag{3.27}$$

$$\nu B_{i,n} = \nu. \tag{3.28}$$

We write (3.26) as

$$\mathfrak{v}_n f = \zeta_{n,0} + \sum_{i=1}^{n} \zeta_{n,i}, \tag{3.29}$$

where

$$\zeta_{n,0} = \zeta_{n,0}(f) := \mathfrak{v}_0 B_{0,n} f, \tag{3.30}$$

$$\zeta_{n,i} = \zeta_{n,i}(f) := \gamma_{i-1} \Delta M_i B_{i,n} f, \qquad 1 \leq i \leq n. \tag{3.31}$$

The main part of the proof is to use the assumptions to show that the random variables $\zeta_{n,i}$ are suitably small. Note that

$$\mathfrak{v}_n 1 = \widetilde{\mathfrak{m}}_n 1 - \nu 1 = 0, \tag{3.32}$$

since both $\widetilde{\mathfrak{m}}_n$ and ν are probability measures. Note also that (B) implies that

$$\Delta M_i 1 = R^{(i)}_{Y_i} 1 - \mathbb{E}_{i-1} R^{(i)}_{Y_i} 1 = 1 - 1 = 0 \qquad \text{a.s.} \tag{3.33}$$

Hence (3.30)–(3.31) and (3.27) show that taking $f = 1$, we obtain $\zeta_{n,i}(1) = 0$ a.s. for every $i \geq 0$. Consequently, by linearity, for any $i \geq 0$ and any constant c,

$$\zeta_{n,i}(f) = \zeta_{n,i}(f - c) \qquad \text{a.s.} \tag{3.34}$$

Recall (see [5, (VII.4.5)] or [7, Section VII.3]) that if T is a bounded operator on a complex Banach space, with spectrum $\sigma(T)$, and h is a function that is analytic in a neighbourhood of $\sigma(T)$, then $h(T)$ is the bounded operator defined by

$$h(T) := \frac{1}{2\pi \mathrm{i}} \oint_\Gamma h(z)(z - T)^{-1}\,\mathrm{d}z, \tag{3.35}$$

integrating over a union Γ of rectifiable closed curves that encircle each component of $\sigma(T)$ once in the positive direction, such that furthermore h is analytic on Γ and in the interior of each of the curves. For properties of the map $h \mapsto h(T)$ see [5, Theorem VII.4.7]. In particular, note that if $h = h_1 h_2$, with h_1 and h_2 analytic in a neighbourhood of $\sigma(T)$, then

$$h(T) = h_1(T) h_2(T). \tag{3.36}$$

Furthermore, the resolvent $z \mapsto (z-T)^{-1}$ is analytic outside $\sigma(T)$ [5, Theorem VII.3.6], and thus $\|(z-T)^{-1}\|$ is bounded on Γ; hence (3.35) implies the existence of a constant C_Γ (depending also on T) such that

$$\|h(T)\| \le C_\Gamma \sup_{z\in\Gamma} |h(z)|. \tag{3.37}$$

Recall also that (3.35) extends the elementary definition of $h(T)$ for polynomials h. Hence,

$$B_{m,n} = b_{m,n}(\mathbf{R}), \tag{3.38}$$

where $b_{m,n}(z)$ is the polynomial

$$b_{m,n}(z) := \prod_{k=m}^{n-1} \bigl(1+\gamma_k(z-1)\bigr). \tag{3.39}$$

Moreover, for any complex c, the function e^{cz} is entire so e^{cT} can be defined by (3.35) as a bounded operator; this agrees with the definition using the usual power series expansion. In particular, if $t>0$, this defines $t^T = \mathrm{e}^{(\log t)T}$.

Lemma 3.3. *For each compact set $K \subset \mathbb{C}$, we have uniformly for $z \in K$ and $0 \le m \le n$, with $n \ge 1$,*

$$|b_{m,n}(z)| \le C\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+(m\vee 1)}\right)^{\operatorname{Re} z-1}. \tag{3.40}$$

Furthermore, there exists a family of analytic functions $h_{m,n}: \mathbb{C}\to\mathbb{C}$ defined by

$$b_{m,n}(z) = (1+h_{m,n}(z))\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+(m\vee 1)}\right)^{z-1} \tag{3.41}$$

$$= (1+h_{m,n}(z))\,\exp\bigl[(z-1)\bigl(\log(\mathfrak{m}_0(E)+n) - \log(\mathfrak{m}_0(E)+(m\vee 1))\bigr)\bigr] \tag{3.42}$$

such that uniformly for $z\in K$ and $0\le m\le n$,

$$|h_{m,n}(z)| \le \frac{C}{m\vee 1}. \tag{3.43}$$

Proof. The function $h_{m,n}$ defined by (3.41) is analytic, and hence it only remains to prove (3.43), since then (3.40) follows from (3.41).

Let $C_K := \sup_{z\in K}|z-1|$. We may in the sequel assume $n\ge m\ge 2C_K$, and in particular that $m\vee 1 = m$. The result for smaller m then follows from the result for $m = \lceil 2C_K\rceil$ because each factor in (3.39) is bounded by $1+C_K$ on K. (The case $n < \lceil 2C_K\rceil$ is trivial.)

For $k\ge m\ge 2C_K$ and $z\in K$, we have $\gamma_k \le 1/k \le 1/(2C_K)$, and thus $|\gamma_k(z-1)|\le 1/2$. Hence,

$$\bigl|\log\bigl(1+\gamma_k(z-1)\bigr) - \gamma_k(z-1)\bigr| \le \gamma_k^2|z-1|^2 \le \frac{C_K^2}{k^2}. \tag{3.44}$$

Consequently,

$$b_{m,n}(z) = \exp\Bigl(\sum_{k=m}^{n-1}\log\bigl(1+\gamma_k(z-1)\bigr)\Bigr) = \exp\Bigl(\sum_{k=m}^{n-1}\frac{z-1}{\mathfrak{m}_0(E)+1+k} + O(1/m)\Bigr)$$

$$= \exp\bigl[(z-1)\bigl(\log(\mathfrak{m}_0(E)+n) - \log(\mathfrak{m}_0(E)+m)\bigr) + O(1/m)\bigr], \tag{3.45}$$

where the implicit constant in $O(1/m)$ does not depend on $\mathfrak{m}_0$, and the result (3.43) follows. ∎

Remark 3.4. Alternatively, one can show (3.40) and (3.43) using the exact formula

$$b_{m,n}(z) = \prod_{k=m}^{n-1}\frac{\mathfrak{m}_0(E)+k+z}{\mathfrak{m}_0(E)+k+1} = \frac{\Gamma(n+\mathfrak{m}_0(E)+z)}{\Gamma(n+\mathfrak{m}_0(E)+1)}\frac{\Gamma(m+\mathfrak{m}_0(E)+1)}{\Gamma(m+\mathfrak{m}_0(E)+z)} \tag{3.46}$$

and Stirling's formula. ∎

Lemma 3.5. *Assume that the conditions of Theorem 2.10 hold. Then* $\Pi_1 f = (\nu f)1$*, for all* $f \in D$*. As a consequence,*

$$(\mathbf{I} - \Pi_1)D = \{f \in D : \Pi_1 f = 0\} = \{f \in D : \nu f = 0\}. \tag{3.47}$$

In particular, if Δ *is a clopen subset of* $\sigma(\mathbf{R}_D) \setminus \{1\}$*, and* $f \in \Pi_\Delta D$*, then* $\nu f = 0$*.*

Proof. Define $\Pi f := (\nu f)1$. Then Π is a bounded operator in $D \subseteq B(W)$ because $\nu W < \infty$ by (N) and $1 \in D$. Furthermore, Π is a projection in D (since $\nu 1 = 1$), and (2.19) and (2.20) imply that $\mathbf{R}\Pi = \Pi = \Pi\mathbf{R}$. Thus Π commutes with $\mathbf{R}$, and therefore with Π_1 (see [5, Proposition VII.4.9]). Furthermore, Π and Π_1 are both projections with rank 1, and the eigenfunction 1 belongs to both their ranges. Hence Π and Π_1 are both projections onto the subspace of constant functions. We thus get that, for any $f \in D$,

$$\Pi_1 f = \Pi\,\Pi_1 f = \Pi_1 \Pi f = \Pi f, \tag{3.48}$$

as stated. The equalities (3.47) follow. Finally, if $1 \notin \Delta$ and $f \in \Pi_\Delta D$, then $\Pi_1 f = \Pi_1\Pi_\Delta f = \Pi_{\{1\}\cap\Delta} f = \Pi_\emptyset f = 0$, see e.g. [7, Corollary VII.3.21], and thus $\nu f = 0$. □

3.2. **Proof of** (2.23) **and** (2.25) **of Theorem 2.10.** We prove first some lemmata.

Lemma 3.6. *Assume that the conditions of Theorem 2.10 hold.*

(i) *For every* $\delta \in (0, 1-\theta_D)$*, there exists a constant* C_δ *such that for every* $f \in D$ *with* $\nu f = 0$*,* $n \geq 1$ *and* $0 \leq m \leq n$*,*

$$\left\|B_{m,n} f\right\|_{B(W)} \leq C_\delta \left(\frac{\mathfrak{m}_0(E) + (m \vee 1)}{\mathfrak{m}_0(E) + n}\right)^\delta \|f\|_{B(W)}. \tag{3.49}$$

(ii) *If* $f \in B(W)$ *and* $(\mathbf{R} - \lambda\mathbf{I})^\kappa f = 0$ *for some* $\lambda \in \mathbb{C}$ *and* $\kappa \geq 1$*, then for* $n \geq 1$ *and* $0 \leq m \leq n$*,*

$$\left\|B_{m,n} f\right\|_{B(W)} \leq C_{\lambda,\kappa} \left(\frac{\mathfrak{m}_0(E) + (m \vee 1)}{\mathfrak{m}_0(E) + n}\right)^{1-\mathrm{Re}(\lambda)} \left[1 + \left(\log\left(\frac{\mathfrak{m}_0(E) + n}{\mathfrak{m}_0(E) + (m \vee 1)}\right)\right)^{\kappa-1}\right] \|f\|_{B(W)}, \tag{3.50}$$

for some constant $C_{\lambda,\kappa}$ *not depending on* f*.*

Proof. (i): Note that $f \in (\mathbf{I} - \Pi_1)D$ by Lemma 3.5. We let $\mathbf{R}'$ denote the restriction of $\mathbf{R}$ (or $\mathbf{R}_D$) to $(\mathbf{I} - \Pi_1)D$; $\mathbf{R}'$ is a bounded operator with spectrum $\sigma(\mathbf{R}') = \sigma(\mathbf{R}_D) \setminus \{1\}$ (see for instance [7, Theorem VII.3.20]).

Fix $\delta \in (0, 1-\theta_D)$. Then $\sup_{z\in\sigma(\mathbf{R}')} \mathrm{Re}(z) = \theta_D < 1-\delta$, and thus we can find a rectifiable curve Γ in $\mathbb{C}$ that encircles $\sigma(\mathbf{R}')$ such that $\sup_{z\in\Gamma} \mathrm{Re}(z) \leq 1-\delta$. Consequently, by (3.38) and (3.35),

$$B_{m,n} f = b_{m,n}(\mathbf{R}) f = b_{m,n}(\mathbf{R}') f = \frac{1}{2\pi \mathrm{i}} \oint_\Gamma b_{m,n}(z)(z - \mathbf{R}')^{-1} f \,\mathrm{d}z. \tag{3.51}$$

Furthermore, Lemma 3.3 implies that for $z \in \Gamma$,

$$|b_{m,n}(z)| \leq C \left(\frac{\mathfrak{m}_0(E) + n}{\mathfrak{m}_0(E) + (m \vee 1)}\right)^{\mathrm{Re}\, z - 1} \leq C \left(\frac{\mathfrak{m}_0(E) + n}{\mathfrak{m}_0(E) + (m \vee 1)}\right)^{-\delta} = C \left(\frac{\mathfrak{m}_0(E) + (m \vee 1)}{\mathfrak{m}_0(E) + n}\right)^{\delta}. \tag{3.52}$$

The result (3.49) follows from (3.51) and (3.52), see (3.37).

(ii): We use the factorization (3.41) and (3.36). Thus,

$$B_{m,n} f = b_{m,n}(\mathbf{R}) f = \big(\mathbf{I} + h_{m,n}(\mathbf{R})\big) \left(\frac{\mathfrak{m}_0(E) + n}{\mathfrak{m}_0(E) + (m \vee 1)}\right)^{\mathbf{R}-\mathbf{I}} f. \tag{3.53}$$

Furthermore, by (3.43), the functions $h_{m,n}$, for $n \geq 1$ and $0 \leq m \leq n$, are uniformly bounded on any fixed compact subset of $\mathbb{C}$, and thus (3.37) implies that the operators $h_{m,n}(\mathbf{R})$ are uniformly bounded on $B(W)$ by a constant that does not depend on $\mathfrak{m}_0$.

Moreover, for all functions $f \in D$ such that $(\mathbf{R}-\lambda\mathbf{I})^\kappa f = 0$, for $m \geq 1$,

$$\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+m}\right)^{\mathbf{R}-\mathbf{I}} f = \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+m}\right)^{\lambda-1} \exp\left(\log\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+m}\right)(\mathbf{R}-\lambda\mathbf{I})\right) f$$
$$= \left(\frac{\mathfrak{m}_0(E)+m}{\mathfrak{m}_0(E)+n}\right)^{1-\lambda} \sum_{k=0}^{\kappa-1} \log^k\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+m}\right)\frac{(\mathbf{R}-\lambda\mathbf{I})^k f}{k!}. \tag{3.54}$$

The result follows by (3.53) and (3.54) since the operators $(\mathbf{R}-\lambda\mathbf{I})^k$ are bounded. □

Recall from Section 1.4 that we use $\mathfrak{C}$ for constants that may depend on $\mathfrak{m}_0$.

Lemma 3.7. *Assume that the conditions of Theorem 2.10 hold.*

(i) *For every $\delta \in (0, 1-\theta_D)$, there exists a constant $\mathfrak{C}_\delta < \infty$ such that for every $f \in D$,*

$$|\zeta_{n,0}| \leq \mathfrak{C}_\delta ({}^1\!/\!{}_n)^\delta \, \|f\|_{B(W)}, \qquad n \geq 1, \tag{3.55}$$

$$\left(\mathbb{E}_{i-1} |\zeta_{n,i}|^q\right)^{1/q} \leq \frac{\mathfrak{C}_\delta}{i} ({}^i\!/\!{}_n)^\delta \, \|f\|_{B(W)} \left(\widetilde{\mathfrak{m}}_{i-1}(V)\right)^{1/q}, \qquad n \geq i \geq 1. \tag{3.56}$$

(ii) *If $\lambda \in \mathbb{C}$ and $\kappa \geq 1$, there exists a constant $\mathfrak{C}_{\lambda,\kappa}$ such that if $f \in B(W)$ and $(\mathbf{R}-\lambda\mathbf{I})^\kappa f = 0$, then*

$$|\zeta_{n,0}| \leq \mathfrak{C}_{\lambda,\kappa} ({}^1\!/\!{}_n)^{1-\mathrm{Re}(\lambda)} (\log n)^{\kappa-1} \, \|f\|_{B(W)}, \qquad n \geq 2, \tag{3.57}$$

$$\left(\mathbb{E}_{i-1} |\zeta_{n,i}|^q\right)^{1/q} \leq \frac{\mathfrak{C}_{\lambda,\kappa}}{i} ({}^i\!/\!{}_n)^{1-\mathrm{Re}(\lambda)} \left[1 + (\log({}^n\!/\!{}_i))^{\kappa-1}\right] \|f\|_{B(W)}\, \widetilde{\mathfrak{m}}_{i-1}(V)^{1/q}, \qquad n \geq i \geq 1. \tag{3.58}$$

Proof. By homogeneity, we may without loss of generality assume $\|f\|_{B(W)} = 1$. Furthermore, by (3.34) we may replace f by $f - \nu f$; hence we may also assume $\nu f = 0$.

(i): Fix $\delta \in (0, 1-\theta_D)$. According to Lemma 3.6, we have for all $n \geq 1$ and $0 \leq i \leq n$,

$$|B_{i,n} f| \leq \mathfrak{C}'_\delta \left(\frac{i \vee 1}{n}\right)^\delta W. \tag{3.59}$$

First, taking $i = 0$, we obtain (3.55) from (3.30) and (3.59), since $\mathfrak{v}_0 \in \mathcal{M}(W)$ by Lemma 3.2.

For $n \geq i \geq 1$, we have by (3.31), Lemma 3.2 (Equation (3.13)) and (3.59),

$$\mathbb{E}_{i-1} |\zeta_{n,i}|^q = \gamma_{i-1}^q \, \mathbb{E}_{i-1} |\Delta M_i B_{i,n} f|^q \leq \frac{C}{i^q} \|B_{i,n} f\|_{B(W)}^q \, \widetilde{\mathfrak{m}}_{i-1}(V) \leq \frac{\mathfrak{C}''_\delta}{i^q} \left(\frac{i}{n}\right)^{\delta q} \widetilde{\mathfrak{m}}_{i-1}(V). \tag{3.60}$$

This concludes the proof of (3.56).

(ii): The same arguments but using (3.50) instead of (3.49) lead to (3.57) and (3.58). □

For technical reasons, we have stated Theorem 2.10 for an invariant subspace D containing the constant functions; these are eigenfunctions with eigenvalue 1 by (2.19), and thus $1 \in \sigma(\mathbf{R}_D)$. It will now be convenient to consider also invariant subspaces not containing constants; we then use the generic notation D' to help the reader distinguish the assumptions.

Lemma 3.8. *Suppose that D' is an $\mathbf{R}$-invariant subspace of $B(W)$ and that $\theta' \in \mathbb{R}$ is such that* $\sup \mathrm{Re}\, \sigma(\mathbf{R}_{D'}) < \theta'$. *Then,*

$$\mathbb{E}\, |\mathfrak{v}_n f|^2 \leq C\, \widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+n}\right)^{1\wedge 2(1-\theta')} \|f\|_{B(W)}^2, \qquad f \in D',\ n \geq 1. \tag{3.61}$$

Proof. The terms in (3.26) are orthogonal, and thus, using (3.11) and (3.12) in Lemma 3.2,

$$\mathbb{E}\, |\mathfrak{v}_n f|^2 = \mathbb{E}\left[|\mathfrak{v}_0 B_{0,n} f|^2\right] + \sum_{i=1}^n \gamma_{i-1}^2 \, \mathbb{E}\left[\left|\Delta M_i B_{i,n} f\right|^2\right]$$

$$\le C\widetilde{\mathfrak{m}}_0 V\|B_{0,n}f\|^2_{B(W)} + C\sum_{i=1}^{n}(\mathfrak{m}_0(E)+i)^{-2}\,\widetilde{\mathfrak{m}}_0 V\,\|B_{i,n}f\|^2_{B(W)}. \tag{3.62}$$

We apply (3.37) to $B_{m,n} = b_{m,n}(\mathbf{R})$ as an operator on D'. By the assumption, we may choose a curve Γ encircling $\sigma(\mathbf{R}_{D'})$ such that $\operatorname{Re} z \le \theta'$ for $z \in \Gamma$, and then (3.37) and (3.40) yield, uniformly for $0 \le m \le n$,

$$\|B_{m,n}\|_{D'} \le C\Big(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+(m\vee 1)}\Big)^{\theta'-1}. \tag{3.63}$$

By homogeneity, we may assume $\|f\|_{B(W)} = 1$, and then (3.62) and (3.63) yield

$$\begin{aligned}\frac{\mathbb{E}\,|\mathfrak{v}_n f|^2}{\widetilde{\mathfrak{m}}_0 V} &\le C\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{2(\theta'-1)} + C\sum_{i=1}^{n}(\mathfrak{m}_0(E)+i)^{-2}\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+i}\right)^{2(\theta'-1)}\\ &\le C\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{2\theta'-2} + C\,(\mathfrak{m}_0(E)+n)^{2\theta'-2}\sum_{i=1}^{n}(\mathfrak{m}_0(E)+i)^{-2\theta'}\\ &\le C\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{2\theta'-2} + C\begin{cases}(\mathfrak{m}_0(E)+n)^{-1} & \text{if } \theta' < 1/2\\ (\mathfrak{m}_0(E)+n)^{-1}\log(\mathfrak{m}_0(E)+n) & \text{if } \theta' = 1/2\\ (\mathfrak{m}_0(E)+n)^{2\theta'-2} & \text{if } \theta' > 1/2.\end{cases}\end{aligned} \tag{3.64}$$

This yields (3.61) when $\theta' \ne 1/2$. If $\theta' = 1/2$, then one can replace θ' by some new $\theta' < 1/2$; then (3.64) yields (3.61) in this case too. ☐

The estimates (2.23) and (2.25) of Theorem 2.10 directly follow from the following result. Recall that $\mathfrak{v}_n := \widetilde{\mathfrak{m}}_n - \nu$.

Lemma 3.9. *Assume that the conditions of Theorem 2.10 hold.*

(i) *For every $\delta \in (0, 1-\theta_D)$, there exists a constant $C_\delta < \infty$ such that, for any $f \in D$ and any $n \ge 1$,*

$$\mathbb{E}\,|\mathfrak{v}_n f|^2 \le C_\delta\,\widetilde{\mathfrak{m}}_0 V\left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+n}\right)^{2\delta\wedge 1}\|f\|^2_{B(W)}. \tag{3.65}$$

(ii) *If, furthermore, $\mathbf{R}$ is slqc on $B(W)$, then, for every $\delta \in (0, 1-\theta)$, there exists a constant $C_\delta < \infty$ such that, for all $f \in B(W^2)$,*

$$\mathbb{E}\,|\mathfrak{v}_n f| \le C_\delta\,\widetilde{\mathfrak{m}}_0 V\left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+n}\right)^{(2\delta\wedge 1)\frac{q/2-1}{q-1}}\|f\|_{B(W^2)}. \tag{3.66}$$

Proof. (i): This is essentially equivalent to Lemma 3.8. Recalling Lemma 3.5, we define

$$D' := (1-\Pi_1)D = \{f \in D : \nu f = 0\}. \tag{3.67}$$

Then, as in the proof of Lemma 3.6, D' is an invariant subspace of $B(W)$ and $\mathbf{R}_{D'}$ has spectrum $\sigma(\mathbf{R}_{D'}) = \sigma(\mathbf{R}_D) \setminus \{1\}$, and thus $\sup \operatorname{Re}\sigma(\mathbf{R}_{D'}) = \theta_D$. We define $\theta' := 1-\delta$, and note that the assumption implies $\theta' > \theta_D$. Hence, (3.61) applies and yields (3.65) for $f \in D'$. Finally, for a general $f \in D$, we apply instead (3.61) to $f - (\nu f)1 = (1-\Pi_1)f \in D'$ (recalling Lemma 3.5), noting that $\mathfrak{v}_n 1 = 0$ by (3.32).

(ii): We now assume that the operator $\mathbf{R}$ is slqc, so we may take $D = B(W)$ in (i). For an arbitrary $f \in B(W^2)$, we will use truncations: For all $K \ge 1$,

$$\mathbb{E}\,|\mathfrak{v}_n f| \le \big|\mathbb{E}[\mathfrak{v}_n(f\mathbf{1}_{W^2\le K})]\big| + \mathbb{E}[\widetilde{\mathfrak{m}}_n|f\mathbf{1}_{W^2>K}|] + \nu|f\mathbf{1}_{W^2>K}|. \tag{3.68}$$

First, since $|f(x)|\mathbf{1}_{W(x)^2\le K} \le \|f\|_{B(W^2)}W(x)\sqrt{K}$, we deduce that $\|f\mathbf{1}_{W^2\le K}\|_{B(W)} \le \|f\|_{B(W^2)}\sqrt{K}$. Therefore, by (3.65) applied to $f\mathbf{1}_{W^2\le K} \in B(W)$, we get, for any fixed

$\delta \in (0, 1-\theta)$ (the constants C below do not depend on f or K),

$$\mathbb{E}\big|\mathfrak{v}_n(f\mathbf{1}_{W^2\le K})\big| \le \mathbb{E}[|\mathfrak{v}_n(f\mathbf{1}_{W^2\le K})|^2]^{1/2} \le C\,\widetilde{\mathfrak{m}}_0 V\left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+n}\right)^{\frac{2\delta\wedge 1}{2}} \|f\|_{B(W^2)}\sqrt{K}. \tag{3.69}$$

On the other hand, if $W^2 > K$ then $VW^{-2} = W^{q-2} > K^{q/2-1}$ and thus

$$\begin{aligned}\mathbb{E}[\widetilde{\mathfrak{m}}_n|f\mathbf{1}_{W^2>K}|] &\le \mathbb{E}[\widetilde{\mathfrak{m}}_n(W^2\mathbf{1}_{W^2>K})]\,\|f\|_{B(W^2)} \le K^{1-q/2}\,\mathbb{E}[\widetilde{\mathfrak{m}}_n V]\,\|f\|_{B(W^2)}\\ &\le CK^{1-q/2}\,\|f\|_{B(W^2)}\,\widetilde{\mathfrak{m}}_0 V,\end{aligned} \tag{3.70}$$

where we used (3.10). The same computation also holds for $\nu|f\mathbf{1}_{W^2>K}|$, since, by assumption (N), $\nu V < \infty$.

Finally, choosing $K = ((\mathfrak{m}_0(E)+1)/(\mathfrak{m}_0(E)+n))^{-(2\delta\wedge 1)/(q-1)}$ and using (3.69) and (3.70), we deduce (3.66). □

3.3. **Proof of** (2.24)**.** We improve the estimate in Lemma 3.9(i) to an estimate for a maximum over a restricted range.

Lemma 3.10. *Suppose that D' is an $\mathbf{R}$-invariant subspace of $B(W)$ and that $\theta' > 1/2$ is such that $\sup\operatorname{Re}\sigma(\mathbf{R}_{D'}) < \theta'$. Moreover, let $0 < \tau < 1$. Then there exists a constant $C = C(\tau)$ such that*

$$\mathbb{E}\sup_{N-N^\tau\le n\le N}|\mathfrak{v}_n f|^2 \le C\widetilde{\mathfrak{m}}_0 V\left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+N}\right)^{2(1-\theta')+\tau} N^\tau\,\|f\|^2_{B(W)}, \qquad f\in D',\ N\ge 1. \tag{3.71}$$

Proof. If $0 \le i \le n \le N$, then the definition (3.3) implies

$$B_{i,n}B_{n,N} = B_{i,N}. \tag{3.72}$$

Let $f \in D'$. For $0 \le n \le N$, we apply (3.26) to $B_{n,N}f$ and obtain, using (3.72),

$$\mathfrak{v}_n B_{n,N} f = \mathfrak{v}_0 B_{0,N} f + \sum_{i=1}^{n} \gamma_{i-1}\Delta M_i B_{i,N} f. \tag{3.73}$$

Since $\mathbb{E}_{i-1}\Delta M_i = 0$, (3.73) shows that $\big(\mathfrak{v}_n B_{n,N} f\big)_{0\le n\le N}$ is a martingale for each fixed N. (Note that the terms on the right-hand side do not depend on n.) Consequently, Doob's maximal inequality yields, together with $B_{N,N} = \mathbf{I}$ and Lemma 3.8,

$$\begin{aligned}\mathbb{E}\big|\sup_{n\le N}\mathfrak{v}_n B_{n,N} f\big|^2 &\le 4\,\mathbb{E}\big|\mathfrak{v}_N B_{N,N} f\big|^2 = 4\,\mathbb{E}\big|\mathfrak{v}_N f\big|^2\\ &\le C\,\widetilde{\mathfrak{m}}_0 V\left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+N}\right)^{2(1-\theta')}\|f\|^2_{B(W)}.\end{aligned} \tag{3.74}$$

Let K be a compact neighbourhood of $\sigma(\mathbf{R}_{D'})$ such that $\sup\operatorname{Re}K < \theta'$. Let $n = N - m$, where $0 \le m \le N^\tau$, and suppose that N is so large that $N^\tau \le N/2$. Also, let $L := \lfloor 1/(1-\tau)\rfloor$. Then, $m \le N/2$ and thus

$$\begin{aligned}\log\left(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\right) &= \left|\log\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+N}\right)\right| = \left|\log\left(\frac{\mathfrak{m}_0(E)+N-m}{\mathfrak{m}_0(E)+N}\right)\right|\\ &= \left|\log\Big(1-\frac{m}{\mathfrak{m}_0(E)+N}\Big)\right| \le \frac{2m}{\mathfrak{m}_0(E)+N} \le 2(\mathfrak{m}_0(E)+N)^{\tau-1}.\end{aligned} \tag{3.75}$$

Since $n = N - m \ge N/2$, (3.43) yields, for all $z \in K$,

$$|h_{n,N}(z)| \le \frac{C}{n} \le \frac{C}{N}. \tag{3.76}$$

Assume in the sequel that N and $n \le N$ are as above, and also that N is so large that (3.76) implies $|h_{n,N}(z)| \le 1/2$ when $z \in K$. Then, (3.42)–(3.43) imply that, uniformly for $z \in K$ and all such n and N,

$$\begin{aligned}
\frac{1}{b_{n,N}(z)} &= \Big(1 + O\Big(\frac{1}{N}\Big)\Big) \exp\left(-(z-1)\log\left(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\right)\right) \\
&= \exp\left(-(z-1)\log\left(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\right)\right) + O\Big(\frac{1}{N}\Big) \\
&= \sum_{\ell=0}^{L} \frac{1}{\ell!}\log^{\ell}\Big(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\Big)(1-z)^{\ell} + O\Big(\log^{L+1}\Big(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\Big)\Big) + O\Big(\frac{1}{N}\Big). \\
&= \sum_{\ell=0}^{L} \frac{1}{\ell!}\log^{\ell}\Big(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\Big)(1-z)^{\ell} + O\Big(\frac{1}{N}\Big), \qquad (3.77)
\end{aligned}$$

where the last equality uses (3.75) and $(L+1)(1-\tau) > 1$. In particular, $b_{n,N}^{-1}(z)$ is finite for $z \in K$, so $b_{n,N}^{-1}(z)$ is analytic in a neighbourhood of $\sigma(\mathbf{R}_{D'})$; hence $B_{n,N}$ is invertible on D', with $B_{n,N}^{-1} = b_{n,N}^{-1}(\mathbf{R})$. Define the operator on D'

$$V_{n,N} := B_{n,N}^{-1} - \sum_{\ell=0}^{L} \frac{1}{\ell!}\log^{\ell}\Big(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\Big)(\mathbf{I}-\mathbf{R})^{\ell}. \qquad (3.78)$$

It follows from (3.77) and (3.37) that

$$\|V_{n,N}\|_{D'} = O\left(\frac{1}{N}\right). \qquad (3.79)$$

Moreover, (3.78) yields, with $f_\ell := (\mathbf{I}-\mathbf{R})^{\ell} f/\ell!$ and $g_{n,N} := V_{n,N} f$,

$$B_{n,N}^{-1} f = \sum_{\ell=0}^{L} \log^{\ell}\Big(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\Big) f_\ell + g_{n,N}, \qquad (3.80)$$

and thus

$$\mathfrak{v}_n f = \mathfrak{v}_n B_{n,N} B_{n,N}^{-1} f = \sum_{\ell=0}^{L} \log^{\ell}\Big(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\Big)\mathfrak{v}_n B_{n,N} f_\ell + \mathfrak{v}_n B_{n,N} g_{n,N}. \qquad (3.81)$$

According to (3.75), we have $0 \le \log\big(\frac{\mathfrak{m}_0(E)+N}{\mathfrak{m}_0(E)+n}\big) \le 2N^{\tau-1} \le 2$ and hence, with constants C depending on τ in the remainder of the proof,

$$\begin{aligned}
\sup_{N-N^\tau \le n \le N} |\mathfrak{v}_n f|^2 &\le C \sum_{\ell=0}^{L} \sup_{N-N^\tau \le n \le N} \big|\mathfrak{v}_n B_{n,N} f_\ell\big|^2 + C \sup_{N-N^\tau \le n \le N} \big|\mathfrak{v}_n B_{n,N} g_{n,N}\big|^2 \\
&\le C \sum_{\ell=0}^{L} \sup_{n \le N} \big|\mathfrak{v}_n B_{n,N} f_\ell\big|^2 + C \sum_{N-N^\tau \le n \le N} \big|\mathfrak{v}_n B_{n,N} g_{n,N}\big|^2. \qquad (3.82)
\end{aligned}$$

Furthermore, for n and N as above, (3.79) holds; hence

$$\|g_{n,N}\|_{B(W)} = \|V_{n,N} f\|_{B(W)} \le C N^{-1} \|f\|_{B(W)} \le C \frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+N} \|f\|_{B(W)}. \qquad (3.83)$$

Taking the expectation in (3.82) and using (3.74) and (3.83) yields

$$\begin{aligned}
&\mathbb{E} \sup_{N-N^\tau \le n \le N} |\mathfrak{v}_n f|^2 \\
&\le C \sum_{\ell=0}^{L} \widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+N}\right)^{2(1-\theta')} \|f_\ell\|_{B(W)}^2 + C \sum_{N-N^\tau \le n \le N} \widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+N}\right)^{2(1-\theta')} \|g_{n,N}\|_{B(W)}^2
\end{aligned}$$

$$\leq C\widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+N}\right)^{2(1-\theta')} \|f\|^2_{B(W)} + C\widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+N}\right)^{2(1-\theta')+2} N^\tau \|f\|^2_{B(W)}. \tag{3.84}$$

This shows (3.71) when N is large enough since $\tau < 1 < 2$. The remaining cases are trivial, since (3.71) for any fixed N follows from Lemma 3.8. □

We are now ready to prove (2.24) and thus conclude the proof of Theorem 2.10:

Lemma 3.11. (i) *Suppose that D' is an $\mathbf{R}$-invariant subspace of $B(W)$ and that $\theta' > 1/2$ is such that $\sup \operatorname{Re} \sigma(\mathbf{R}_{D'}) < \theta'$. Then, for every $f \in D'$, a.s. and in L^2 as $n \to \infty$,*

$$n^{1-\theta'} \mathfrak{v}_n f \to 0. \tag{3.85}$$

(ii) *Assume that the conditions of Theorem 2.10 holds, and let $\delta \in (0, (1-\theta_D) \wedge 1/2)$. Then $n^\delta \mathfrak{v}_n f \to 0$ a.s. and in L^2 as $n \to \infty$, for every $f \in D$.*

Proof. (i): Let $\theta := \sup \operatorname{Re} \sigma(\mathbf{R}_{D'})$ and choose $\theta'' \in (\theta \vee 1/2, \theta')$. Then Lemma 3.8 applied with θ'' yields, for any $f \in D'$,

$$\begin{aligned}\mathbb{E}\left|\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{1-\theta'} \mathfrak{v}_n f\right|^2 &\leq C\widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{2(1-\theta')+2(\theta''-1)} \|f\|^2_{B(W)} \\ &= C\widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{2(\theta''-\theta')} \|f\|^2_{B(W)} = o(1). \end{aligned} \tag{3.86}$$

This implies the convergence (3.85) in L^2.

To show the convergence a.s., choose $\tau \in (0,1)$ with $\tau > 1+\theta''-\theta'$. Define an increasing sequence (n_k) by $n_0 := 1$ and $n_{k+1} := n_k + \lfloor n_k^\tau \rfloor$. Then Lemma 3.10 applied with θ'' yields, for every $k \geq 1$, (here $\mathfrak{C}$ are constants that may depend on $\mathfrak{m}_0$ and f)

$$\begin{aligned}\mathbb{E} \sup_{n_{k-1}<n\leq n_k} |n^{1-\theta'} \mathfrak{v}_n f|^2 &\leq C n_k^{2-2\theta'} \mathbb{E} \sup_{n_k - n_k^\tau \leq n \leq n_k} |\mathfrak{v}_n f|^2 \\ &\leq \mathfrak{C} n_k^{2-2\theta'+2(\theta''-1)} = \mathfrak{C} n_k^{2\theta''-2\theta'} \\ &\leq \mathfrak{C} \sum_{n=n_{k-1}+1}^{n_k} n^{2\theta''-2\theta'-\tau}. \end{aligned} \tag{3.87}$$

The exponent in the final sum is

$$2\theta''-2\theta'-\tau < 2\theta''-2\theta'-(1+\theta''-\theta') = \theta''-\theta'-1 < -1. \tag{3.88}$$

Consequently,

$$\mathbb{E}\sum_{k=1}^\infty \sup_{n_{k-1}<n\leq n_k} |n^{1-\theta'} \mathfrak{v}_n f|^2 \leq \mathfrak{C} \sum_{k=1}^\infty \sum_{n=n_{k-1}+1}^{n_k} n^{2\theta''-2\theta'-\tau} = \mathfrak{C} \sum_{n=2}^\infty n^{2\theta''-2\theta'-\tau} < \infty. \tag{3.89}$$

Hence, a.s.,

$$\sum_{k=1}^\infty \sup_{n_{k-1}<n\leq n_k} |n^{1-\theta'} \mathfrak{v}_n f|^2 < \infty, \tag{3.90}$$

which implies that $\sup_{n_{k-1}\leq n\leq n_k} |n^{1-\theta'} \mathfrak{v}_n f|^2 \to 0$ as $k \to \infty$, and thus $n^{1-\theta'} \mathfrak{v}_n f \to 0$ as $n \to \infty$.

(ii): Let, as in the proof of Lemma 3.9, $D' := (1-\Pi_1)D$ and apply (i) to D' and $f - (\nu f)1 \in D'$ with $\theta' := 1-\delta > \theta_D = \sup \operatorname{Re} \sigma(\mathbf{R}_{D'})$. □

Remark 3.12. We observe that, if $\mathfrak{m}_0$ is random, then (3.86) holds conditioned on $\mathfrak{m}_0$. Hence, if $\mathbb{E}\big[(\mathfrak{m}_0(E)+1)^{2(1-\theta')}\widetilde{\mathfrak{m}}_0 V\big] < \infty$, then, using dominated convergence,

$$\mathbb{E}\big|(\mathfrak{m}_0(E)+n)^{1-\theta'} \mathfrak{v}_n f\big|^2 \leq C\,\mathbb{E}\Big[(\mathfrak{m}_0(E)+1)^{2(1-\theta')}\widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{2(\theta''-\theta')}\Big] \|f\|^2_{B(W)}$$

$$\to 0. \tag{3.91}$$

Hence, the convergence (3.85) still holds in L^2. □

4. Proof of Theorem 2.13

4.1. **Proofs of Theorem 2.13(1) and (2).** Recall that $\widetilde{\mathfrak{m}}_n - \nu = \mathfrak{v}_n$ and that $\mathfrak{v}_n 1 = 0$ by (3.32), which implies that $\mathfrak{v}_n f$ is not affected if we subtract a constant from f. It is also obvious that subtracting a constant from f does not affect $\chi(f)$ and $\sigma^2(f)$ in (2.28)–(2.29) and (2.34)–(2.35). Hence, replacing f by $f - \nu f$, we may in the proof assume that $\nu f = 0$. For convenience, we also assume $\|f\|_{B(W)} \le 1$, as we may by homogeneity.

We will prove (1) and (2) in parallel, since most the arguments are the same for both cases. Our proof relies on a central theorem for martingales given by Hall & Heyde [13, Corollary 3.1] (see [13] for other versions and references). This theorem in [13] is stated there for real-valued variables, but it extends immediately to vector-valued variables (in a finite-dimensional space) by the Cramér–Wold device [12, Theorem 5.10.5]; in particular, the theorem holds for complex-valued variables by considering the real and imaginary parts, and can then be stated as follows. (In general, χ and σ^2 may be random, but we are only interested in the special case when they are constant.)

Theorem 4.1 ([13, Corollary 3.1]). *Let $(\hat{\zeta}_{n,i}, n \ge 0, 1 \le i \le n)$ be a complex-valued martingale difference array. If there exist $\chi \in \mathbb{C}$ and $\sigma^2 \ge 0$ such that, in probability when $n \to \infty$,*

(a) $\sum_{i=1}^n \mathbb{E}_{i-1}[|\hat{\zeta}_{n,i}|^2 \mathbf{1}_{|\hat{\zeta}_{n,i}| \ge \varepsilon}] \to 0$ *for all $\varepsilon > 0$, and*
(b) $\sum_{i=1}^n \mathbb{E}_{i-1}[\hat{\zeta}_{n,i}^2] \to \chi$*, and*
(c) $\sum_{i=1}^n \mathbb{E}_{i-1}[|\hat{\zeta}_{n,i}|^2] \to \sigma^2$,

then, in distribution when $n \to \infty$, $\sum_{i=1}^n \hat{\zeta}_{n,i} \Rightarrow \Lambda_1 + i\Lambda_2$ where the random vector (Λ_1, Λ_2) has a centered Gaussian distribution with covariance matrix

$$\frac{1}{2}\begin{pmatrix} \sigma^2 + \mathrm{Re}(\chi) & \mathrm{Im}(\chi) \\ \mathrm{Im}(\chi) & \sigma^2 - \mathrm{Re}(\chi) \end{pmatrix} \tag{4.1}$$

In (1), we assume that $\mathbf{R}_D$ is a small operator. In this case, recall from Lemma 2.9 that $\theta_D < 1/2$. We may thus choose $\delta \in (1/2, 1)$ such that $\delta < 1 - \theta_D$; we fix such a δ for the rest of the proof.

In (2), we assume that $\mathbf{R}_D$ and $\mathbf{R}$ are slqc operators and that the spectrum of $\mathbf{R}_D$ is given by

$$\sigma(\mathbf{R}_D) = \{1, \lambda_1, \dots, \lambda_p\} \cup \Delta, \tag{4.2}$$

where $\mathrm{Re}(\lambda_1) = \dots = \mathrm{Re}(\lambda_p) = 1/2$, and $\sup \mathrm{Re}(\Delta) < 1/2$. Thus $\theta_D = 1/2$. Let $\Delta' := \Delta \cup \{1\}$. Then

$$f = \Big(\Pi_{\Delta'} + \sum_{j=1}^p \Pi_{\lambda_j}\Big) f = \Pi_{\Delta'} f + \sum_{j=1}^p \Pi_{\lambda_j} f. \tag{4.3}$$

Furthermore, $\mathbf{R}$ is a small operator in $D' := \Pi_{\Delta'} D$. Hence, according to (2.23) of Theorem 2.10 applied to D' with $\delta = 1/2$,

$$\mathbb{E}\Big|\frac{\sqrt{n}}{(\log n)^{\kappa - 1/2}} \mathfrak{v}_n(\Pi_{\Delta'} f)\Big|^2 \le \mathfrak{C}(\log n)^{1-2\kappa} \xrightarrow[n\to+\infty]{} 0, \tag{4.4}$$

and hence it is sufficient to prove (2.33) for $f - \Pi_{\Delta'} f$ instead of f. In other words, in (2) we may assume that

$$f = f - \Pi_{\Delta'} f = \sum_{j=1}^p \Pi_{\lambda_j} f. \tag{4.5}$$

Note that $\|\Pi_{\lambda_j} f\|_{B(W)} \le C$, since each Π_{λ_j} is a bounded operator.

Returning to treating (1) and (2) together, we use (3.29), which we now write as

$$a_n \mathfrak{v}_n f = \sum_{i=0}^{n} a_n \zeta_{n,i} = \sum_{i=0}^{n} \hat{\zeta}_{n,i} \tag{4.6}$$

where

$$a_n := \begin{cases} n^{1/2} & \text{under the conditions of (1),} \\ \frac{n^{1/2}}{(\log n)^{\kappa-1/2}} & \text{under the conditions of (2),} \end{cases} \tag{4.7}$$

and $\hat{\zeta}_{n,i} := a_n \zeta_{n,i}$. For $1 \le i \le n$, we also set, for (2) considering in the sequel only $n \ge 2$,

$$d_{i,n} := \begin{cases} i^{2\delta-2}\, n^{1-2\delta} & \text{under the conditions of (1),} \\ i^{-1}(\log n)^{-1} & \text{under the conditions of (2).} \end{cases} \tag{4.8}$$

By Lemma 3.7, we have, using part (i) for (1) and part (ii) together with the decomposition (4.5) for (2), recalling (2.32),

$$|\hat{\zeta}_{n,0}| \le \mathfrak{C}\, d_{1,n}^{1/2} \xrightarrow[n\to+\infty]{} 0, \tag{4.9}$$

meaning that the $\hat{\zeta}_{n,0}$ may be ignored in (4.6).

We check that the $\hat{\zeta}_{n,i}$ satisfy conditions (a), (b) and (c) of Theorem 4.1:

Condition (a). We want to show the conditional Lindeberg condition

$$\sum_{i=1}^{n} \mathbb{E}_{i-1}\big[|\hat{\zeta}_{n,i}|^2 \mathbf{1}_{|\hat{\zeta}_{n,i}|\ge\varepsilon}\big] \xrightarrow[n\to+\infty]{\mathrm{p}} 0, \qquad \text{for every } \varepsilon > 0. \tag{4.10}$$

From Lemma 3.7 and (4.7),

$$\mathbb{E}_{i-1}\, |\hat{\zeta}_{n,i}|^q \le \mathfrak{C}\, d_{i,n}^{q/2} \widetilde{\mathfrak{m}}_{i-1}(V), \tag{4.11}$$

where $d_{i,n}$ is defined in (4.8). By (3.10) in Lemma 3.2, this implies

$$\mathbb{E}\left[|\hat{\zeta}_{n,i}|^2 \mathbf{1}_{|\hat{\zeta}_{n,i}|\ge\varepsilon}\right] \le \varepsilon^{2-q}\, \mathbb{E}\big|\hat{\zeta}_{n,i}\big|^q \le \mathfrak{C}_\varepsilon d_{i,n}^{q/2}, \tag{4.12}$$

for some constant $\mathfrak{C}_\varepsilon$ which may depend on $\mathfrak{m}_0$ and ε. We deduce that

$$\begin{aligned} \mathbb{E} \sum_{i=1}^{n} \mathbb{E}_{i-1}\big[|\hat{\zeta}_{n,i}|^2 \mathbf{1}_{|\hat{\zeta}_{n,i}|\ge\varepsilon}\big] &= \sum_{i=1}^{n} \mathbb{E}\left[|\hat{\zeta}_{n,i}|^2 \mathbf{1}_{|\hat{\zeta}_{n,i}|\ge\varepsilon}\right] \le \mathfrak{C}_\varepsilon \sum_{i=1}^{n} d_{i,n}^{q/2} \\ &\le \mathfrak{C}_\varepsilon \big(\max_{i\le n} d_{i,n}\big)^{q/2-1} \sum_{i=1}^{n} d_{i,n} \to 0, \end{aligned} \tag{4.13}$$

as $n \to \infty$, since (4.8) implies $\max_{i\le n} d_{i,n} = d_{1,n} \to 0$ and $\sum_{i=1}^{n} d_{i,n} \le C$. Hence, (4.10) holds, which is Condition (a) of Theorem 4.1.

Condition (b). First note that for $i \ge 1$, using the fact that $\hat{\zeta}_{n,i} = a_n \zeta_{n,i}$ and the definition of $\zeta_{n,i}$ in (3.31), and setting $f_{i,n} := B_{i,n} f$, we obtain

$$\begin{aligned} \mathbb{E}_{i-1}[\hat{\zeta}_{n,i}^2] &= a_n^2 \gamma_{i-1}^2\, \mathbb{E}_{i-1}\left[(\Delta M_i f_{i,n})^2\right] = a_n^2 \gamma_{i-1}^2\, \mathbb{E}_{i-1}\big[\big(R_{Y_i}^{(i)} f_{i,n} - \mathbb{E}_{i-1} R_{Y_i}^{(i)} f_{i,n}\big)^2\big] \\ &= a_n^2 \gamma_{i-1}^2 \big(\mathbb{E}_{i-1}\left[\big(R_{Y_i}^{(i)} f_{i,n}\big)^2\right] - (\mathbb{E}_{i-1} R_{Y_i}^{(i)} f_{i,n})^2\big) \\ &= a_n^2 \gamma_{i-1}^2 \big(\mathbb{E}_{i-1} \mathbf{B}_{Y_i}(f_{i,n}) - (\mathbb{E}_{i-1} R_{Y_i}^{(i)} f_{i,n})^2\big) \\ &= a_n^2 \gamma_{i-1}^2 \big(\widetilde{\mathfrak{m}}_{i-1} \mathbf{B}(f_{i,n}) - \big(\widetilde{\mathfrak{m}}_{i-1} \mathbf{R} f_{i,n}\big)^2\big) \\ &= a_n^2 \gamma_{i-1}^2 \big(\nu \mathbf{B}(f_{i,n}) + \mathfrak{v}_{i-1} \mathbf{B}(f_{i,n}) - \big(\widetilde{\mathfrak{m}}_{i-1} \mathbf{R} f_{i,n}\big)^2\big). \end{aligned} \tag{4.14}$$

We treat the three terms in the final parenthesis separately. We start with the third term; by (2.20) and (3.28),

$$\nu \mathbf{R} f_{i,n} = \nu f_{i,n} = \nu B_{i,n} f = \nu f = 0. \tag{4.15}$$

Hence, according to (3.65), and using $\mathfrak{v}_0 W < \infty$ when $i = 1$, there exists $\varepsilon > 0$ such that

$$a_n^2 \gamma_{i-1}^2 \,\mathbb{E}\left[|\widetilde{\mathfrak{m}}_{i-1}\mathbf{R} f_{i,n}|^2\right] = a_n^2 \gamma_{i-1}^2 \,\mathbb{E}\left[|\mathfrak{v}_{i-1}\mathbf{R} f_{i,n}|^2\right] \le \mathfrak{C} a_n^2 \gamma_{i-1}^2 i^{-\varepsilon} \|f_{i,n}\|_{B(W)}^2. \tag{4.16}$$

Furthermore, by (3.5) and Lemma 3.6, again using the decomposition (4.5) for (2), we have for all $n \ge i \ge 1$,

$$a_n^2 \gamma_{i-1}^2 \|f_{i,n}\|_{B(W)}^2 = a_n^2 \gamma_{i-1}^2 \|B_{i,n} f\|^2 \le \mathfrak{C} d_{i,n}. \tag{4.17}$$

We thus get that

$$\mathbb{E}\Big|\sum_{i=1}^n a_n^2 \gamma_{i-1}^2 (\widetilde{\mathfrak{m}}_{i-1}\mathbf{R} f_{i,n})^2\Big| \le \mathbb{E}\sum_{i=1}^n a_n^2 \gamma_{i-1}^2 \big|\widetilde{\mathfrak{m}}_{i-1}\mathbf{R} f_{i,n}^2\big| \le \mathfrak{C} \sum_{i=1}^n i^{-\varepsilon} d_{i,n} \xrightarrow[n\to+\infty]{} 0. \tag{4.18}$$

We now treat the second term in (4.14). Using (3.66) in Lemma 3.9, together with the fact that $\mathbf{B}$ is a bounded quadratic operator $B(W) \to B(W^2)$ (see (2.43)), and (4.17), we obtain that there exists $\varepsilon > 0$ such that

$$\begin{aligned} a_n^2 \gamma_{i-1}^2 \,\mathbb{E}\,|\mathfrak{v}_{i-1}\mathbf{B}(f_{i,n})| &\le \mathfrak{C} i^{-\varepsilon} a_n^2 \gamma_{i-1}^2 \|\mathbf{B}(f_{i,n})\|_{B(W^2)} \le \mathfrak{C} i^{-\varepsilon} a_n^2 \gamma_{i-1}^2 \|f_{i,n}\|_{B(W)}^2 \\ &\le \mathfrak{C} i^{-\varepsilon}\, d_{i,n} \end{aligned} \tag{4.19}$$

and hence

$$\sum_{i=1}^n a_n^2 \gamma_{i-1}^2 \,\mathbb{E}\big|\mathfrak{v}_{i-1}\mathbf{B}(f_{i,n})\big| \le \mathfrak{C} \sum_{i=1}^n i^{-\varepsilon} d_{i,n} \xrightarrow[n\to+\infty]{} 0. \tag{4.20}$$

We now consider the first term of (4.14), which needs a different treatment under the conditions of (1) and (2), so we treat the two cases separately.

Under the conditions of (1)*,* we rewrite the sum of the terms corresponding to the first term in (4.14) (note that this sum is non-random) as an integral:

$$\sum_{i=1}^n a_n^2 \gamma_{i-1}^2 \nu \mathbf{B}(f_{i,n}) = \sum_{i=1}^n n \gamma_{i-1}^2 \nu \mathbf{B}(f_{i,n}) = \int_0^1 n^2 \gamma_{\lceil nx \rceil - 1}^2 \nu \mathbf{B}(f_{\lceil xn \rceil, n}) \,\mathrm{d}x. \tag{4.21}$$

Using $\nu W^2 < \infty$ (implied by (N)), (2.43), and (4.17), we obtain that

$$\begin{aligned} n^2 \gamma_{\lceil nx \rceil - 1}^2 \big|\nu \mathbf{B}(f_{\lceil nx \rceil, n})\big| &\le n^2 \gamma_{\lceil nx \rceil - 1}^2 \|\mathbf{B}(f_{\lceil nx \rceil, n})\|_{B(W^2)} \le C n^2 \gamma_{\lceil nx \rceil - 1}^2 \|f_{\lceil nx \rceil, n}\|_{B(W)}^2 \\ &\le \mathfrak{C} n d_{\lceil nx \rceil, n} = \mathfrak{C} \lceil nx \rceil^{2\delta - 2} n^{2 - 2\delta} \le \mathfrak{C} x^{2\delta - 2}. \end{aligned} \tag{4.22}$$

Furthermore, for every fixed $x \in (0,1)$, we have by Lemma 3.3, uniformly for z in a compact set and all $n \ge 1$,

$$b_{\lceil nx \rceil, n}(z) = \Big(1 + O\Big(\frac{1}{\lceil nx \rceil}\Big)\Big) \left(\frac{n}{\lceil nx \rceil}\right)^{z-1} = x^{1-z} + O(1/n). \tag{4.23}$$

Hence, (3.37) shows that

$$\|B_{\lceil nx \rceil, n} - x^{1-\mathbf{R}}\| = \|b_{\lceil nx \rceil, n}(\mathbf{R}) - x^{1-\mathbf{R}}\| = O(1/n), \tag{4.24}$$

and, in particular,

$$f_{\lceil nx \rceil, n} = B_{\lceil nx \rceil, n} f \xrightarrow[n\to+\infty]{} x^{1-\mathbf{R}} f = x x^{-\mathbf{R}} f = x \mathrm{e}^{-(\log x)\mathbf{R}} f \tag{4.25}$$

in $B(W)$. Furthermore, $g \mapsto \nu g$ is a continuous linear functional on $B(W^2)$, since $\nu W^2 < \infty$, and thus, recalling Remark 2.15, $f \mapsto \nu \mathbf{B}(f)$ is a continuous quadratic form on $B(W)$. Hence, (4.25) implies

$$\nu \mathbf{B}(f_{\lceil xn \rceil, n}) \xrightarrow[n\to+\infty]{} \nu \mathbf{B}\big(x \mathrm{e}^{-(\log x)\mathbf{R}} f\big) = x^2 \nu \mathbf{B}\big(\mathrm{e}^{-(\log x)\mathbf{R}} f\big). \tag{4.26}$$

Moreover, we have $n^2 \gamma_{\lceil nx \rceil - 1}^2 \to x^{-2}$ when $n \to +\infty$. Consequently, for every fixed $x \in (0,1)$,

$$n^2 \gamma_{\lceil nx \rceil - 1}^2 \nu \mathbf{B}(f_{\lceil xn \rceil, n}) \xrightarrow[n\to+\infty]{} \nu \mathbf{B}\big(\mathrm{e}^{-(\log x)\mathbf{R}} f\big). \tag{4.27}$$

Consequently, by (4.21) and dominated convergence justified by (4.22), followed by a change of variables,

$$\sum_{i=1}^{n} a_n^2 \gamma_{i-1}^2 \nu\mathbf{B}(f_{i,n}) \xrightarrow[n\to+\infty]{} \int_0^1 \nu\mathbf{B}\big(\mathrm{e}^{-(\log x)\mathbf{R}} f\big)\,\mathrm{d}x = \int_0^\infty \nu\mathbf{B}\big(\mathrm{e}^{sR} f\big)\mathrm{e}^{-s}\,\mathrm{d}s$$
$$= \chi(f), \tag{4.28}$$

where it also follows that the integral is absolutely convergent as claimed in Theorem 2.10. The final equality in (2.28) follows by Fubini's theorem.

Under the conditions of (2), we observe that for $n, m \geq 1$,

$$\Big(\frac{n}{m}\Big)^{\mathbf{R}-\mathbf{I}} \Pi_{\lambda_j} f = \Big(\frac{n}{m}\Big)^{\lambda_j - 1} \Big(\frac{n}{m}\Big)^{\mathbf{R}-\lambda_j} \Pi_{\lambda_j} f = \Big(\frac{n}{m}\Big)^{\lambda_j - 1} \sum_{k\geq 0} \frac{1}{k!} (\log {}^{n}\!/\!{}_{m})^k (\mathbf{R}-\lambda_j)^k \Pi_{\lambda_j} f, \tag{4.29}$$

where the terms with $k \geq \kappa_j$ are null by (2.32). We deduce from (3.53), (4.5) and (4.29) that, for $n \geq m \geq 1$,

$$f_{m,n} = B_{m,n} f = (1 + h_{m,n}(\mathbf{R})) \sum_{j=1}^{p} \sum_{k=0}^{\kappa_j - 1} \Big(\frac{n}{m}\Big)^{\lambda_j - 1} \frac{1}{k!} \left(\log({}^{n}\!/\!{}_{m})\right)^k (\mathbf{R} - \lambda_j \mathbf{I})^k \Pi_{\lambda_j} f \tag{4.30}$$

where $\|h_{m,n}(\mathbf{R})\| = \mathcal{O}({}^{1}\!/\!{}_{m})$ on D by (3.43) and (3.37). We deduce that

$$\sum_{i=1}^{n} a_n^2 \gamma_{i-1}^2 \nu\mathbf{B}(f_{i,n}) = \sum_{j,j'=1}^{p} \sum_{k=0}^{\kappa_j - 1} \sum_{\ell=0}^{\kappa_{j'} - 1} \sum_{i=1}^{n} a_{n,i}^{(k,\ell,j,j')} x_{n,i}^{(k,\ell,j,j')}, \tag{4.31}$$

where

$$a_{n,i}^{(k,\ell,j,j')} = i^{-\lambda_j - \lambda_{j'}} n^{\lambda_j + \lambda_{j'} - 1} (\log {}^{n}\!/\!{}_{i})^{k+\ell} / (\log n)^{2\kappa - 1}, \tag{4.32}$$

and, using that $\widetilde{\mathbf{B}}$ is a bounded bilinear operator (Remark 2.15) and thus $\nu\widetilde{\mathbf{B}}$ is a bounded bilinear form on $B(W)$, and also (3.5),

$$\begin{aligned} x_{n,i}^{(k,\ell,j,j')} &= \frac{\gamma_{i-1}^2 i^2}{k!\ell!} \nu\widetilde{\mathbf{B}} \Big((1 + h_{i,n}(\mathbf{R}))(\mathbf{R} - \lambda_j \mathbf{I})^k \Pi_{\lambda_j} f, (1 + h_{i,n}(\mathbf{R}))(\mathbf{R} - \lambda_{j'} \mathbf{I})^\ell \Pi_{\lambda_{j'}} f \Big) \\ &= \frac{\gamma_{i-1}^2 i^2}{k!\ell!} \nu\widetilde{\mathbf{B}} \Big((\mathbf{R} - \lambda_j \mathbf{I})^k \Pi_{\lambda_j} f, (\mathbf{R} - \lambda_{j'} \mathbf{I})^\ell \Pi_{\lambda_{j'}} f \Big) + O({}^{1}\!/\!{}_{i}) \\ &= \frac{1}{k!\ell!} \nu\widetilde{\mathbf{B}} \Big((\mathbf{R} - \lambda_j \mathbf{I})^k \Pi_{\lambda_j} f, (\mathbf{R} - \lambda_{j'} \mathbf{I})^\ell \Pi_{\lambda_{j'}} f \Big) + O({}^{1}\!/\!{}_{i}). \end{aligned} \tag{4.33}$$

Fix k, ℓ, j, j' as in (4.31). By definition of $a_{n,i}^{(k,\ell,j,j')}$ (see (4.32)), and because $\mathrm{Re}(\lambda_j) = \mathrm{Re}(\lambda_{j'}) = {}^{1}\!/\!{}_{2}$, we get

$$\begin{aligned} \sum_{i=1}^{n} a_{n,i}^{(k,\ell,j,j')} &= \frac{n^{\mathrm{i}\,\mathrm{Im}(\lambda_j + \lambda_{j'})}}{(\log n)^{2\kappa - 1}} \sum_{i=1}^{n} i^{-1-\mathrm{i}\,\mathrm{Im}(\lambda_j + \lambda_{j'})} \log^{k+\ell}({}^{n}\!/\!{}_{i}) \\ &= \begin{cases} \frac{1+o(1)}{2\kappa - 1} & \text{if } \mathrm{Im}(\lambda_j + \lambda_{j'}) = 0 \text{ and } k = \ell = \kappa - 1, \\ o(1) & \text{if } \mathrm{Im}(\lambda_j + \lambda_{j'}) \neq 0 \text{ or } k < \kappa - 1 \text{ or } \ell < \kappa - 1, \end{cases} \end{aligned} \tag{4.34}$$

when $n \to +\infty$, where we refer to Lemma C.1 for detailed calculations. Furthermore,

$$\sum_{i=1}^{n} \big| a_{n,i}^{(k,\ell,j,j')} \big| \frac{1}{i} = \frac{1}{(\log n)^{2\kappa - 1}} \sum_{i=1}^{n} i^{-2} \log^{k+\ell}({}^{n}\!/\!{}_{i}) \leq \frac{(\log n)^{k+\ell}}{(\log n)^{2\kappa - 1}} \sum_{i=1}^{n} i^{-2} \leq \frac{C}{\log n} \to 0. \tag{4.35}$$

It folllows from (4.31), (4.33), (4.34) and (4.35) that, as $n \to \infty$,

$$\sum_{i=1}^{n} a_n^2 \gamma_{i-1}^2 \nu\mathbf{B}(f_{i,n}) \to \chi(f) :=$$

$$\sum_{j,j'=1}^{p} \frac{\mathbf{1}_{\kappa_j=\kappa_{j'}=\kappa,\ \overline{\lambda_j}=\lambda_{j'}}}{(2\kappa-1)((\kappa-1)!)^2} \nu\widetilde{\mathbf{B}}\left((\mathbf{R}-\lambda_j\mathbf{I})^{\kappa-1}\Pi_{\lambda_j}f,(\mathbf{R}-\lambda_{j'}\mathbf{I})^{\kappa-1}\Pi_{\lambda_{j'}}f\right). \quad (4.36)$$

Condition (c). Condition (c) of Theorem 4.1 is verified in the same way as Condition (b) above, with mainly notational differences. We therefore omit the details and only give a sketch. For $i \geq 1$, we have, corresponding to (4.14),

$$\mathbb{E}_{i-1}[|\hat{\zeta}_{n,i}^2|] = a_n^2\gamma_{i-1}^2\big(\nu\mathbf{C}(f_{i,n}) + \mathfrak{v}_{i-1}\mathbf{C}(f_{i,n}) - \left|\widetilde{\mathfrak{m}}_{i-1}\mathbf{R}f_{i,n}\right|^2\big), \quad (4.37)$$

with $\mathbf{C}$ defined in (2.30). As for (b), the two last terms can be neglected and, concerning the first term, we have the following convergence results, which depend on whether we work under the conditions of (1) or (2):

Under the conditions of (1). One shows that

$$\sum_{i=1}^{n} a_n^2\gamma_{i-1}^2\nu\mathbf{C}(f_{i,n}) \xrightarrow[n\to+\infty]{} \int_0^1 \nu\mathbf{C}\big(\mathrm{e}^{-(\log x)\mathbf{R}}f\big)\,\mathrm{d}x = \int_0^\infty \nu\mathbf{C}\big(\mathrm{e}^{sR}f\big)\mathrm{e}^{-s}\,\mathrm{d}s$$
$$= \sigma^2(f), \quad (4.38)$$

where the integral is absolutely convergent, as claimed in Theorem 2.10.

Under the conditions of (2). Using the same approach as above, but conjugating the second argument of $\widetilde{\mathbf{B}}$, we obtain, as $n \to +\infty$,

$$\sum_{i=1}^{n} a_n^2\gamma_{i-1}^2\nu\mathbf{C}(f_{i,n}) \to \sum_{j,j'=1}^{p} \frac{\mathbf{1}_{\kappa_j=\kappa_{j'}=\kappa,\ \lambda_j=\lambda_{j'}}}{(2\kappa-1)((\kappa-1)!)^2} \nu\widetilde{\mathbf{B}}\left((\mathbf{R}-\lambda_j\mathbf{I})^{\kappa-1}\Pi_{\lambda_j}f,\overline{(\mathbf{R}-\lambda_{j'}\mathbf{I})^{\kappa-1}\Pi_{\lambda_{j'}}f}\right). \quad (4.39)$$

Note that the condition $\lambda_j = \overline{\lambda_{j'}}$ in (4.36) has been changed into $\lambda_j = \lambda_{j'}$, so that the sum in (4.39) really is a single sum; hence (4.39) can be written, recalling (2.30),

$$\sum_{i=1}^{n} a_n^2\gamma_{i-1}^2\nu\mathbf{C}(f_{i,n}) \xrightarrow[n\to+\infty]{} \sum_{j=1}^{p} \frac{\mathbf{1}_{\kappa_j=\kappa}}{(2\kappa-1)((\kappa-1)!)^2} \nu\mathbf{C}\left((\mathbf{R}-\lambda_j\mathbf{I})^{\kappa-1}\Pi_{\lambda_j}f\right) = \sigma^2(f). \quad (4.40)$$

We have thus checked that under the conditions of either Theorem 2.13(1) or Theorem 2.13(2), the $\hat{\zeta}_{n,i}$ $(1 \leq i \leq n)$ satisfy Conditions (a), (b) and (c) of Theorem 4.1. The values of χ and σ^2 in (b) and (c) are given by (4.28) and (4.38) under the conditions of (1), and by (4.36) and (4.40) under the conditions of (2). Therefore, since $a_n\mathfrak{v}_nf = \sum_{i=0}^n \hat{\zeta}_{n,i}$, and since we have shown that $\zeta_{n,0} \to 0$ a.s., Theorem 4.1 yields the results (2.27) and (2.33).

4.2. **Proof of Theorem 2.13(3).** Next, consider the case of a generalized eigenfunction corresponding to an eigenvalue λ with $\operatorname{Re}\lambda > 1/2$. We state a lemma under slightly more general assumptions.

Lemma 4.2. *Suppose that D' is an $\mathbf{R}$-invariant subspace of $B(W)$ such that $\sigma(\mathbf{R}_{D'}) = \{\lambda\}$ consists of a single point λ with $1/2 < \operatorname{Re}\lambda \leq 1$. Then each operator $B_{m,n}$, $0 \leq m \leq n$, is invertible on D'. If $f \in D'$, then there exists a complex random variable Λ_f such that*

$$\mathfrak{v}_nB_{0,n}^{-1}f \to \Lambda_f \quad (4.41)$$

a.s. and in L^2 as $n \to \infty$; moreover, for any $0 < \varepsilon < \operatorname{Re}\lambda - 1/2$, there exists a constant $C_\varepsilon > 0$ such that

$$\mathbb{E}\left|\mathfrak{v}_nB_{0,n}^{-1}f - \Lambda_f\right|^2 \leq \frac{C_\varepsilon\,\widetilde{\mathfrak{m}}_0V}{n^{2(\operatorname{Re}\lambda-\varepsilon)-1}}\|f\|_{B(W)}^2. \quad (4.42)$$

Furthermore,

$$\mathbb{E}\,\Lambda_f = \mathfrak{v}_0f = \widetilde{\mathfrak{m}}_0f - \nu f \quad (4.43)$$

and

$$\mathbb{E}\,|\Lambda_f|^2 \le C\widetilde{\mathfrak{m}}_0 V\,\|f\|^2_{B(W)}. \tag{4.44}$$

Proof. First note that, since $\operatorname{Re}\lambda > 1/2$ and $0 < \gamma_k < 1$, we have

$$\big|1+\gamma_k(\lambda-1)\big| \ge \operatorname{Re}\big(1+\gamma_k(\lambda-1)\big) = 1-\gamma_k+\gamma_k\operatorname{Re}\lambda > 1/2, \qquad k \ge 0, \tag{4.45}$$

and thus $b_{m,n}(\lambda) \neq 0$ by (3.39). Hence $b_{m,n} \neq 0$ on $\sigma(\mathbf{R}_{D'})$, so $b_{m,n}^{-1}$ is analytic in a neighbourhood of $\sigma(\mathbf{R}_{D'}) = \{\lambda\}$ and it follows that, as an operator on D', $B_{m,n} = b_{m,n}(\mathbf{R}_{D'})$ is invertible with inverse $B_{m,n}^{-1} = b_{m,n}^{-1}(\mathbf{R}_{D'})$.

If $0 \le i \le n$, then (3.72) (or (3.3)) shows that $B_{0,i}B_{i,n} = B_{0,n}$, which yields

$$B_{i,n}B_{0,n}^{-1} = B_{0,i}^{-1}. \tag{4.46}$$

Let $f \in D'$. By (3.26), applied to $B_{0,n}^{-1}f$, and (4.46), we have

$$\mathfrak{v}_n B_{0,n}^{-1} f = \mathfrak{v}_0 f + \sum_{i=1}^{n} \gamma_{i-1}\Delta M_i B_{0,i}^{-1} f. \tag{4.47}$$

Since $\mathbb{E}_{i-1}\Delta M_i = 0$, (4.47) shows that $\big(\mathfrak{v}_n B_{0,n}^{-1}f\big)_{n\ge0}$ is a martingale. (Cf. the closely related (3.73).) We will show that the martingale (4.47) is L^2 bounded; the result (4.41) then follows by the martingale convergence theorem.

Let $0 < \varepsilon < \operatorname{Re}\lambda - 1/2$, and let K be the closed disc $\{z : |z-\lambda| \le \varepsilon\}$. Then $\operatorname{Re} z > 1/2$ for all $z \in K$, and it follows, as in (4.45), that $|1+\gamma_k(z-1)| > 1/2$ on K. Hence, each $b_{m,n}$ is non-zero on K, and thus is invertible on K; furthermore, (3.39) gives the trivial bound

$$\big|b_{m,n}(z)^{-1}\big| \le 2^{n-m}, \qquad 0 \le m \le n,\ z \in K. \tag{4.48}$$

To get a better bound, we fix $m \ge 1$ such that (3.45) holds for all $n \ge m$ and $z \in K$. Thus, using (4.48) for $b_{0,m}(z)^{-1}$,

$$\begin{aligned}\big|b_{0,n}(z)^{-1}\big| &= \big|b_{0,m}(z)^{-1}b_{m,n}(z)^{-1}\big| \le C|b_{m,n}(z)|^{-1} = C\big|\mathrm{e}^{-(z-1)\log\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+m}+O(1)}\big| \\ &\le C\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+m}\right)^{1-\operatorname{Re} z} \le C\left(\mathfrak{m}_0(E)+n\right)^{1-\operatorname{Re}\lambda+\varepsilon}. \end{aligned} \tag{4.49}$$

By (4.48), the same bound holds trivially (with a suitable C) also for $1 \le n < m$, so the estimate (4.49) holds for all $n \ge 1$ and $z \in K$.

Since $\sigma(\mathbf{R}_{D'}) = \{\lambda\}$, it follows from (3.37), taking Γ to be the circle $\{z : |z-\lambda| = \varepsilon\} \subset K$, together with (4.49) that

$$\|B_{0,n}^{-1}\|_{D'} \le C\sup_{z\in K}|b_{0,n}^{-1}(z)| \le C\left(\mathfrak{m}_0(E)+n\right)^{1-\operatorname{Re}\lambda+\varepsilon}, \qquad n \ge 1. \tag{4.50}$$

For $f \in D'$ and $i \ge 1$, we thus have, by (3.12) and (4.50),

$$\begin{aligned}\mathbb{E}\big|\gamma_{i-1}\Delta M_i B_{0,i}^{-1} f\big|^2 &\le C\gamma_{i-1}^2\widetilde{\mathfrak{m}}_0 V\,\|B_{0,i}^{-1}f\|^2_{D'} \\ &\le C\widetilde{\mathfrak{m}}_0 V\,(\mathfrak{m}_0(E)+i)^{-2}(\mathfrak{m}_0(E)+i)^{2(1-\operatorname{Re}\lambda+\varepsilon)}\|f\|^2_{D'} \\ &= C\widetilde{\mathfrak{m}}_0 V\,(\mathfrak{m}_0(E)+i)^{-2(\operatorname{Re}\lambda-\varepsilon)}\|f\|^2_{D'}. \end{aligned} \tag{4.51}$$

Since $2(\operatorname{Re}\lambda-\varepsilon) > 1$, it follows from (4.47) (where the terms are orthogonal), (3.11), and (4.51), that

$$\mathbb{E}\,\big|\mathfrak{v}_n B_{0,n}^{-1} f\big|^2 \le |\mathfrak{v}_0 f|^2 + \sum_{i=1}^{\infty}\mathbb{E}\big|\gamma_{i-1}\Delta M_i B_{0,i}^{-1} f\big|^2 \le C\widetilde{\mathfrak{m}}_0 V\,\|f\|^2_{D'}, \tag{4.52}$$

and thus the martingale (4.47) converges in L^2 and a.s., as claimed. The properties (4.42), (4.43) and (4.44) immediately follow from (4.47) and (4.51)–(4.52). ☐

We combine Lemma 4.2 with a standard result for functions of nilpotent operators.

Lemma 4.3. *Suppose that D' is an $\mathbf{R}$-invariant subspace of $B(W)$ such that $(\mathbf{R}_{D'} - \lambda)^\kappa = 0$ for some complex λ and integer $\kappa \geq 1$. Let h be a function that is analytic in a neighbourhood of λ. Then, for $f \in D'$,*

$$h(\mathbf{R}_{D'})f = \sum_{k=0}^{\kappa-1} \frac{h^{(k)}(\lambda)}{k!}(\mathbf{R} - \lambda)^k f. \tag{4.53}$$

Proof. A Taylor expansion yields, for some function h_κ analytic in the same domain as h,

$$h(z) = \sum_{k=0}^{\kappa-1} \frac{h^{(k)}(\lambda)}{k!}(z - \lambda)^k + (z - \lambda)^\kappa h_\kappa(z). \tag{4.54}$$

We have $(\mathbf{R}_{D'} - \lambda)^\kappa = 0$ by assumption, and thus (4.54) yields, using (3.36),

$$h(\mathbf{R}_{D'}) = \sum_{k=0}^{\kappa-1} \frac{h^{(k)}(\lambda)}{k!}(\mathbf{R}_{D'} - \lambda)^k, \tag{4.55}$$

as operators on D', which is (4.53). □

We can now show the convergence (2.38) for f in a generalized eigenspace.

Lemma 4.4. *Suppose that D' is an $\mathbf{R}$-invariant subspace of $B(W)$ such that $(\mathbf{R}_{D'} - \lambda)^\kappa = 0$ for some complex λ with $1/2 < \operatorname{Re}\lambda \leq 1$ and some integer $\kappa \geq 1$. If $f \in D'$, then, for some complex random variable Λ,*

$$\frac{n^{1-\lambda}}{\log^{\kappa-1} n}\, \mathfrak{v}_n f \to \Lambda \tag{4.56}$$

a.s. and in L^2 as $n \to \infty$. Furthermore,

$$\mathbb{E}\,\Lambda = \frac{\Gamma(\mathfrak{m}_0(E) + 1)}{(\kappa - 1)!\,\Gamma(\mathfrak{m}_0(E) + \lambda)} \mathfrak{v}_0(\mathbf{R} - \lambda)^{\kappa-1} f. \tag{4.57}$$

Proof. Note that the assumption $(\mathbf{R}_{D'} - \lambda)^\kappa = 0$ implies that $\sigma(\mathbf{R}_{D'}) = \{\lambda\}$, for example by the spectral mapping theorem [5, Theorem VII.4.10]. Hence, Lemma 4.2 applies.

We use also Lemma 4.3 with $h = b_{0,n}$. This yields, defining $f_k := (\mathbf{R} - \lambda)^k f / k!$,

$$B_{0,n} f = b_{0,n}(\mathbf{R}_{D'}) f = \sum_{k=0}^{\kappa-1} b_{0,n}^{(k)}(\lambda) f_k \tag{4.58}$$

and thus

$$\mathfrak{v}_n f = \mathfrak{v}_n B_{0,n}^{-1} B_{0,n} f = \sum_{k=0}^{\kappa-1} b_{0,n}^{(k)}(\lambda) \mathfrak{v}_n B_{0,n}^{-1} f_k. \tag{4.59}$$

Each random variable $\mathfrak{v}_n B_{0,n}^{-1} f_k$ converges a.s. and in L^2 as $n \to \infty$ by Lemma 4.2, and it remains to study the coefficients $b_{0,n}^{(k)}(\lambda)$. By (3.41) we have

$$b_{0,n}(z) = \bigl(1 + h_{0,n}(z)\bigr) \left(\frac{\mathfrak{m}_0(E) + n}{\mathfrak{m}_0(E) + 1}\right)^{z-1}. \tag{4.60}$$

In a fixed neighbourhood of λ, the functions $h_{0,n}(z)$, $n \geq 1$, are uniformly bounded by (3.43), and thus Cauchy's estimates show that for each fixed $k \geq 0$,

$$\bigl|h_{0,n}^{(k)}(\lambda)\bigr| \leq C. \tag{4.61}$$

Furthermore,

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k} \left(\frac{\mathfrak{m}_0(E) + n}{\mathfrak{m}_0(E) + 1}\right)^{z-1} = \log^k \left(\frac{\mathfrak{m}_0(E) + n}{\mathfrak{m}_0(E) + 1}\right) \left(\frac{\mathfrak{m}_0(E) + n}{\mathfrak{m}_0(E) + 1}\right)^{z-1}. \tag{4.62}$$

Hence, using (4.60) and Leibniz' rule, for a fixed k and $n \geq 1$,

$$\left| b_{0,n}^{(k)}(\lambda) - \big(1 + h_{0,n}(\lambda)\big) \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{\lambda-1} \log^k \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right) \right| \\ \leq C \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{\operatorname{Re}\lambda-1} \left(1 \vee \log^{k-1}\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)\right). \tag{4.63}$$

By (4.63), each coefficient $b_{0,n}^{(k)}(\lambda)$ in (4.59) with $k < \kappa - 1$ satisfies

$$\left| \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{1-\lambda} \frac{b_{0,n}^{(k)}(\lambda)}{1 \vee \log^{\kappa-1}\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)} \right| \leq \frac{C}{1 \vee \log\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)}. \tag{4.64}$$

Using (4.63) with $k = \kappa - 1$, we similarly deduce that

$$\left| \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{1-\lambda} \frac{b_{0,n}^{(\kappa-1)}(\lambda)}{1 \vee \log^{\kappa-1}\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)} - \big(1 + h_{0,n}(\lambda)\big) \right| \leq \frac{C}{1 \vee \log\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)}. \tag{4.65}$$

We obtain from (4.59), (4.64), (4.65), and the fact that $h_{0,n}(\lambda)$ is uniformly bounded, with $\Lambda_{f_{\kappa-1}}$ from Lemma 4.2,

$$\left| \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{1-\lambda} \frac{\mathfrak{v}_n f}{1 \vee \log^{\kappa-1}\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)} - \big(1 + h_{0,n}(\lambda)\big)\Lambda_{f_{\kappa-1}} \right| \\ \leq C \frac{\sum_{k=0}^{\kappa-1} |\mathfrak{v}_n B_{0,n}^{-1} f_k|}{1 \vee \log\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)} + C \left| \mathfrak{v}_n B_{0,n}^{-1} f_{\kappa-1} - \Lambda_{f_{\kappa-1}} \right|. \tag{4.66}$$

In addition, (3.41), (3.46) and a well-known consequence of Stirling's formula (see e.g. [29, 5.11.13]) imply that for any fixed z,

$$\begin{aligned} 1 + h_{0,n}(z) = \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{1-z} b_{0,n}(z) &= \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{1-z} \frac{\Gamma(n+\mathfrak{m}_0(E)+z)}{\Gamma(n+\mathfrak{m}_0(E)+1)} \frac{\Gamma(\mathfrak{m}_0(E)+1)}{\Gamma(\mathfrak{m}_0(E)+z)} \\ &= \frac{1}{(\mathfrak{m}_0(E)+1)^{1-z}} \frac{\Gamma(\mathfrak{m}_0(E)+1)}{\Gamma(\mathfrak{m}_0(E)+z)} \Big(1 + O\Big(\frac{1}{n}\Big)\Big) \\ &= \frac{1}{(\mathfrak{m}_0(E)+1)^{1-z}} \frac{\Gamma(\mathfrak{m}_0(E)+1)}{\Gamma(\mathfrak{m}_0(E)+z)} + O\Big(\frac{1}{n}\Big). \end{aligned} \tag{4.67}$$

Lemma 4.2 implies that the right-hand side of (4.66) tends to 0 a.s. as $n \to \infty$. Hence, (4.66) and (4.67) imply that

$$(\mathfrak{m}_0(E)+n)^{1-\lambda} \frac{\mathfrak{v}_n f}{\log^{\kappa-1}\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)} \to \Lambda \tag{4.68}$$

holds a.s. with

$$\Lambda = \frac{\Gamma(\mathfrak{m}_0(E)+1)}{\Gamma(\mathfrak{m}_0(E)+\lambda)} \Lambda_{f_{\kappa-1}}. \tag{4.69}$$

We may simplify (4.68) and conclude that (4.56) holds a.s.

Moreover, (4.42) and (4.44) imply that for every fixed $k \geq 0$,

$$\mathbb{E}\big|\mathfrak{v}_n B_{0,n}^{-1} f_k\big|^2 \leq C \widetilde{\mathfrak{m}}_0 V \|f_k\|_{B(W)}^2 + C\,\mathbb{E}\big|\Lambda_{f_k}\big|^2 \leq C \widetilde{\mathfrak{m}}_0 V \|f_k\|_{B(W)}^2 \leq C \widetilde{\mathfrak{m}}_0 V \|f\|_{B(W)}^2. \tag{4.70}$$

Taking the expectation of the square in (4.66), we deduce using (4.70) and (4.42),

$$\mathbb{E} \left| \left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)^{1-\lambda} \frac{\mathfrak{v}_n f}{1 \vee \log^{\kappa-1}\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)} - \big(1 + h_{0,n}(\lambda)\big)\Lambda_{f_{\kappa-1}} \right|^2 \leq \frac{C\,\widetilde{\mathfrak{m}}_0 V\,\|f\|_{B(W)}^2}{1 \vee \log^2\left(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\right)} \tag{4.71}$$

Furthermore, by (4.67), (4.69) and (4.44),

$$\mathbb{E}\left|\bigl(1+h_{0,n}(\lambda)\bigr)\Lambda_{f_{\kappa-1}}-\frac{1}{(m_0(E)+1)^{1-\lambda}}\Lambda\right|^2\leq\frac{C}{n^2}\,\mathbb{E}\bigl|\Lambda_{f_{\kappa-1}}\bigr|^2\leq\frac{C}{n^2}\widetilde{\mathfrak{m}}_0V\|f\|^2_{B(W)}. \tag{4.72}$$

Combining (4.71) and (4.72), we obtain

$$\begin{aligned}\mathbb{E}\left|\bigl(\mathfrak{m}_0(E)+n\bigr)^{1-\lambda}\frac{\mathfrak{v}_nf}{1\vee\log^{\kappa-1}\bigl(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\bigr)}-\Lambda\right|^2&\\ \leq C\bigl(\mathfrak{m}_0(E)+1\bigr)^{2(1-\operatorname{Re}\lambda)}\frac{\widetilde{\mathfrak{m}}_0V}{1\vee\log^2\bigl(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\bigr)}\|f\|^2_{B(W)}.&\end{aligned} \tag{4.73}$$

Hence, (4.68) holds also in L^2, and thus so does (4.56). Finally, (4.57) follows by (4.69), (4.43), and the definition of $f_{\kappa-1}$, which completes the proof. □

We are now ready to prove Theorem 2.13(3); we assume that the operator $\mathbf{R}_D$ is slqc and that the spectrum of $\mathbf{R}_D$ is given by

$$\sigma(\mathbf{R}_D)=\{1,\lambda_1,\dots,\lambda_p\}\cup\Delta,\qquad p\geq 1, \tag{4.74}$$

where $\operatorname{Re}(\lambda_1)=\dots=\operatorname{Re}(\lambda_p)=\theta_D\in(1/2,1)$, and $\sup\operatorname{Re}(\Delta)<\theta_D$. Note that this implies that $\lambda_1,\dots,\lambda_j$ are isolated points of the spectrum $\sigma(\mathbf{R}_D)$, and that Δ is a clopen subset. Thus the spectral projections Π_{λ_j} and Π_Δ are defined and, for any $f\in D$, recalling Lemma 3.5,

$$f=\Pi_1f+\sum_{j=1}^p\Pi_{\lambda_j}f+\Pi_\Delta f=\nu f+\sum_{j=1}^p\Pi_{\lambda_j}f+\Pi_\Delta f. \tag{4.75}$$

Hence, it suffices to prove (2.38) for the functions νf, $\Pi_{\lambda_j}f$ and $\Pi_\Delta f$ separately; in other words, it suffices to consider the cases $f=c$ constant, $f\in\Pi_{\lambda_j}D$ and $f\in\Pi_\Delta D$. Recall that $\widetilde{\mathfrak{m}}_n-\nu=\mathfrak{v}_n$.

First, we may ignore the constant term νf in (4.75), since $\mathfrak{v}_n1=0$.

Secondly, Lemma 4.4 applies to each space $\Pi_{\lambda_j}D$, since we assume (2.32) and thus $(\mathbf{R}-\lambda_j)^\kappa=0$ on $D_j:=\Pi_{\lambda_j}D$. It follows that, for some complex random variable $\Lambda_j\in L^2$,

$$\frac{n^{1-\operatorname{Re}\lambda_j}}{\log^{\kappa-1}n}\mathfrak{v}_n\Pi_{\lambda_j}f-n^{\mathrm{i}\operatorname{Im}\lambda_j}\Lambda_j\to 0 \tag{4.76}$$

a.s. and in L^2. Furthermore, (2.20) and Lemma 3.5 imply that

$$\nu(\mathbf{R}-\lambda_j)^{\kappa-1}\Pi_{\lambda_j}f=(1-\lambda_j)^{\kappa-1}\nu\Pi_{\lambda_j}f=0, \tag{4.77}$$

so that (4.57) yields (2.39).

Thirdly, Lemma 3.11 applies to $\Pi_\Delta D$ and $\theta':=\theta_D=\operatorname{Re}\lambda_1$, and shows

$$n^{1-\operatorname{Re}\lambda_1}\mathfrak{v}_n\Pi_\Delta f\to 0 \tag{4.78}$$

a.s. and in L^2.

Theorem 2.13(3) follows by combining (4.75) with (4.76) and (4.78). This completes the proof of Theorem 2.13. □

Remark 4.5. Note that (4.73) implies an upper bound $O\bigl(1/\log n\bigr)$ for the speed of convergence in L^2 of (4.68), which yields the same rate in (4.56) in Lemma 4.4. Since, in addition, (3.86) in the proof of Lemma 3.11 yields an upper bound $O(n^{-\varepsilon})$ (for some $\varepsilon=\theta'-\theta''>0$) for the speed of convergence in L^2 in (3.85), one finds $O\bigl(1/\log n\bigr)$ as an explicit upper bound for the speed of convergence in L^2 of Theorem 2.13(3). □

Remark 4.6. If $\mathfrak{m}_0$ is random, then under the conditions of Lemma 4.4, (4.73) holds conditioned on $\mathfrak{m}_0$. Taking the expectation, we see by dominated convergence that if further

$$\mathbb{E}\big[\big(\mathfrak{m}_0(E)+1\big)^{2(1-\operatorname{Re}\lambda)}\widetilde{\mathfrak{m}}_0 V\big] < \infty, \tag{4.79}$$

then the left-hand side of (4.73) converges to 0 as $n\to\infty$. With the notation

$$a_n := \frac{(\mathfrak{m}_0(E)+n)^{1-\lambda}}{1\vee \log^{\kappa-1}\big(\frac{\mathfrak{m}_0(E)+n}{\mathfrak{m}_0(E)+1}\big)}, \tag{4.80}$$

this says that $a_n\mathfrak{v}_n f \to \Lambda$ in L^2. Since we also have convergence a.s. (by (4.68) and conditioning on $\mathfrak{m}_0$), this implies that the sequence $|a_n\mathfrak{v}_n f|^2$ is uniformly integrable, see e.g. [12, Theorem 5.5.2]. Let $b_n := n^{1-\lambda}/\log^{\kappa-1} n$. Then, for $n\geq 3$, $|b_n|\leq |a_n|$, and it follows that also $|b_n\mathfrak{v}_n f|^2$ is uniformly integrable. Furthermore, also $b_n\mathfrak{v}_n f\to\Lambda$ a.s., and thus [12, Theorem 5.5.2] again shows that $b_n\mathfrak{v}_n f \to \Lambda$ in L^2. Consequently, under the assumption (4.79), (4.56) holds both a.s. and in L^2.

By combining this and Remark 3.12, it follows as above that (2.38) in Theorem 2.13(3) holds also in L^2 for random $\mathfrak{m}_0$ that satisfies (2.41), as claimed in Remark 2.14. □

5. Proof of Theorems 2.25–2.27.

In this section we prove Theorems 2.25–2.27 on possible degeneracies in the limit distributions in Theorem 2.13.

Lemma 5.1. *Suppose that $f\in B(W)$ and that $\nu|f|=0$, or, equivalently,*

$$f(x)=0 \text{ for } \nu\text{-a.e. } x. \tag{5.1}$$

(i) *Then $\nu|\mathbf{R}f|=0$, i.e., (5.1) holds for $\mathbf{R}f$ too.*
(ii) *Moreover, for ν-a.e. x,*

$$R_x^{(1)}f=0 \qquad a.s. \tag{5.2}$$

Proof. By linearity we may assume that $f\geq 0$. Let $N:=\{x: f(x)\neq 0\}$; then $\nu N=0$ by the assumption (5.1). If $x\notin N$, then $R_x^{(1)}f\geq 0$, because $R_x^{(1)}$ is positive on $E\setminus\{x\}$ and $f(x)=0$. Hence, by taking the expectation, also

$$\mathbf{R}f(x)=\mathbb{E}\,R_x^{(1)}f\geq 0, \qquad x\notin N. \tag{5.3}$$

Thus $\mathbf{R}f\geq 0$ ν-a.e.

On the other hand, by (2.20) and the assumption (5.1),

$$\nu(\mathbf{R}f)=(\nu\mathbf{R})f=\nu f=0. \tag{5.4}$$

It follows from (5.3) and (5.4) that $\mathbf{R}f=0$ ν-a.e., which proves (i).

Moreover, let $N_1:=\{x:\mathbf{R}f(x)\neq 0\}$. If $x\notin N\cup N_1$, then, as just shown, $R_x^{(1)}f\geq 0$, and also $\mathbb{E}\,R_x^{(1)}f=\mathbf{R}f(x)=0$; hence, (5.2) holds. This proves (ii), since $\nu(N_1)=0$ by (i), and thus $\nu(N\cup N_1)=0$. □

Proof of Theorem 2.25. By replacing f by $f-\nu f$, we may for simplicity assume $\nu f=0$.

Note first that (2.26) implies $\sigma^2(f)=0 \implies \chi(f)=0$. Hence, (i) $\iff$ (ii) follows from the formula for $\Sigma(f)$ in (2.27).

Next, $s\mapsto e^{s\mathbf{R}}f$ is a continuous map $[0,\infty)\to B(W)$, and thus, by Remark 2.15 and $\nu W^2<\infty$, $s\mapsto \nu\mathbf{C}\big(e^{s\mathbf{R}}f\big)$ is a continuous function of $s\geq 0$. Furthermore, by (2.30), we have $\mathbf{C}(e^{s\mathbf{R}}f)\geq 0$, and thus $\nu\mathbf{C}(e^{s\mathbf{R}}f)\geq 0$. Consequently, by (2.29),

$$\sigma^2(f)=0 \iff \nu\mathbf{C}(e^{s\mathbf{R}}f)=0 \qquad \text{for every } s\geq 0. \tag{5.5}$$

In particular, taking $s=0$, we see that (ii) $\implies$ (iii).

Furthermore, $\nu\mathbf{C}(f)=0 \iff \mathbf{C}_x f=0$ for ν-a.e. x, which by (2.30) is equivalent to $R_x^{(1)}f=0$ a.s., for ν-a.e. x. Hence, (iii) $\iff$ (iv).

Finally assume (iv), and let $N \subset E$ be a set with $\nu(N) = 0$ such that $R_x^{(1)} f = 0$ a.s. when $x \notin N$. By taking the expectation, we obtain $\mathbf{R}f(x) = \mathbb{E}\, R_x^{(1)} f = 0$ for $x \notin N$. Hence, $\mathbf{R}f = 0$ ν-a.e., i.e., $\mathbf{R}f$ satisfies (5.1). We may thus apply Lemma 5.1 to $\mathbf{R}f$ and conclude by induction that $\mathbf{R}^k f = 0$ ν-a.e., for every $k \geq 1$. Consequently, for any $s \geq 0$,

$$\mathrm{e}^{s\mathbf{R}} f - f = \sum_{k=1}^{\infty} \frac{s^k}{k!} \mathbf{R}^k f = 0 \qquad \nu\text{-a.e.} \tag{5.6}$$

We apply Lemma 5.1 again, this time to $\mathrm{e}^{s\mathbf{R}} f - f$, and conclude by Lemma 5.1(ii) that for ν-a.e. x,

$$R_x^{(1)}\big(\mathrm{e}^{s\mathbf{R}} f - f\big) = 0 \qquad a.s. \tag{5.7}$$

Together with the assumption $R_x^{(1)} f = 0$ a.s. for ν-a.e. x, this shows that for ν-a.e. x,

$$R_x^{(1)}\big(\mathrm{e}^{s\mathbf{R}} f\big) = 0 \qquad a.s. \tag{5.8}$$

Hence, (2.30) yields $\mathbf{C}_x(e^{s\mathbf{R}} f) = 0$ for ν-a.e. x, and thus $\nu\mathbf{C}(e^{s\mathbf{R}} f) = 0$, for every $s \geq 0$. Consequently, (5.5) shows that $\sigma^2(f) = 0$. We have shown that (iv) $\Longrightarrow$ (ii), which completes the proof. ☐

Proof of Theorem 2.26. The equivalence (i) $\Longleftrightarrow$ (ii) follows as in the proof of Theorem 2.25.

Let $g_j := (\mathbf{R} - \lambda_j \mathbf{I})^{\kappa-1} \Pi_{\lambda_j} f$. Note that by (2.32), $g_j = 0$ if $\kappa_j < \kappa$, which shows the equivalence (iii) $\Longleftrightarrow$ (iv).

By (2.35), it remains only to show that

$$\nu\mathbf{C}(g_j) = 0 \iff g_j = 0 \ \nu\text{-a.e.} \tag{5.9}$$

To see this, we first note that by definition of κ,

$$(\mathbf{R} - \lambda_j) g_j = (\mathbf{R} - \lambda_j)^{\kappa} \Pi_j f = 0. \tag{5.10}$$

and thus

$$\mathbf{R} g_j = \lambda_j g_j. \tag{5.11}$$

In other words, g_j is (if non-zero) an eigenfunction with eigenvalue $\lambda_j \neq 0$.

Assume now $\nu\mathbf{C}(g_j) = 0$. Then, using (2.30) again, for ν-a.e. x, we have $\mathbf{C}_x(g_j) = 0$ and thus $R_x^{(1)} g_j = 0$ a.s. Taking the expectation shows that for such x, we have $\mathbf{R}g_j(x) = \mathbb{E}\, R_x^{(1)} g_j = 0$. Consequently, $\mathbf{R}g_j = 0$ ν-a.e., and (5.11) implies $g_j = 0$ ν-a.e. This shows one implication in (5.9).

Conversely, assume $g_j = 0$ ν-a.e. Then Lemma 5.1 shows that for ν-a.e. x, we have $R_x^{(1)} g_j = 0$ a.s., and thus $\mathbf{C}_x(g_j) = 0$ by (2.30). Hence, $\nu\mathbf{C}(g_j) = 0$. This completes the proof of (5.9), and thus of (ii) $\Longleftrightarrow$ (iv), and of the theorem. ☐

Proof of Theorem 2.27. We note first that (5.11) holds in the present case too, and thus g_j is an eigenfunction of $\mathbf{R}$ with eigenvalue $\lambda_j \neq 1$. Hence, ν and g_j are left and right eigenvectors of $\mathbf{R}$ with different eigenvalues (recall (2.20)), which implies, as is well known,

$$\nu g_j = 0, \tag{5.12}$$

because we have

$$\nu g_j = (\nu\mathbf{R}) g_j = \nu(\mathbf{R} g_j) = \lambda_j (\nu g_j). \tag{5.13}$$

(i) $\Longleftrightarrow$ (ii): Obvious.

(ii) $\Longrightarrow$ (iv),(vi): Suppose now that (ii) holds, i.e., $\Lambda_j = \mathbb{E}\,\Lambda_j$ a.s. The proofs of Theorem 2.13(3) and Lemma 4.4 (in particular (4.69)) show that

$$\Lambda_j = c\Lambda_{g_j} \tag{5.14}$$

where $c > 0$ is an explicit constant and Λ_{g_j} is given by Lemma 4.2. Since Λ_{g_j} is constructed in the proof of Lemma 4.2 as the limit of the martingale (4.47) (with f replaced by g_j), it follows that $\Lambda_j = \mathbb{E}\,\Lambda_j$ a.s. if and only if all martingale differences in (4.47) vanish a.s., i.e.,

$$\Lambda_j = \mathbb{E}\,\Lambda_j \text{ a.s.} \iff \Delta M_i B_{0,i}^{-1} g_j = 0 \text{ a.s., for every } i \geq 1. \tag{5.15}$$

Moreover, as remarked above, $\mathbf{R} g_j = \lambda_j g_j$, Hence, by (3.38),

$$B_{0,i}^{-1} g_j = b_{0,i}(\mathbf{R})^{-1} g_j = b_{0,i}(\lambda_j)^{-1} g_j, \tag{5.16}$$

where $b_{0,i}(\lambda_j) \neq 0$ by (3.39), $0 < \gamma_n < 1$, and $\operatorname{Re}\lambda_j > 0$. Thus (5.15) yields

$$\Delta M_i g_j = 0 \text{ a.s., for every } i \geq 1. \tag{5.17}$$

Using also (3.4) (and replacing i by $n+1$), (5.17) says that, for every $n \geq 0$,

$$R_{Y_{n+1}}^{(n+1)} g_j = \widetilde{\mathfrak{m}}_n \mathbf{R} g_j \quad \text{a.s.} \tag{5.18}$$

Conditioning on $\widetilde{\mathfrak{m}}_n$, and recalling that Y_{n+1} has the conditional distribution $\widetilde{\mathfrak{m}}_n$, we see that (5.18) implies that $\widetilde{\mathfrak{m}}_n$ is a.s. such that, conditioned on $\widetilde{\mathfrak{m}}_n$,

$$R_x^{(n+1)} g_j = \widetilde{\mathfrak{m}}_n \mathbf{R} g_j \quad \text{a.s., for } \widetilde{\mathfrak{m}}_n\text{-a.e. } x. \tag{5.19}$$

Consider first the case $n = 0$. Recall that $\widetilde{\mathfrak{m}}_0$ is non-random, and let $a := \widetilde{\mathfrak{m}}_0 \mathbf{R} g_j$ (a non-random real number). Then the case $n = 0$ of (5.19) says

$$R_x^{(1)} g_j = a \quad \text{a.s., for } \widetilde{\mathfrak{m}}_0\text{-a.e. } x. \tag{5.20}$$

Now return to a general $n \geq 0$. Since $\widetilde{\mathfrak{m}}_n$ is a positive number times $\mathfrak{m}_n$, we may in (5.19) equivalently write "for $\mathfrak{m}_n$-a.e. x". Furthermore, $\mathfrak{m}_n \geq \mathfrak{m}_0$, and thus (5.19) implies that the equality holds for $\mathfrak{m}_0$-a.e. x. Moreover, $R_x^{(n+1)}$ is independent of $\widetilde{\mathfrak{m}}_n$, and thus its conditional distribution equals the distribution of $R_x^{(1)}$. Hence, (5.20) shows that, also conditioned on $\widetilde{\mathfrak{m}}_n$,

$$R_x^{(n+1)} g_j = a \quad \text{a.s., for } \widetilde{\mathfrak{m}}_0\text{-a.e. } x. \tag{5.21}$$

Consequently, comparing (5.19) and (5.21), we obtain, for every $n \geq 0$,

$$\widetilde{\mathfrak{m}}_n \mathbf{R} g_j = a \quad \text{a.s.} \tag{5.22}$$

Thus, (5.19) shows that, for every $n \geq 0$, $\widetilde{\mathfrak{m}}_n$ is a.s. such that

$$R_x^{(1)} g_j = a \quad \text{a.s., for } \widetilde{\mathfrak{m}}_n\text{-a.e. } x. \tag{5.23}$$

By again conditioning on $\widetilde{\mathfrak{m}}_n$, it follows from (5.23) that

$$R_{Y_{n+1}}^{(n+1)} g_j = a \quad \text{a.s.} \tag{5.24}$$

Consequently, by (2.2) and induction,

$$\mathfrak{m}_n g_j = \mathfrak{m}_0 g_j + na \quad \text{a.s.} \tag{5.25}$$

On the other hand, taking the expectation in (5.23) yields that $\widetilde{\mathfrak{m}}_n$ is a.s. such that

$$\mathbf{R} g_j(x) = \mathbb{E}\, R_x^{(1)} g_j = a \quad \text{for } \widetilde{\mathfrak{m}}_n\text{-a.e. } x. \tag{5.26}$$

Recalling (5.11), this implies that

$$\widetilde{\mathfrak{m}}_n g_j = \lambda_j^{-1} \widetilde{\mathfrak{m}}_n(\mathbf{R} g_j) = \lambda_j^{-1} a \quad \text{a.s.} \tag{5.27}$$

Consequently, recalling (2.3),

$$\mathfrak{m}_n g_j = \mathfrak{m}_n(E) \widetilde{\mathfrak{m}}_n g_j = \bigl(\mathfrak{m}_0(E) + n\bigr)\lambda_j^{-1} a = \mathfrak{m}_0(E) a/\lambda_j + na/\lambda_j \quad \text{a.s.} \tag{5.28}$$

Comparing (5.25) and (5.28), we see that $a = a/\lambda_j$, and thus (since $\lambda_j \neq 1$), $a = 0$. Consequently, (5.27) says $\widetilde{\mathfrak{m}}_n g_j = 0$ a.s., which completes the proof of (ii) $\Longrightarrow$ (iv).

Moreover, (5.26) with $a = 0$ and (5.11) show that a.s., $g_j(x) = 0$ for $\widetilde{\mathfrak{m}}_n$-a.e. x, which is the same as $\widetilde{\mathfrak{m}}_n |g_j| = 0$. Hence, also (ii) $\Longrightarrow$ (vi).

(vi) $\Longrightarrow$ (iv): Trivial.

(iv) $\Longrightarrow$ (iii): Now suppose $\mathfrak{m}_n g_j = 0$ a.s., for every $n \geq 0$. Then $\widetilde{\mathfrak{m}}_n g_j = 0$ a.s., and, using (5.12),

$$\mathfrak{v}_n g_j = \widetilde{\mathfrak{m}}_n g_j - \nu g_j = 0 \quad \text{a.s.} \tag{5.29}$$

Furthermore, g_j is an eigenfunction of $\mathbf{R}$ by (5.11), and thus also an eigenfunction of $B_{0,n} = b_{0,n}(\mathbf{R})$. Hence, (5.29) implies

$$\mathfrak{v}_n B_{0,n}^{-1} g_j = 0 \quad \text{a.s.} \tag{5.30}$$

for every $n \geq 0$. By Lemma 4.2, we have $\mathfrak{v}_n B_{0,n}^{-1} g_j \xrightarrow{\text{a.s.}} \Lambda_{g_j}$. Thus (5.30) implies $\Lambda_{g_j} = 0$ a.s., which by (5.14) yields $\Lambda_j = 0$ a.s. and thus shows (iii).

(iii) $\Longrightarrow$ (ii): Trivial.

(iv) $\Longleftrightarrow$ (v): Obvious by (2.2).

(vi) $\Longleftrightarrow$ (vii): Trivial.

(vi) $\Longleftrightarrow$ (viii): (vi) is equivalent to

$$\mathbb{E}\, \widetilde{\mathfrak{m}}_n |g_j| = 0 \quad \text{for every } n \geq 0. \tag{5.31}$$

By (3.18) and induction, recalling (3.3) and (3.38),

$$\mathbb{E}\, \widetilde{\mathfrak{m}}_n = \widetilde{\mathfrak{m}}_0 B_{0,n} = \widetilde{\mathfrak{m}}_0 b_{0,n}(\mathbf{R}). \tag{5.32}$$

Since $b_{0,n}(\mathbf{R})$ is a polynomial in $\mathbf{R}$ of degree (exactly) n, (5.31) is equivalent to (viii), which shows the equivalence (vi) $\Longleftrightarrow$ (viii).

(vi) $\Longrightarrow$ (ix) when $\mathbf{R}$ is slqc: Theorem 2.10 then applies to all functions in $B(W)$, and in particular to $|g_j|$. Hence, $\widetilde{\mathfrak{m}}_n |g_j| \xrightarrow{\text{a.s.}} \nu |g_j|$. The condition (vi) thus implies $\nu|g_j| = 0$, which is (ix).

(ix) $\Longrightarrow$ (viii) when $\mathfrak{m}_0$ is absolutely continuous w.r.t. ν: By Lemma 5.1 and induction, (ix) implies $\mathbf{R}^n |g_j| = 0$ ν-a.e. for every $n \geq 0$. Our assumption $\mathfrak{m}_0 \ll \nu$ then yields (viii). □

Example 5.2. Let $E = [0,1]$, and let μ be the Lebesgue measure on E. Let $0 < \theta < 1$ and let $R_x^{(1)}$ be the (non-random) replacement kernel given by

$$R_x^{(1)} = \overline{R}_x = \begin{cases} \mu, & x \neq 0 \\ \theta\delta_0 + (1-\theta)\mu, & x = 0. \end{cases} \tag{5.33}$$

We take $W = V = 1$, and it is trivial to verify (B), (H), and (N), with $\nu = \mu$. The operator $\mathbf{R}$ (considered on $B(E)$ as usual) has rank 2 and it is easily seen that $\sigma(\mathbf{R}) = \{0, \theta, 1\}$, with the spectral projections Π_1 and Π_θ both having rank 1 and corresponding eigenvectors 1 and $\mathbf{1}_{\{0\}}$. Hence $\mathbf{R}$ is always slqc on $B(W)$, and small if and only if $\theta < 1/2$; moreover, our parameter θ is as in (2.22).

If we start with $\mathfrak{m}_0 = \mu$, or with δ_x for any $x \neq 0$, then a.s. $\mathfrak{m}_n = \mathfrak{m}_0 + n\mu$, so the MVPP is deterministic. However, if we take $\mathfrak{m}_0 = \delta_0$, then the evolution is different; the MVPP then is essentially a triangular urn of the type considered in e.g. [20], where its asymptotic distribution can be found. (To see this, call colour 0 'white' and lump all other colours in E together as 'non-white'.)

In particular, if $1/2 < \theta < 1$, then Theorem 2.13(3) applies with $D = B(E)$, $p = 1$ and $\lambda_1 = \theta$. Moreover, if we take $f := \mathbf{1}_{\{0\}}$, then Theorem 2.27 applies with $g_1 = f = \mathbf{1}_{\{0\}}$. It follows easily that the limit $\Lambda_1 = 0$ in (2.38) if and only if $\mathfrak{m}_0\{0\} = 0$.

This, admittedly artifical, example shows that one cannot always ignore functions that are ν-a.e. 0; thus some care may be required when considering $\mathbf{R}$ as acting on $L^\infty(E, \nu)$. □

Example 5.3. We may vary Example 5.2 by fixing 3 distinct points $x_0, x_1, x_2 \in [0,1]$ and defining (non-random)

$$R_x^{(1)} = \overline{R}_x = \begin{cases} \frac{1}{2}\delta_{x_1} + \frac{1}{2}\delta_{x_2}, & x = x_0, \\ \theta\delta_x + (1-\theta)\mu, & x = x_1, x_2, \\ \mu, & \text{otherwise}. \end{cases} \tag{5.34}$$

The spectrum is still $\{0, \theta, 1\}$, and the range of the spectral projection Π_θ has dimension 2. Let $\frac{1}{2} < \theta < 1$, so that Theorem 2.13(3) applies. One can easily check that $f := \mathbf{1}_{x_1} - \mathbf{1}_{x_2} \in \Pi_\theta$ and that Theorem 2.27 applies with $g_1 = f$. Using Theorem 2.27(viii), it follows that $\Lambda_1 = 0$ if and only if $\mathfrak{m}_0\{x_0, x_1, x_2\} = 0$. In particular, note that Λ_1 is non-random if $\mathfrak{m}_0 = \delta_{x_0}$; this shows that in Theorem 2.27, it is not enough to assume $\mathfrak{m}_0|g_j| = 0$. □

6. Examples

We consider some examples, in separate subsections.

6.1. **Out-degree distribution in the random recursive tree.** This example is already considered by [19] in the Pólya urn context and in [27] in the MVPP context. The random recursive tree is built recursively as follows: at time 1 the tree has one node, its root, and, at every discrete time-step, we add one node to the tree, and this new node chooses its parent uniformly at random among the nodes that are already in the tree. The out-degree of a node is its number of children. For all $n \geq 1$ and $k \geq 0$, we set $U_k(n)$ the number of vertices of out-degree k in the n-node RRT, and

$$\mathfrak{m}_n := \sum_{k\geq 0} U_k(n)\delta_k. \tag{6.1}$$

(We start this process at time $n = 1$; this is just a matter of notation.) We show that

Proposition 6.1. *If $f : \mathbb{N}_0 \to \mathbb{C}$ satisfies $f(k) = O(r^k)$ for some $r < \sqrt{2}$, then there exists a covariance matrix $\Sigma(f)$ such that, as $n \to +\infty$,*

$$n^{1/2} \sum_{k\geq 0} \Big(\frac{U_k(n)}{n} - 2^{-k-1}\Big) f(k) \xrightarrow{\mathrm{d}} \mathcal{N}\big(0, \Sigma(f)\big). \tag{6.2}$$

We show how to calculate $\Sigma(f)$ at the end of the section, at least in some cases.

Remark 6.2. We compare Proposition 6.1 to the results of [19] and [27]. The results in [19] give an equivalent of Proposition 6.1 but only for functions f with finite support. The results of [27] apply to unbounded functions f as long as they are negligible in front of $x \mapsto 2^{x-\varepsilon}$ for some $\varepsilon > 0$. This class of functions is larger than the one in Proposition 6.1, but [27] proves a.s. convergence of $\frac{1}{n}\sum_{k\geq 0} U_k(n) f(k)$ to $\sum_{k\geq 0} 2^{-k-1} f(k)$ while Proposition 6.1 gives the fluctuations around this almost-sure limit. □

Proof. To prove this proposition, first note that $(\mathfrak{m}_n)_{n\geq 1}$ is an MVPP with $E = \mathbb{N}_0 := \{0, 1, 2, \dots\}$, and deterministic $R^{(n)} = R^{(1)} = \overline{R}$ such that, for all $k \geq 0$,

$$R_k^{(1)} = \delta_{k+1} - \delta_k + \delta_0, \qquad k \geq 0. \tag{6.3}$$

Note that this includes subtracting the drawn ball k (unless $k = 0$). In other words, the operator $\mathbf{R}$ is defined by (2.18) as

$$\mathbf{R}f(k) = \overline{R}_k f = f(k+1) - f(k) + f(0), \qquad k \geq 0. \tag{6.4}$$

Dually,

$$\delta_k \mathbf{R} = \overline{R}_k = \delta_{k+1} - \delta_k + \delta_0, \tag{6.5}$$

and thus, for any complex measure μ on $\mathbb{N}_0$, (with $\mu\{-1\} := 0$)

$$(\mu\mathbf{R})(k) = \mu\{k-1\} - \mu\{k\} + \mathbf{1}_{k=0}(\mu 1). \tag{6.6}$$

The urn is balanced by (6.3), i.e., (B) holds.

We first choose $W = V = 1$. Then (H) holds since $\|R^{(1)}_k\| \le 3$ for every $k \ge 0$, see Remark 2.2. Furthermore, it is easily checked from (6.6) that the probability measure

$$\nu\{k\} = 2^{-k-1}, \qquad k \in \mathbb{N}_0, \tag{6.7}$$

(i.e., a geometric distribution Ge(1/2)) is an eigenvector satisfying $\nu\mathbf{R} = \nu$, and thus (N) holds too.

We next show that $\mathbf{R}$ is a small operator on $B(W) = B(E)$. To do so, we show first that the dual operator $\mathbf{R}^*$ is a small operator on $\mathcal{M}(E)$; recall that $\mathbf{R}^*$ is the operator in (6.5)–(6.6) which we there, as usually, denote by $\mathbf{R}$ (acting on the right).

The space $\mathcal{M}(E)$ of complex measures is naturally identified with ℓ^1; we also identify it with the space

$$\mathcal{A} := \Bigl\{\sum_{k=0}^{\infty} a_k z^k : \sum_{k=0}^{\infty} |a_k| < \infty\Bigr\} \tag{6.8}$$

of analytic functions. (The functions in $\mathcal{A}$ are thus the analytic functions in the unit disc with a Taylor series that is absolutely convergent on the closed unit disc.) The identification is the obvious one, mapping a measure $\mu \in \mathcal{M}(E)$ to $\sum_{k=0}^{\infty} \mu\{k\} z^k$. Note that $\mathcal{A}$ is a Banach algebra under pointwise multiplication. (The norm in $\mathcal{A}$ is inherited from $\mathcal{M}(E) = \ell^1$.)

The operator $\mathbf{R}^*$ acting on $\mathcal{M}(E)$ by (6.5) corresponds to the operator $\hat{\mathbf{R}} : \mathcal{A} \to \mathcal{A}$ given by

$$\hat{\mathbf{R}} z^k = z^{k+1} - z^k + 1. \tag{6.9}$$

This means that, for all $f \in \mathcal{A}$, cf. (6.6),

$$\hat{\mathbf{R}} f(z) = z f(z) - f(z) + f(1) = (z-1) f(z) + f(1). \tag{6.10}$$

We first show that

$$\sigma(\hat{\mathbf{R}}) \subseteq \{\lambda : |\lambda + 1| \le 1\} \cup \{1\}; \tag{6.11}$$

this implies that (QC2) holds. Fix $\lambda \in \mathbb{C}$ such that $|\lambda+1| > 1$ and $\lambda \neq 1$; our aim is to show that $\lambda \notin \sigma(\hat{\mathbf{R}})$, i.e. $\lambda \in \rho(\hat{\mathbf{R}})$. To do so, we fix $g \in \mathcal{A}$ and consider the equation $(\lambda - \hat{\mathbf{R}}) f = g$. By (6.10), the equation can be written

$$(1 + \lambda - z) f(z) - f(1) = g(z). \tag{6.12}$$

In particular, taking $z = 1$ yields

$$(\lambda - 1) f(1) = g(1). \tag{6.13}$$

Then, (6.13) gives $f(1) = g(1)/(\lambda - 1)$, and (6.12) is solved (uniquely) by

$$f(z) = \frac{g(z) + f(1)}{1 + \lambda - z} = \frac{g(z) + g(1)/(\lambda - 1)}{1 + \lambda - z}. \tag{6.14}$$

Furthermore, this solution f belongs to $\mathcal{A}$, since $1/(1+\lambda-z) \in \mathcal{A}$ when $|\lambda+1| > 1$ and $\mathcal{A}$ is a Banach algebra. Hence, $(\lambda - \hat{\mathbf{R}}) f = g$ has a unique solution $f \in \mathcal{A}$ for every $g \in \mathcal{A}$; in other words, $\lambda \in \rho(\hat{\mathbf{R}})$, which concludes the proof of (6.11) and thus of (QC2).

Furthermore, the resolvent $(\lambda - \hat{\mathbf{R}})^{-1} g$ is given by (6.14), and thus, by [5, Equation VII.6.9], the spectral projection Π_1 is given by

$$\Pi_1 g(z) = \frac{1}{2\pi \mathrm{i}} \oint_\Gamma (\lambda - \hat{\mathbf{R}})^{-1} g(z) \,\mathrm{d}\lambda = \frac{1}{2\pi \mathrm{i}} \oint_\Gamma \frac{g(z) + g(1)/(\lambda - 1)}{1 + \lambda - z} \,\mathrm{d}\lambda, \qquad |z| \le 1, \tag{6.15}$$

where Γ is a small circle around 1. (Any circle of radius less that 1 will do.) If $|z| \le 1$, then $1/(1+\lambda-z)$ is an analytic function of λ on and inside Γ, and it follows by the residue theorem that the integral (6.15) equals the residue at $\lambda = 1$, which is $g(1)/(2-z)$. Thus,

$$\Pi_1 g(z) = \frac{g(1)}{2 - z}, \qquad |z| \le 1, \tag{6.16}$$

which together with (6.11) shows that (QC1) holds with the eigenfunction $1/(2-z) = \sum_{k=0}^{\infty} 2^{-k-1}z^k$. This eigenfunction corresponds to ν in (6.7), which shows again that $\nu\mathbf{R} = \nu$.

Therefore, $\hat{\mathbf{R}}$ is an slqc operator on $\mathcal{A}$. Furthermore, we conclude from (6.11) that it is a small operator on $\mathcal{A}$, and thus that $\mathbf{R}^*$ is a small operator on $\mathcal{M}(E)$. By Corollary B.3, with $\mathcal{X} = B(E)$ and $\mathcal{Y} = \mathcal{M}(E)$, this implies that $\mathbf{R}$ is a small operator on $B(E)$.

We have verified the conditions of Theorem 2.13(1), which thus applies and shows asymptotic normality of $\mathfrak{m}_n f$ as in (2.27) for every $f \in B(W) = B(E)$.

We can extend the range of this result by considering other functions W. Fix $r \geq 1$ and take now

$$W(k) = W_r(k) := r^k. \tag{6.17}$$

Thus $V(k) = W(k)^q = r^{qk}$ for some $q > 2$. Recall that (N) requires $\nu V < \infty$. Since ν still is given by (6.7), this is equivalent to $r^q < 2$. Similarly, (6.4) shows that

$$\overline{R}_k V = V(k+1) - V(k) + V(0) = (r^q - 1)V(k) + 1, \tag{6.18}$$

and thus (H)(i) holds if and only if $r^q < 2$. It is easily seen that (H)(ii) holds for every $r \geq 1$. Furthermore, the urn starts with the composition δ_0, and thus (H)(iii) is trivial. Hence, (H) and (N) both hold if and only if $r^q < 2$. Since (B) holds regardless of W, we conclude that

$$\text{(B), (H), and (N) hold for some } q > 2 \iff r < \sqrt{2}. \tag{6.19}$$

We now have to find the spectral gap of $\mathbf{R}$ as an operator on $B(W_r)$. We argue as in the case $r = 1$ above, and begin by noting that $\mathcal{M}(W_r) = \bigl\{\mu : \sum_0^\infty |\mu\{k\}| r^k < \infty\bigr\}$ is a norm-determining subspace of $B(W_r)^*$. Moreover, $\mathcal{M}(W_r)$ may be identified with the space

$$\mathcal{A}_r := \Bigl\{\sum_{k=0}^{\infty} a_k z^k : \sum_{k=0}^{\infty} |a_k| r^k < \infty\Bigr\} \tag{6.20}$$

of analytic functions. The functions in $\mathcal{A}_r$ are continuous in the closed disc $\{z : |z| \leq r\}$ and analytic in its interior. $\mathcal{A}_r$ is, as $\mathcal{A} = \mathcal{A}_1$ studied above, a Banach algebra under pointwise multiplication.

As in the case $r = 1$, the operator $\mathbf{R}^*$ on $\mathcal{M}(W_r)$ corresponds to an operator $\hat{\mathbf{R}}$ on $\mathcal{A}_r$ given by (6.9) and (6.10). The argument above then shows that $\lambda \in \rho(\hat{\mathbf{R}})$ provided $\lambda \neq 1$ and $1/(1+\lambda-z) \in \mathcal{A}_r$, i.e., if $\lambda \neq 1$ and $|1+\lambda| > r$. Consequently, (6.11) is replaced by

$$\sigma(\mathbf{R}^*) \subseteq \{\lambda : |\lambda+1| \leq r\} \cup \{1\}. \tag{6.21}$$

Hence, on $B(W_r)$, using Lemma B.2(ii) and with $\sigma(\mathbf{R}^*)\widehat{\ }$ defined in Definition B.1,

$$\sigma(\mathbf{R}) \subseteq \sigma(\mathbf{R}^*)\widehat{\ } \subseteq \{\lambda : |\lambda+1| \leq r\} \cup \{1\}. \tag{6.22}$$

In particular,

$$\theta_{B(W_r)} \leq r - 1. \tag{6.23}$$

We have seen above that we have to take $r < \sqrt{2}$ in order to have (H) and (N), and (6.23) shows that in this case $\theta < 1/2$ follows. Consequently, if $r < \sqrt{2}$, then the asymptotic normality (2.27) extends to all $f \in B(W_r)$, i.e., all f such that $f(k) = O(r^k)$. This completes the proof. □

Remark 6.3. It is easy to see that we have equality in (6.22)–(6.23). In fact, we know that 1 is an eigenvalue by (2.19). Moreover, if $|1+\lambda| \leq r$ and $\lambda \neq 0$, then $f(k) := (1+\lambda)^k + 1/(\lambda-1)$ satisfies $f \in B(W_r)$ and $\mathbf{R}f = \lambda f$ by (6.4), see also (6.25) below, so λ is an eigenvalue of $\mathbf{R}$ and thus $\lambda \in \sigma(\mathbf{R})$. Hence, we have equality in (6.22)–(6.23) too. (For $\lambda = 0$ and $r > 1$, $f(k) = k - 1$ is an eigenfunction, but this case follows also because $\sigma(\mathbf{R})$ is closed.) Consequently, $\mathbf{R}$ is a small operator in $B(W_r) \iff r < 3/2$. □

In the rest of this subsection, we show how to calculate the asymptotic covariance matrix $\Sigma(f)$ in Proposition 6.1 for the following functions f: Fix $r \in [1, \sqrt{2})$ and let, for a complex a with $|a| \le r$,

$$f_a(k) := a^k. \tag{6.24}$$

Then $f_a \in B(W_r)$, $\nu f_a = \sum_{k=0}^{\infty} 2^{-k-1} a^k = 1/(2-a)$, and, by (6.4),

$$\mathbf{R}(f_a - \nu f_a) = a^{k+1} - a^k + 1 - \frac{1}{2-a} = a^{k+1} - a^k + \frac{1-a}{2-a} = (a-1)(f_a - \nu f_a). \tag{6.25}$$

In other words, provided $a \neq 1$ (so the function does not vanish), $\tilde{f}_a := f_a - \nu f_a \in B(W_r)$ is an eigenfunction of $\mathbf{R}$ with eigenvalue $a-1$. This makes it easy to compute asymptotic variances and covariances in Theorem 2.13 for the functions f_a.

Let a and b be complex numbers with $|a|, |b| \le r$. First, note that by (2.36), since $R_x^{(1)} = \overline{R}_x$ is deterministic,

$$\begin{aligned}\widetilde{\mathbf{B}}(\tilde{f}_a, \tilde{f}_b) &= \overline{R}_{\bullet}\tilde{f}_a \cdot \overline{R}_{\bullet}\tilde{f}_b = (\mathbf{R}\tilde{f}_a)\cdot(\mathbf{R}\tilde{f}_b) = (a-1)(b-1)\tilde{f}_a\tilde{f}_b \\ &= (a-1)(b-1)\big(f_{ab} - (\nu f_a) f_b - (\nu f_b) f_a + (\nu f_a)(\nu f_b)\big).\end{aligned} \tag{6.26}$$

Hence,

$$\begin{aligned}\nu\widetilde{\mathbf{B}}(\tilde{f}_a, \tilde{f}_b) &= (a-1)(b-1)\big(\nu f_{ab} - (\nu f_a)(\nu f_b)\big) \\ &= (a-1)(b-1)\Big(\frac{1}{2-ab} - \frac{1}{(2-a)(2-b)}\Big) = \frac{2(a-1)^2(b-1)^2}{(2-ab)(2-a)(2-b)}\end{aligned} \tag{6.27}$$

and thus, recalling again (6.25),

$$\begin{aligned}\int_0^\infty \nu\widetilde{\mathbf{B}}\big(e^{s\mathbf{R}}\tilde{f}_a, e^{s\mathbf{R}}\tilde{f}_b\big)e^{-s}\,ds &= \int_0^\infty \nu\widetilde{\mathbf{B}}\big(e^{s(a-1)}\tilde{f}_a, e^{s(b-1)}\tilde{f}_b\big)e^{-s}\,ds \\ &= \int_0^\infty \nu\widetilde{\mathbf{B}}\big(\tilde{f}_a, \tilde{f}_b\big)e^{-(3-a-b)s}\,ds \\ &= \frac{2(a-1)^2(b-1)^2}{(3-a-b)(2-ab)(2-a)(2-b)}.\end{aligned} \tag{6.28}$$

Taking $b = a$ in (6.28) gives $\chi(f_a)$ in (2.28), and taking $b = \bar{a}$ gives $\sigma^2(f_a)$ in (2.29). (See Remark 2.15). In particular, for a real with $|a| < \sqrt{2}$, Theorem 2.13 shows (see Remark 2.12) that

$$n^{-1/2}\Big(\sum_{k=0}^{\infty} U_k(n)a^k - \frac{n}{2-a}\Big) = n^{1/2}\big(\widetilde{\mathfrak{m}}_n f_a - \nu f_a\big) \xrightarrow{\text{d}} \mathcal{N}\big(0, \sigma^2(f_a)\big), \tag{6.29}$$

with

$$\sigma^2(f_a) = \frac{2(a-1)^4}{(3-2a)(2-a^2)(2-a)^2}. \tag{6.30}$$

More generally, we have joint convergence for several (real or complex) a, with asymptotic covariances easily found from (6.28).

Remark 6.4. It follows that the asymptotic variances and covariances of $n^{-1/2}U_k(n)$ can be obtained as Taylor coefficients of the bivariate rational function in (6.28); this was earlier shown in [19] by related calculations using urns with finitely many colours. □

Remark 6.5. Moreover, using Fourier analysis, any function f in Proposition 6.1 may be expressed as an integral of functions f_a: for any $\rho \in (r, \sqrt{2})$,

$$f = \frac{1}{2\pi}\int_0^{2\pi} \widehat{f}\big(\rho^{-1}e^{-it}\big) f_{\rho e^{it}}\,dt \tag{6.31}$$

where $\widehat{f}(z) := \sum_{k=0}^{\infty} f(k)z^k$. By substituting (6.31) in (2.28) and (2.29), and using (6.28), one can obtain integral formulas for $\chi(f)$ and $\sigma^2(f)$, and thus for $\Sigma(f)$. The result is rather complicated, however, and we leave the details to the reader. □

Remark 6.6. The asymptotic variance in (6.30) diverges as $a \nearrow \sqrt{2}$, and thus the result cannot be extended (in this form at least) to $a \geq \sqrt{2}$. Hence, the condition $r < \sqrt{2}$ in Proposition 6.1 and the argument above is not just a technical condition required by our proofs; it is essential for (6.29)–(6.30), which strongly suggests that it is necessary in Proposition 6.1 too. This also shows that the technical conditions (H) and (N) are more or less best possible; in particular, it is not enough to take $q < 2$ in (H). □

We do not know what happens for functions f that grow faster than allowed in Proposition 6.1. In particular, the following case seems interesting.

Problem 6.7. What is the asymptotic distribution of $\sum_{k=0}^{\infty} U_k(n)a^k$ for $a \geq \sqrt{2}$?

Is there any difference between the cases $a < 3/2$ and $a > 3/2$? (Recall that $\mathbf{R}$ is a small operator in $B(W_a)$ for $a < 3/2$, but not for larger a.)

6.2. **The heat kernel on the square.** Imagine some flowers planted in a closed square room: we start with one flower in the room (say at the centre of the room). Each flower blooms at exponential rate, independently from the others, and when a flower blooms, it sends one seed in the air, which travels in the air according to a Brownian motion reflected at the walls for a unit-time, then fall onto the ground and instantly becomes a new flower. We assume that the rate of blooming is so small that we can imagine that the seeds perform their unit-one Brownian motions instantly. We set τ_n to be the instant of the n-th bloom ($\tau_0 := 0$), and ξ_n to be the position of the n-th flower in $[0, \ell]^2$ ($\xi_0 = (\ell/2, \ell/2)$). We are interested in the long-term behaviour of the distribution of flowers in the room:

$$\Xi_n = \sum_{i\geq 0} \delta_{\xi_i}. \tag{6.32}$$

It is expected that Ξ_n/n converges to the uniform distribution on the square, and this is indeed confirmed by Theorem 2.10(ii); Theorem 2.13 allows to study the fluctuations around this limit. This yields the following.

Proposition 6.8. *For all bounded measurable functions $f : [0, \ell]^2 \mapsto \mathbb{R}$,*

$$\frac{1}{n}\Xi_n f = \frac{1}{n}\sum_{i=0}^{n} f(\xi_i) \to \frac{1}{\ell^2}\int_{[0,\ell]^2} f(x)\,\mathrm{d}x, \qquad \textit{almost surely when } n \to +\infty. \tag{6.33}$$

For all $m, p \in \mathbb{N}_0^2$, set

$$\lambda_{m,p} := \exp\left(-\frac{\pi^2(m^2 + p^2)}{2\ell^2}\right), \tag{6.34}$$

and

$$\varphi_{m,p}(x, y) := \cos\left(\frac{\pi m x}{\ell}\right)\cos\left(\frac{\pi p y}{\ell}\right). \tag{6.35}$$

Also, set $I(\ell) := \{(m, p) \in \mathbb{N}_0^2 \colon \lambda_{m,p} < 1/2\}$ and let D be the closed linear span in $B([0, \ell]^2)$ of 1 and $\{\varphi_{m,p} \colon (m, p) \in I(\ell)\}$. Similarly, set $J(\ell) := \{(m, p) \in \mathbb{N}_0^2 \colon \lambda_{m,p} \leq 1/2\}$ and let D' be the closed linear span of 1 and $\{\varphi_{m,p} \colon (m, p) \in J(\ell)\}$.

(i) *For every function $f \in D$, there exists a covariance matrix $\Sigma(f)$ such that*

$$n^{1/2}\left(\frac{1}{n}\sum_{i=0}^{n} f(\xi_i) - \frac{1}{\ell^2}\int_{[0,\ell]^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y\right) \to \mathcal{N}\big(0, \Sigma(f)\big), \tag{6.36}$$

in distribution as $n \to +\infty$.

(ii) *If $\frac{2\log 2}{\pi^2}\ell^2 \in \{m^2 + p^2 \colon (m, p) \in \mathbb{N}_0^2\}$, so $J(\ell) \neq I(\ell)$, then for every function $f \in D'$, there exists a covariance matrix $\Sigma(f)$ such that*

$$\frac{n^{1/2}}{(\log n)^{1/2}}\left(\frac{1}{n}\sum_{i=0}^{n} f(\xi_i) - \frac{1}{\ell^2}\int_{[0,\ell]^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y\right) \to \mathcal{N}\big(0, \Sigma(f)\big), \tag{6.37}$$

in distribution as $n \to +\infty$.

(iii) *If $\ell > \pi/\sqrt{2\log 2}$, then for every function $f \in B([0,\ell]^2)$, there exists a random variable $W(f)$ such that*

$$n^{1-\exp(-\pi^2/2\ell^2)} \left(\frac{1}{n} \sum_{i=0}^{n} f(\xi_i) - \frac{1}{\ell^2} \int_{[0,\ell]^2} f(x,y) \, \mathrm{d}x \, \mathrm{d}y \right) \to W(f), \tag{6.38}$$

almost surely and in L^2 when $n \to +\infty$.

Remark 6.9. If $\ell < \pi/\sqrt{2\log 2}$, then $D = B([0,\ell]^2)$, and then (i) applies to all bounded f. Similarly, if $\ell = \pi/\sqrt{2\log 2}$, then $D' = B([0,\ell]^2)$ and (ii) applies to all bounded f. □

Proof. First note that Ξ_n is an MVPP with colour space $E = [0,\ell]^2$, initial composition $\delta_{(\ell/2,\ell/2)}$, and random replacement kernel

$$R_x^{(1)} = \delta_{B_1^{(x)}}, \tag{6.39}$$

where $B = (B_t)_{t\geq 0}$ is the standard Brownian motion on the square of side-length ℓ started at $B_0^{(x)} = x$ and reflected at the boundary. Note that $R_x^{(1)}$ is a positive measure. We have

$$\overline{R}_x = \mathcal{L}\big(B_1^{(x)}\big), \tag{6.40}$$

the distribution of the reflected Brownian motion. Hence, for any probability measure μ on E,

$$\mu\mathbf{R} = \mathcal{L}\big(B_1^{\mu}\big), \tag{6.41}$$

the distribution of the reflecting Brownian motion at time 1 when started according to μ.

This MVPP satisfies Assumption (B). We choose $W = V = 1$, and then (H) holds by Remark 2.2. Furthermore, (N) holds because the uniform distribution ν on $[0,\ell]^2$ is invariant for the reflected Brownian motion and thus satisfies $\nu\mathbf{R} = \nu$ by (6.41).

The kernel $\overline{R}_x$ in (6.40) of $\mathbf{R}$ is known as the *heat kernel with Neumann boundary conditions*. Its eigenvalues and eigenfunctions are well known, and can be found e.g. as follows. (We give a sketch, omitting the standard details.) First, since the kernel is absolutely continuous, and depends continuously on x, it is easily seen that it does not matter whether we consider $\mathbf{R}$ as an operator on $B(E)$ or $L^\infty(E)$. (See Lemma B.5, with $\mathcal{N}$ the space of bounded functions that are 0 a.e.) Furthermore, the density of $\overline{R}_x$ is bounded, uniformly in x, and it follows that $\mathbf{R}$ maps $L^2(E)$ into $L^\infty(E)$. Hence, Lemma B.4 shows that eigenvalues and other spectral properties are the same in $L^\infty(E)$ and in $L^2(E)$ (except possibly at 0, which is not important for us). Finally, we regard $L^2(E) = L^2([0,\ell]^2)$ as the subspace of $L^2([-\ell,\ell]^2)$ consisting of functions that are even in each variable, and then extend these functions periodically to $\mathbb{R}^2$. We then can replace the reflecting Brownian motion by ordinary Brownian motion on $\mathbb{R}^2$, and it follows that the functions $\varphi_{m,p}$ in (6.35) form a complete orthogonal set of eigenfunctions in $L^2(E)$, with corresponding eigenvalues $\lambda_{m,p}$ given by (6.34). (In this example, $\mathbf{R}$ is a self-adjoint operator on L^2, which makes the spectral theory in L^2 particularly simple.)

Since $\lambda_{m,p} \to 0$ as $m + p \to \infty$, it follows that

$$\sigma(\mathbf{R}) = \{\lambda_{m,p} : m, p \in \mathbb{N}_0\} \cup \{0\}, \tag{6.42}$$

in $L^2(E)$, and by Lemma B.4 as indicated above, also in $L^\infty(E)$ and in $B(E)$.

The eigenvalue 1 is obtained only for $m = p = 0$, and thus it follows from (6.42) that $\mathbf{R}$ is slqc. Moreover, the second largest eigenvalue is $\lambda_{1,0} = \lambda_{0,1} = \exp\big(-\pi^2/(2\ell^2)\big)$, and thus $\mathbf{R}$ is small if and only if $\pi^2/(2\ell^2) > \log 2$, i.e., if $\ell < \pi/\sqrt{2\log 2}$.

The almost sure convergence in (6.33) is thus a direct consequence of Theorem 2.10(i), which also gives an (upper) estimate of the rate.

Next, we show that

$$\sigma(\mathbf{R}_D) = \{\lambda_{m,p} : (m,p) \in I(\ell) \cup \{(0,0)\}\} = \{\lambda_{m,p} : \lambda_{m,p} < 1/2\} \cup \{1\}. \tag{6.43}$$

To see this, we first note that if $\widehat{D}$ is the closure of D in $L^2(E)$, i.e., the closed linear span in $L^2(E)$ of 1 and $\{\varphi_{m,p}\colon (m,p)\in I(\ell)\}$, then $\sigma(\mathbf{R}_{\widehat{D}})$ is given by (6.43), since the functions $\varphi_{m,p}$ are orthogonal eigenfunctions. Then, (6.43) follows by Lemma B.4, because $\mathbf{R}:\widehat{D}\to D$.

It follows from (6.43) that $\mathbf{R}_D$ is a small operator, and thus (i) is a direct consequence of Theorem 2.13(1).

Similarly, by the same argument,

$$\sigma(\mathbf{R}_D)=\{\lambda_{m,p}:(m,p)\in J(\ell)\cup\{(0,0)\}\}=\{\lambda_{m,p}:\lambda_{m,p}\le 1/2\}\cup\{1\}. \tag{6.44}$$

and (ii) follows from Theorem 2.13(2), with $p=1$, $\lambda_1=\frac12$, and $\kappa=\kappa_1=1$.

Finally, (iii) follows from Theorem 2.13(3), with $p=1$, $\lambda_1=\mathrm{e}^{-\pi^2/(2\ell^2)}$, and $\kappa=\kappa_1=1$. □

Remark 6.10. The covariance matrices of the limits in (6.36) and (6.37) can easily be computed from the formulas in Theorem 2.13 and a Fourier expansion of f into the functions $\varphi_{m,p}$; we leave the details to the reader. □

We can use Theorems 2.25–2.27 to see whether the limit distributions in Proposition 6.8 are degenerate. Note that if $\lambda\neq 0$, then Π_λ is a projection onto a finite-dimensional space spanned by some $\varphi_{m,p}$; these are all continuous, and thus $\Pi_\lambda f$ is continuous for any $f\in B(E)$.

First, for (i), it is easily seen from Theorem 2.25 that the limit in (6.36) is degenerate only if $f=c$ a.e. for some constant c.

Secondly, for (ii), Theorem 2.26 (with $\kappa=1$ and $p=1$) shows that the limit is degenerate if and only if $\Pi_{1/2}f=0$ a.e.; since $\Pi_{1/2}f$ is continuous, this holds if and only if $\Pi_{1/2}f=0$. It is easily seen that this holds if and only if $f\in D$ (and thus (i) applies, and gives a more precise result).

Similarly, for (iii), Theorem 2.27 shows that the limit is degenerate if and only if $\Pi_{\lambda_1}f=0$, where $\lambda_1=e^{-\pi^2/(2\ell^2)}$. Assume this. The next largest eigenvalue of $\mathbf{R}$ is $\lambda_2=e^{-\pi^2/\ell^2}$. Hence, if $\lambda_2\le 1/2$, we can apply (i) or (ii) to f. If $\lambda_2>1/2$, we may instead apply Theorem 2.13(3) to the subspace $D_1:=(\mathbf{I}-\Pi_{\lambda_1})B(E)$; note that T is slqc in D_1 and $\theta_{D_1}=\lambda_2=e^{-\pi^2/\ell^2}$.

Remark 6.11. In this example, the generalized eigenspaces Π_λ ($\lambda\neq 0$) are all spanned by eigenvectors. Hence, $\kappa=1$ in Theorem 2.13, regardless of the multiplicities of the eigenvalues. The multiplicities show up when considering joint convergence of several f, as discussed in Remark 2.28. In fact, in Proposition 6.8(iii), the dominating eigenvalue $\lambda_{1,0}=\lambda_{0,1}$ has multiplicity 2, and thus there is a two-dimensional space of limits.

In Proposition 6.8(ii), the dimension of the space of limits equals the multiplicity of the eigenvalue $1/2$, which equals the number of solutions to $m^2+p^2=N:=(2\log 2)\ell^2/\pi^2$. A formula for the number of such solutions is well known (and was stated already by Gauss), see [14, Theorem 278 and Notes p. 243], as well as a criterion for the existence of any solutions at all (so $D'\neq D$) [14, Theorem 366]. □

Remark 6.12. We could replace $[0,\ell]^2$ by any finite measure space (E,μ) and the Brownian motion $B_1^{(x)}$ by jumps according to any transition kernel $P(x,\mathrm{d}y)$ on E that has a density with respect to μ that is bounded (or, more generally, in $L^2(\mu)$), uniformly in $x\in E$. The operator $\mathbf{R}$ then maps $L^2(\mu)\to B(E)$. Moreover, $\mathbf{R}$ is a Hilbert–Schmidt integral operator on (E,μ), and thus $\mathbf{R}$ is a compact operator on $L^2(\mu)$. By the spectral theorm for compact operators, [5, Theorem VII.7.1], the spectrum $\sigma(\mathbf{R})$ can be written as $\{\lambda_i\}_{i=1}^N\cup\{0\}$ for some $N\le\infty$ and eigenvalues $\lambda_i\neq 0$; either $N<\infty$ or $\lambda_i\to 0$ as $i\to\infty$. Furthermore, $\Pi_{\lambda_i}(L^2(E.\mu))$ has finite dimension for every λ_i. $\mathbf{R}$ is a bounded operator also on $L^\infty(E)$ and $B(E)$, and by Lemmas B.4 and B.5, the spectrum of T is the same for these spaces as for $L^2(E,\mu)$.

The function 1 is an eigenfunction with eigenvalue 1, so $1\in\sigma(T)$, and $|\lambda_i|\le 1$ for all i since $\|\mathbf{R}\|_{B(E)}=1$. In particular, $\mathbf{R}$ is slqc provided $\Pi_1(\mathbf{R})$ does not contain any non-constant function. Assuming the latter property, we thus obtain the same type of behaviour as in Proposition 6.8.

The main advantage of choosing the Brownian motion on $E = [0, \ell]^2$ is that its spectral decomposition is explicitly known and very simple. (That the operator is self-adjoint on L^2 helps but is not essential.) Other examples for which the spectral decomposition is fully known are the reflected Brownian motion on the rectangle, on the isosceles triangle (see e.g. [16, Chapter 5]) or on the annulus (see [25] and [11] for surveys on eigenfunctions and eigenvalues of the heat kernel). □

6.3. **A branching random walk.** The following branching random walk is studied in [17]. Let G be a compact group, and let $(Y_n)_1^\infty$ be an i.i.d. sequence of random variables in G with some distribution $\mu \in \mathcal{P}(G)$. Let $X_0 \in G$ be given. (In [17], X_0 may be random. We assume here that X_0 is non-random; otherwise we may condition on X_0, cf. Remark 2.1.) For $n \geq 1$, let I_n be uniformly distributed on $\{0, \dots, n-1\}$ and assume that all I_n and Y_m are independent. Then define $X_n \in G$ inductively by

$$X_n := X_{I_n} Y_n, \qquad n \geq 1. \tag{6.45}$$

In other words, for each n, we first choose a parent uniformly among $X_0, \dots, X_{n-1}$, and then let X_n be a daughter with a random displacement Y_n from its parent.

This process can be regarded as a MVPP with colour space $E = G$ by defining

$$\mathfrak{m}_n := \sum_{i=0}^{n} \delta_{X_i}. \tag{6.46}$$

The construction of X_n in (6.45) then means that $(\mathfrak{m}_n)_n$ is a MVPP with replacements given by

$$R_x^{(1)} = \delta_{xY_1}, \qquad x \in G. \tag{6.47}$$

We choose $W = V = 1$, and let ν be the normalized Haar measure. The conditions (B), (H), and (N) are easily verified. We have

$$\overline{R}_x = \mathcal{L}(xY_1), \tag{6.48}$$

which is μ left translated by x. Hence, $\mathbf{R}$ acts on functions by convolution $\mathbf{R}f = f * \check{\mu}$, where $\check{\mu}$ is the distribution of Y^{-1}.

The results in [17] are about asymptotic normality, under certain conditions, of the sums

$$S_n(f) := \sum_{i=0}^{n} f(X_i) = \mathfrak{m}_n f \tag{6.49}$$

for suitable functions f. (The proof uses the method of moments.)

Consider for simplicity the case when G is commutative. (The case of non-commutative G is similar but more technical and requires study of the irreducible representations of G; see [17].) Let $\widehat{G}$ be the dual group, consisting of all characters on G (i.e., continuous homomorphisms $G \to \{z \in \mathbb{C} : |z| = 1\}$), and define the Fourier transform of μ by

$$\widehat{\mu}(\gamma) := \int_G \gamma(g) \,\mathrm{d}\mu(g) = \mathbb{E}\,\gamma(Y_1), \qquad \gamma \in \widehat{G}. \tag{6.50}$$

Then, every character γ is an eigenfunction of $\mathbf{R}$, with

$$\mathbf{R}\gamma = \widehat{\mu}(\gamma)\gamma. \tag{6.51}$$

Hence, on the Hilbert space $L^2(G)$, $\mathbf{R}$ has an ON basis of eigenfunctions, and

$$\sigma(\mathbf{R}) = \{\widehat{\mu}(\gamma) : \gamma \in \widehat{G}\}. \tag{6.52}$$

If we assume (as in [17]) that μ is not supported on any proper closed subgroup of G, then $\widehat{\mu}(\gamma) \neq 1$ and thus $\operatorname{Re}\widehat{\mu}(\gamma) < 1$ for every $\gamma \neq 1$. If we further assume, for example, that μ is absolutely continuous w.r.t. the Haar measure ν, then $\widehat{\mu} \in c_0(\widehat{G})$ by (a general version of) the Riemann–Lebesgue lemma, and it follows that $\mathbf{R}$ is slqc on $L^2(G)$. Moreover, if the density $\mathrm{d}\mu/\mathrm{d}\nu$ of μ is in $L^2(G)$, then $\mathbf{R} : L^2 \to B(G)$, and it follows from Lemmas B.4 and B.5 that $\mathbf{R}$ is slqc also on $L^\infty(G)$ and on $B(G)$.

Theorem 2.13 then applies and yields asymtotic normality of $S_n(f)$ if

$$\theta := \sup\{\operatorname{Re}\widehat{\mu}(\gamma) : \gamma \neq 1\} \leq 1/2; \tag{6.53}$$

this is essentially [17, Theorems 3.1 and 3.2], although the technical conditions there on f and μ are somewhat different from ours. (They neither imply or are implied by our conditions here; an example where Theorem 2.13 applies but not [17] is when $\mathrm{d}\mu/\mathrm{d}\nu \in L^2(G) \setminus L^\infty(G)$, and $f \in B(G) \setminus C(G)$.) Moreover, if $1/2 < \theta < 1$, then Theorem 2.13(3) applies, and extends the brief comments given in [17] for that case.

Remark 6.13. [17] considers also a generalization to compact homogeneous spaces; this is treated by constructing a branching random walk as above on a compact group G, and then considering the projection to G/H for a closed subgroup H of G. (This assumes that the distribution μ is invariant under left or right multiplication by elements of H.) The space $B(G/H)$ can be identified with a subspace of $B(G)$, and thus Theorems 2.10 and 2.13 can be applied in this setting too. □

Remark 6.14. This example is closely related to the one in Section 6.2. In fact, the latter example can, by identifying $[-\ell, \ell]^2$ with the group $\mathbb{T}^2$, be treated as a branching random walk as above on the group $G = \mathbb{T}^2$, but considering only the subspace of bounded functions that are even in each coordinate. □

6.4. **Reinforced process on a countable state space.** In this section, we consider a reinforced process which is a particular case of balanced Pólya urn on a countable state space. Let $(X_n)_{n\in\mathbb{Z}_+}$ be an irreducible Markov chain evolving in a countable state space E, and denote by $\mathbb{P}_x$ and $\mathbb{E}_x$ the law and expectation of the process starting from $X_0 = x \in E$. Similarly, if $\nu \in \mathcal{P}(E)$, we use $\mathbb{P}_\nu$ and $\mathbb{E}_\nu$ for the Markov chain started with a random $X_0 \sim \nu$. We assume that X admits a Lyapunov type function: there exist a function $V : E \to [1, +\infty)$ such that $\{x \in E : V(x) \leq A\}$ is finite for every $A < \infty$, and for some constants $\lambda \in (0, 1)$ and $C < \infty$,

$$\mathbb{E}_x\big[V(X_1)\big] \leq \lambda V(x) + C \qquad \text{for all } x \in E. \tag{6.54}$$

We fix $T \in \{2, 3, \dots\}$ and consider the reinforced process $Z = (Z_n)_{n\geq 0}$ constructed as follows: $Z_0 = z_0 \in E$ is fixed and Z evolves according to the dynamic of X up to time $T - 1$. At time T, it jumps to a random position distributed according to its empirical occupation measure $\frac{1}{T}\sum_{i=0}^{T-1}\delta_{Z_i}$; in other words, the process jumps back to its position at a uniformly random earlier time $i \in [0, T)$. Then Z evolves according to the dynamic of X up to time $2T - 1$ and, at time $2T$, it jumps to a random position distributed according to its current empirical occupation measure, and so on. (The process thus jumps back to a random earlier position at times kT, $k \in \mathbb{N}$.)

Let $\boldsymbol{\mu}_n := \frac{1}{n+1}\sum_{i=0}^{n}\delta_{Z_i}$ denote the empirical occupation measure of Z at time n; i.e.

$$\boldsymbol{\mu}_n f = \frac{1}{n+1}\sum_{i=0}^{n} f(Z_i). \tag{6.55}$$

We show that $\boldsymbol{\mu}_n$ converges almost surely (and in a weak L^2 sense) to the unique invariant distribution of X, and that, at least if T is large enough, $\boldsymbol{\mu}_n$ satisfies a central limit theorem.

Proposition 6.15. *The Markov chain X has a unique invariant distribution ν. Moreover:*

(a) *For any $q > 2$, there exists $\delta = \delta(q) > 0$ such that, for every $f \in B(V^{1/q})$,*

$$\mathbb{E}\,|\boldsymbol{\mu}_n f - \nu f|^2 = O\big(n^{-2\delta}\big). \tag{6.56}$$

and

$$n^\delta\,|\boldsymbol{\mu}_n f - \nu f| \xrightarrow[n\to+\infty]{a.s.} 0. \tag{6.57}$$

(b) *If in addition* $\big(\frac{1}{T}\frac{1-\lambda^T}{1-\lambda}\big)^{1/q} < 1/2$, *then, for any* $f \in B(V^{1/q})$, *one of the conclusions* (1), (2) *or* (3) *of Theorem 2.13 holds with:*
- $(n/T)^{1/2}\,(\boldsymbol{\mu}_n f - \nu f)$ *instead of* $n^{1/2}(\widetilde{\mathfrak{m}}_n f - \nu f)$ *in* (1),
- $\frac{(n/T)^{1/2}}{(\log n)^{\kappa-1/2}}\,(\boldsymbol{\mu}_n f - \nu f)$ *instead of* $\frac{n^{1/2}}{(\log n)^{\kappa-1/2}}(\widetilde{\mathfrak{m}}_n f - \nu f)$ *in* (2),
- $\boldsymbol{\mu}_n f$ *instead of* $\widetilde{\mathfrak{m}}_n f$ *and* $\Lambda'_j := T^{1-\lambda_j}\Lambda_j$ *instead of* Λ_j *in* (2.38).

(c) *There exists* $T_0 = T_0(q) \geq 2$ *such that, for any* $T \geq T_0$, *conclusion* (1) *of Theorem 2.13 holds for all* $f \in B(V^{1/q})$.

The proof uses the following lemma, which we prove at the end of this subsection. We use the notations $r(\mathbf{R})$ and $r_e(\mathbf{R})$ for the spectral radius and essential spectral radius of the operator $\mathbf{R}$; see Definition B.6.

Lemma 6.16. *Let* $\mathbf{R}$ *be the operator given by* (2.18) *for some probability kernel* $\overline{R}$ *from* E *to* E, *and let* $V : E \to [1, +\infty)$ *be a function such that* $\{x \in E : V(x) \leq A\}$ *is finite for every* $A < \infty$. *If there exist* $\vartheta < 1$ *and* $C < \infty$ *such that*

$$\mathbf{R}V \leq \vartheta V + C, \tag{6.58}$$

then, for every $q > 1$, $\mathbf{R}$ *acts as a bounded operator on* $B(V^{1/q})$ *with spectral radius* $r(\mathbf{R}) = 1$ *and essential spectral radius* $r_e(\mathbf{R}) \leq \vartheta^{1/q}$.

In particular, $\mathbf{R}$ then is quasi-compact, see Remark B.8.

Proof of Proposition 6.15. We observe that the sequence

$$\mathfrak{m}_n := \frac{1}{T}\sum_{i=0}^{(n+1)T-1}\delta_{Z_i} \tag{6.59}$$

is an MVPP on the set E with (random) initial measure

$$\mathfrak{m}_0 = \frac{1}{T}\sum_{i=0}^{T-1}\delta_{Z_i} \tag{6.60}$$

and replacement kernel

$$R_x^{(1)} \overset{\mathrm{d}}{=} \frac{1}{T}\sum_{i=0}^{T-1}\delta_{X_i}, \quad \text{where } (X_i)_{i\geq 0} \text{ has law } \mathbb{P}_x. \tag{6.61}$$

We start by proving that $\mathfrak{m}$ **satisfies assumptions (B), (H) and (N).** Assumption (B) holds true since

$$R_x^{(1)}(E) \overset{\mathrm{d}}{=} \frac{1}{T}\sum_{i=0}^{T-1}\delta_{X_i}(E) = 1. \tag{6.62}$$

We now show that Assumption (H) holds with $W := V^{1/q}$ and

$$\vartheta := \frac{1}{T}\sum_{i=0}^{T-1}\lambda^i = \frac{1}{T}\,\frac{1-\lambda^T}{1-\lambda} \in (0,1). \tag{6.63}$$

Note that $\vartheta \in (0,1)$ since $\lambda \in (0,1)$ and $T \geq 2$.

For (H)(i), we obtain from (6.54) used iteratively that, for all $x \in E$ and all $n \geq 0$,

$$\mathbb{E}_x\, V(X_{n+1}) \leq \lambda\,\mathbb{E}_x\, V(X_n) + C \leq \lambda^{n+1}V(x) + C_1, \tag{6.64}$$

where $C_1 := \sum_{i=0}^{\infty}\lambda^i C < +\infty$. Hence

$$\overline{R}_x V = \frac{1}{T}\sum_{i=0}^{T-1}\mathbb{E}_x\big[V(X_i)\big] \leq \frac{1}{T}\sum_{i=0}^{T-1}(\lambda^i V(x) + C_1) = \vartheta V(x) + C_1. \tag{6.65}$$

This proves (H)(i).

(H)(ii) then follows by Remark 2.3, since $R_x^{(1)} \geq 0$ a.s.

For (H)(iii), we simply observe that $\mathfrak{m}_0 V = \sum_{i=0}^{T-1} V(Z_i) < +\infty$. This concludes the proof that Assumption (H) holds true.

We now show that Assumption (N) holds. Recall (6.54) and note that it follows that the set $\{x \in E : \mathbb{E}_x\big[V(X_1)\big] > V(x) - 1\}$ is finite. Hence, by [6, Theorem 7.5.3],[1] it follows from (6.54) and the irreducibility of X that X is positive recurrent and thus, see [6, Theorem 7.2.1 and Definition 7.2.2], that it admits a unique invariant probability measure ν. Thus, for every bounded measurable function f and all $n \geq 0$,

$$\mathbb{E}_\nu\, f(X_n) = \nu f. \tag{6.66}$$

Hence, still for every bounded f,

$$\nu \overline{R} f = \mathbb{E}_\nu \left[\frac{1}{T} \sum_{i=0}^{T-1} f(X_i) \right] = \frac{1}{T} \sum_{i=0}^{T-1} \mathbb{E}_\nu \left[f(X_i) \right] = \nu f \tag{6.67}$$

and thus $\nu \overline{R} = \nu$. It remains to verify that $\nu V < \infty$, which follows by the following standard arguments. By irreducibility of X and the fact that E is countable, we have $\nu(\{x\}) > 0$ for all $x \in E$ and hence (see for instance [6, Theorems 5.2.11 and 5.2.9]), for all $A > 0$ and $x \in E$,

$$\frac{1}{n} \sum_{i=1}^{n} \big(V(X_i) \wedge A \big) \xrightarrow[n \to +\infty]{} \nu(V \wedge A) \quad \mathbb{P}_x\text{-almost surely.} \tag{6.68}$$

By dominated convergence and using (6.64), this implies that

$$\nu(V \wedge A) = \lim_{n \to +\infty} \mathbb{E}_x \frac{1}{n} \sum_{i=1}^{n} \big(V(X_i) \wedge A \big) \leq \liminf_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{n} (\lambda^i V(x) + C_1) = C_1 \tag{6.69}$$

and hence, letting $A \to +\infty$, that $\nu V \leq C_1 < +\infty$. This completes the proof that Assumption (N) holds true.

Furthermore, $\overline{R}$ is the probability kernel of an irreducible Markov chain on E, and thus we deduce from [6, Theorem 7.5.3] and (6.65) that ν is the unique invariant probability measure of $\overline{R}$.

We now show that Theorem 2.10 applies to $\mathfrak{m}$, which implies Proposition 6.15(a). We first show that $\mathbf{R}$ defined by (2.18) is an slqc operator on $B(W) = B(V^{1/q})$, i.e. that it satisfies conditions (QC1) and (QC2) of Definition 2.6, which entails that Theorem 2.10 applies.

Note that (6.58) holds by (6.65). Hence, according to Lemma 6.16, $r_e(\mathbf{R}) \leq \vartheta^{1/q} < 1$ and thus by Definition B.6, for any $\rho \in (\vartheta^{1/q}, 1)$, there exists a decomposition of $B(W)$ into two closed $\mathbf{R}$-invariant subspaces:

$$B(W) = F_\rho \oplus H_\rho, \tag{6.70}$$

such that F_ρ has finite dimension, and the spectral radius of $\mathbf{R}|_{H_\rho}$ is less than ρ. Since the spectrum of F_ρ is finite, this says that the spectrum $\sigma(\mathbf{R})$ contains only a finite number of points λ with $|\lambda| \geq \rho$; moreover, these points satisfy $|\lambda| \leq r(\mathbf{R}) = 1$ and thus $\operatorname{Re} \lambda < 1$ unless $\lambda = 1$. This shows both that (QC2) holds and that 1 is an isolated point in $\sigma(\mathbf{R})$. (As always, $1 \in \sigma(\mathbf{R})$ because $\mathbf{R}1 = 1$.)

The generalized eigenspace of $\mathbf{R}$ corresponding to the eigenvalue 1 is a subspace of F_ρ, and thus has finite dimension. In order to verify (QC1), it remains to show that this dimension is 1, i.e., that the eigenvalue 1 has algebraic multiplicity 1.

We first show that the eigenvalue 1 is simple: The corresponding eigenfunctions of $\mathbf{R}$ satisfy $\mathbf{R}f = f$, which means that they are harmonic functions for the Markov kernel $\overline{R}$. As shown above, ν is the unique invariant probability measure for $\overline{R}$, and furthermore

[1]Theorem 7.5.3 in [6] is not stated correctly, but the direction we use is correct. The other direction becomes correct if, for example, one replaces the irreducibility assumption by a strong irreducibility assumption, which corresponds to our (classical) notion of irreducibility.

$\nu V < \infty$; hence, [6, Proposition 5.2.12] shows that every harmonic function in $B(V)$ is constant ν-a.e., and hence constant everywhere because $\nu(\{x\}) > 0$ for all $x \in E$. This implies that 1 has simple geometric multiplicity: it remains to prove that it also has simple algebraic multiplicity. To do so, let $f \in B(W)$ be such that $(\mathbf{R}-\mathbf{I})^2 f = 0$. Then $(\mathbf{R}-\mathbf{I})f$ is an eigenfunction associated to 1 and hence it is equal to a constant, say $c \in \mathbb{C}$. We deduce that $\mathbf{R}f = f + c$ and hence $\mathbf{R}^n f = f + nc$ for all $n \geq 1$. Moreover, for all $n \geq 1$ and $x \in E$, by iterating (6.65),

$$\mathbf{R}^n V(x) \leq \vartheta \mathbf{R}^{n-1} V(x) + C_1 \leq \vartheta^n V(x) + \sum_{i=0}^{n-1} \vartheta^i C_1 \leq C V(x) \tag{6.71}$$

which implies, by Jensen's inequality,

$$\mathbf{R}^n W(x) \leq (\mathbf{R}^n V(x))^{1/q} \leq C W(x). \tag{6.72}$$

In particular, for all $n \geq 0$,

$$|f(x) + nc| = |\mathbf{R}^n f(x)| \leq \|f\|_{B(W)}\, \mathbf{R}^n W(x) \leq \|f\|_{B(W)} C W(x), \tag{6.73}$$

which implies that $c = 0$ and hence that $(\mathbf{R}-\mathbf{I})f = 0$, so that f is an eigenfunction associated to 1 and hence is constant. We have shown that $\ker\big((\mathbf{R}-\mathbf{I})^2\big) = \ker\big(\mathbf{R}-\mathbf{I}\big) = \{c1 : c \in \mathbb{C}\}$. This implies that the algebraic multiplicity of 1 in the finite-dimensional space F_ρ is 1, and it follows that (QC1) holds true, which completes the proof that $\mathbf{R}$ is slqc.

We set, as in (2.22), $\theta := \sup \operatorname{Re}(\sigma(\mathbf{R}) \setminus \{1\})$. By Theorem 2.10 (with $D := B(W)$), $\theta < 1$ and, for every $\delta \in (0, 1-\theta)$, there exists a constant C_δ such that, for all $f \in B(W)$,

$$\mathbb{E}\left(|\widetilde{\mathfrak{m}}_n - \nu f|^2 \mid \mathfrak{m}_0\right) \leq C_\delta\, \widetilde{\mathfrak{m}}_0 V \left(\frac{\mathfrak{m}_0(E)+1}{\mathfrak{m}_0(E)+n}\right)^{2\delta\wedge 1} \|f\|^2_{B(W)}, \qquad \forall n \geq 1. \tag{6.74}$$

But $\mathfrak{m}_0(E) = 1$ and, by (6.60) and (6.64),

$$\mathbb{E}\big[\widetilde{\mathfrak{m}}_0 V\big] = \frac{1}{T}\sum_{i=0}^{T-1} \mathbb{E}\, V(Z_i) = \frac{1}{T}\sum_{i=0}^{T-1} \mathbb{E}\, V(X_i) < +\infty. \tag{6.75}$$

Hence, (6.74) yields, up to a change of C_δ,

$$\mathbb{E}\left(|\widetilde{\mathfrak{m}}_n - \nu f|^2\right) \leq C_\delta\, n^{-2\delta\wedge 1} \|f\|^2_{B(W)}. \tag{6.76}$$

If furthermore $\delta < 1/2$, then for all $f \in B(W)$,

$$n^\delta\, |\widetilde{\mathfrak{m}}_n - \nu f| \xrightarrow[n\to+\infty]{a.s.} 0. \tag{6.77}$$

For all $n \geq 1$ and $k \in \{0, \dots, T-1\}$, we have

$$\frac{1}{nT+k}\sum_{i=0}^{nT+k-1} \delta_{Z_i} = \frac{nT}{nT+k}\widetilde{\mathfrak{m}}_{n-1} + \frac{1}{nT+k}\sum_{i=nT}^{nT+k-1} \delta_{Z_i}, \tag{6.78}$$

and thus, for all $f \in B(W)$ such that $\|f\|_{B(W)} \leq 1$,

$$\left|\frac{1}{nT+k}\sum_{i=0}^{nT+k-1} f(Z_i) - \frac{nT}{nT+k}\widetilde{\mathfrak{m}}_{n-1} f\right| \leq \frac{1}{nT+k}\sum_{i=nT}^{nT+k-1} W(Z_i) \tag{6.79}$$

$$\leq \frac{(n+1)T}{nT+k}\widetilde{\mathfrak{m}}_n W - \frac{nT}{nT+k}\widetilde{\mathfrak{m}}_{n-1} W. \tag{6.80}$$

Hence,

$$\begin{aligned}
\big|\boldsymbol{\mu}_{nT+k-1} f - \nu f\big| &= \left|\frac{1}{nT+k}\sum_{i=0}^{nT+k-1} f(Z_i) - \nu f\right| \\
&\leq \frac{k}{nT+k}|\nu f| + \frac{nT}{nT+k}|\widetilde{\mathfrak{m}}_{n-1} f - \nu f| + \frac{(n+1)T}{nT+k}|\widetilde{\mathfrak{m}}_n W - \nu W|
\end{aligned}$$

$$+\frac{nT}{nT+k}|\widetilde{\mathfrak{m}}_{n-1}W-\nu W|+\frac{T}{nT+k}\nu W$$
$$\le\frac{1}{n}|\nu f|+|\widetilde{\mathfrak{m}}_{n-1}f-\nu f|+2|\widetilde{\mathfrak{m}}_nW-\nu W|+|\widetilde{\mathfrak{m}}_{n-1}W-\nu W|+\frac{1}{n}\nu W. \tag{6.81}$$

We now obtain (6.56) from (6.81) and (6.76) by Minkowski's inequality. Similarly, (6.57) follows from (6.81) and (6.77), which concludes the proof of Proposition 6.15(a).

We next show that Theorem 2.13 applies to $\mathfrak{m}$ which implies Proposition 6.15(b). We have proved that $\mathbf{R}$ is an slqc operator on $D=B(W)$; moreover, by Lemma 6.16 and (6.65), $r_e(\mathbf{R})\le\vartheta^{1/q}$ where ϑ is given by (6.63). We now assume that $\vartheta^{1/q}<1/2$, and thus $r_e(\mathbf{R})<1/2$. This means that we may take $\rho<1/2$ in (6.70), which entails that the set

$$\{\lambda\in\sigma(\mathbf{R}):\operatorname{Re}\lambda\ge\rho\}=\{\lambda\in\sigma(\mathbf{R}|_{F_\rho}):\operatorname{Re}\lambda\ge\rho\} \tag{6.82}$$

is finite. Since in addition we have $\mathbb{E}((1+\mathfrak{m}_0(E))\widetilde{\mathfrak{m}}_0V)<+\infty$, it follows from Remark 2.14 that one of the cases (1), (2) or (3) in Theorem 2.13 applies to $\mathfrak{m}$ with $D=B(W)$. (The case depends on whether θ is $<1/2$, $=1/2$, or $>1/2$.) Moreover, in cases (2) and (3), we have $\kappa<\infty$ by Remark 2.18.

In addition, by (6.79), for $n\ge1$ and $0\le k<T$,

$$\mathbb{E}\left|\frac{1}{nT+k}\sum_{i=0}^{nT+k-1}f(Z_i)-\frac{nT}{nT+k}\widetilde{\mathfrak{m}}_{n-1}f\right|\le\mathbb{E}\Big[\frac{1}{nT+k}\sum_{i=nT}^{(n+1)T-1}W(Z_i)\Big]$$
$$=\mathbb{E}\Big[\frac{T}{nT+k}\big(\mathfrak{m}_nW-\mathfrak{m}_{n-1}W\big)\Big]=\frac{T}{nT+k}\,\mathbb{E}\big[R^{(n)}_{Y_n}W\big] \tag{6.83}$$

and thus

$$\mathbb{E}\left|\frac{1}{nT+k}\sum_{i=0}^{nT+k-1}f(Z_i)-\widetilde{\mathfrak{m}}_{n-1}f\right|\le\frac{T}{nT+k}\,\mathbb{E}\big[R^{(n)}_{Y_n}W\big]+\frac{k}{nT+k}\,\mathbb{E}\,|\widetilde{\mathfrak{m}}_{n-1}f| \tag{6.84}$$

Both $\mathbb{E}\big[\widetilde{\mathfrak{m}}_{n-1}W\big]$ and $\mathbb{E}\big[R^{(n)}_{Y_n}W\big]$ are uniformly bounded in n by (3.10) in Lemma 3.2 and (3.16) in its proof. Hence, (6.84) yields

$$\mathbb{E}\,|\boldsymbol{\mu}_{nT+k-1}f-\widetilde{\mathfrak{m}}_{n-1}f|=\mathbb{E}\left|\frac{1}{nT+k}\sum_{i=0}^{nT+k-1}f(Z_i)-\widetilde{\mathfrak{m}}_{n-1}f\right|=O(1/n). \tag{6.85}$$

and, in particular, as $n\to\infty$, for any fixed $\alpha<1$,

$$n^\alpha\big(\boldsymbol{\mu}_{nT+k-1}f-\widetilde{\mathfrak{m}}_{n-1}f\big)\overset{\mathrm{p}}{\longrightarrow}0. \tag{6.86}$$

If conclusion (1) of Theorem 2.13 holds for $\mathfrak{m}$, this implies that $(n/T)^{1/2}\,(\boldsymbol{\mu}_nf-\nu f)$ has the same limit in distribution as $(n/T)^{1/2}(\widetilde{\mathfrak{m}}_{\lfloor(n+1)/T\rfloor-1}f-\nu f)$, which equals the limit in (2.27), and similarly for conclusion (2).

If conclusion (3) of Theorem 2.13 holds for $\mathfrak{m}$, we use (6.80) which entails

$$\begin{aligned}\big|\boldsymbol{\mu}_{nT+k-1}f-\widetilde{\mathfrak{m}}_{n-1}f\big|&\le\frac{(n+1)T}{nT+k}\widetilde{\mathfrak{m}}_nW-\frac{nT}{nT+k}\widetilde{\mathfrak{m}}_{n-1}W+\frac{k}{nT+k}|\widetilde{\mathfrak{m}}_{n-1}f|\\&\le\frac{(n+1)T}{nT+k}\widetilde{\mathfrak{m}}_nW-\frac{nT-k}{nT+k}\widetilde{\mathfrak{m}}_{n-1}W\\&=\frac{T+k}{nT+k}\widetilde{\mathfrak{m}}_nW+\frac{nT-k}{nT+k}\big(\widetilde{\mathfrak{m}}_nW-\widetilde{\mathfrak{m}}_{n-1}W\big)\\&\le\frac{2}{n}\widetilde{\mathfrak{m}}_nW+\big|\widetilde{\mathfrak{m}}_nW-\widetilde{\mathfrak{m}}_{n-1}W\big|.\end{aligned} \tag{6.87}$$

In particular, setting $\alpha_n:=n^{1-\operatorname{Re}\lambda_1}/\log^{\kappa-1}n$, we get

$$\left|\alpha_n\big(\boldsymbol{\mu}_{nT+k-1}f-\nu f\big)-\sum_{j=1}^p n^{\mathrm{i}\operatorname{Im}\lambda_j}\Lambda_j\right|\le\left|\alpha_n\left(\widetilde{\mathfrak{m}}_{n-1}f-\nu f\right)-\sum_{j=1}^p n^{\mathrm{i}\operatorname{Im}\lambda_j}\Lambda_j\right|$$

$$+ \alpha_n\big|\widetilde{\mathfrak{m}}_n W - \widetilde{\mathfrak{m}}_{n-1}W\big| + \frac{2\alpha_n}{n}\widetilde{\mathfrak{m}}_n W. \qquad (6.88)$$

The first term on the right-hand side converges to 0 a.s. and in L^2 as $n \to \infty$ according to conclusion (3) of Theorem 2.13 for $\mathfrak{m}$; it is easy to see that we may replace $\widetilde{\mathfrak{m}}_n$ by $\widetilde{\mathfrak{m}}_{n-1}$ in (2.38) because $\alpha_n/\alpha_{n-1} \to 1$. Similarly, the third term in the right-hand side converges to 0 a.s. and in L^2 according to Theorem 2.10(i) for $\mathfrak{m}$ and the fact that $\alpha_n = o(n)$; note that $\mathbb{E}\,|\widetilde{\mathfrak{m}}_n W|^2 = O(1)$ by taking the expectation in (2.23) combined with (6.75). It remains to consider the second term in the right-hand side, for which we observe that

$$\alpha_n\left(\widetilde{\mathfrak{m}}_n W - \widetilde{\mathfrak{m}}_{n-1}W\right) = \alpha_n\left(\widetilde{\mathfrak{m}}_n W - \nu W\right) - \sum_{j=1}^{p} n^{\mathrm{i}\,\mathrm{Im}\,\lambda_j}\Lambda_j$$
$$- \left(\alpha_n\left(\widetilde{\mathfrak{m}}_{n-1} W - \nu W\right) - \sum_{j=1}^{p} n^{\mathrm{i}\,\mathrm{Im}\,\lambda_j}\Lambda_j\right), \qquad (6.89)$$

where both terms go to 0 a.s. and in L^2 as $n \to +\infty$ according to conclusion (3) of Theorem 2.13 for $\mathfrak{m}$. Consequently, the left-hand side of (6.88) converges to 0 a.s. and in L^2. Finally, $\alpha_{nT+k-1}/\alpha_n \to T^{1-\mathrm{Re}\,\lambda_1}$ as $n \to \infty$ and $0 \le k < T$, and it follows easily that

$$\left|\alpha_N\big(\boldsymbol{\mu}_N f - \nu f\big) - \sum_{j=1}^{p} N^{\mathrm{i}\,\mathrm{Im}\,\lambda_j} T^{1-\lambda_j}\Lambda_j\right| \to 0 \qquad (6.90)$$

a.s. and in L^2 as $N \to \infty$.

We conclude by proving Proposition 6.15(c). In order to do so, we note that Lemma 6.16 applied to the transition probability kernel P of X, using (6.54), implies that the corresponding operator $\mathbf{P}$ is quasi-compact on $B(W)$ with $r_e(\mathbf{P}) < r(\mathbf{P}) = 1$. Hence, there exists $\rho < 1$ and a decomposition as in (6.70), and it follows that the spectrum $\sigma(\mathbf{P})$ has only finitely many points λ with $|\lambda| > \rho$, and these points all have $|\lambda| \le 1$. In particular, 1 is isolated in $\sigma(\mathbf{P})$ and thus

$$\eta := \inf\big\{|1-s| : s \in \sigma(\mathbf{P}) \setminus \{1\}\big\} > 0. \qquad (6.91)$$

We have $\mathbf{R} = \frac{1}{T}\sum_{i=0}^{T-1}\mathbf{P}^i$, and thus the spectral mapping theorem [5, Theorem VII.4.10] shows that the spectrum of $\mathbf{R}$ is given by

$$\sigma(\mathbf{R}) = \left\{\frac{1}{T}\sum_{i=0}^{T-1} s^i : s \in \sigma(\mathbf{P})\right\}, \qquad (6.92)$$

and thus

$$\sigma(\mathbf{R}) \setminus \{1\} = \left\{\frac{1}{T}\sum_{i=0}^{T-1} s^i : s \in \sigma(\mathbf{P}) \setminus \{1\}\right\}. \qquad (6.93)$$

For every $s \in \sigma(\mathbf{P}) \setminus \{1\}$ we have $|s| \le 1$ and $|1-s| \ge \eta$, and thus

$$\left|\frac{1}{T}\sum_{i=0}^{T-1} s^i\right| = \left|\frac{1}{T}\frac{1-s^T}{1-s}\right| \le \frac{2}{\eta T} \xrightarrow[T\to+\infty]{} 0. \qquad (6.94)$$

In particular, if we choose T_0 such that $T_0 > 4/\eta$, then for every $T \ge T_0$, we have by (6.93) and (6.94),

$$\theta = \sup\big\{\mathrm{Re}\,\lambda : \lambda \in \sigma(\mathbf{R}) \setminus \{1\}\big\} \le \sup\big\{|\lambda| : \lambda \in \sigma(\mathbf{R}) \setminus \{1\}\big\} < 1/2, \qquad (6.95)$$

and thus case (1) applies in part (b).

This concludes the proof of Proposition 6.15. □

Proof of Lemma 6.16. The proof relies on [15, Theorem XIV.3]. Fix $q > 1$ and set as usual $W := V^{1/q}$. Jensen's inequality and the assumption (6.58) entail that

$$\mathbf{R}W \le (\mathbf{R}V)^{1/q} \le (\vartheta V + C)^{1/q} \le \vartheta^{1/q} W + C^{1/q}. \tag{6.96}$$

In particular, this shows that $\mathbf{R}$ acts as a bounded operator on $B(W)$; we regard in the rest of the proof $\mathbf{R}$ as an operator on $B(W)$. By induction similar to (6.71), (6.96) also implies that $\mathbf{R}^n W \le CW$ for some constant $C > 0$ and all $n \ge 0$. Thus $\|\mathbf{R}^n\|_{B(W)} \le C$ and by the spectral radius formula [5, Proposition VII.3.8], the spectral radius $\sigma(\mathbf{R})$ of $\mathbf{R}$ is at most 1. Since $1 \in B(W)$ and we have $\mathbf{R}1 = 1$, we deduce that 1 is an eigenvalue of $\mathbf{R}$. We can thus conclude that the spectral radius of $\mathbf{R}$, as a bounded operator on $B(W)$, equals 1.

To apply [15, Theorem XIV.3], we consider the Banach space $(B(W), \|\cdot\|_{B(W)})$, endowed with the continuous norm $\|\cdot\|_{B(V)}$. We check that

(i) $\mathbf{R}(\{f \in B(W) : \|f\|_{B(W)} \le 1\})$ is totally bounded in $(B(W), \|\cdot\|_{B(V)})$;
(ii) there exists a constant $M > 0$ such that, for all $f \in B(W)$, $\|\mathbf{R}f\|_{B(V)} \le M\|f\|_{B(V)}$;
(iii) for any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ such that

$$\|\mathbf{R}f\|_{B(W)} \le (\vartheta^{1/q} + \varepsilon)\|f\|_{B(W)} + C_\varepsilon \|f\|_{B(V)}. \tag{6.97}$$

Once this is proved, the conclusion of Lemma 6.16 immediately follows from [15, Theorem XIV.3].

We first prove (i). Recall that a set in a metric space is totally bounded if for every $\varepsilon > 0$ there is a finite ε-net in it, i.e., a finite subset F such that every point in the set has distance at most ε to F. (This is also called precompact, and in a complete metric space it is equivalent to relatively compact. Thus (i) says that $\mathbf{R}$ is a compact operator $B(W) \to B(V)$. See e.g. [7, I.6.14–15].) Let

$$U := \{f \in B(W) : \|f\|_{B(W)} \le 1\} = \{f \in \mathbb{C}^E : |f(x)| \le W(x),\ \forall x \in E\} \tag{6.98}$$

be the unit ball of $B(W)$. Since $\mathbf{R}$ is bounded on $B(W)$, $\mathbf{R}(U) \subseteq CU$ for some constant C, and it suffices to show that U is totally bounded for the norm $\|\cdot\|_{B(V)}$.

Let $\varepsilon > 0$. Fix $M > 0$, and let $K_M := \{x \in E : V(x) \le M\}$; recall that this set is finite. Consider first the restrictions to K_M. $U_M := \{f|_{K_M} : f \in U\}$ is a bounded set in the finite-dimensional space $\mathbb{C}^{K_M}$, and thus it is relatively compact. (In fact, it is compact.) Hence, there exists a finite set $\{f_i\}_{i=1}^N \subset U_M$ such that for every $f \in U$ there exists an f_i with

$$\max_{x \in K_M} |f(x) - f_i(x)| < \varepsilon. \tag{6.99}$$

Extend every f_i to a function on E, still denoted f_i, by $f_i(x) := 0$ for $x \notin K_M$. If $f \in U$ and $x \notin K_M$, then for every $i \in \{1, \dots, N\}$,

$$\frac{|f(x) - f_i(x)|}{V(x)} = \frac{|f(x)|}{V(x)} \le \frac{W(x)}{V(x)} = V(x)^{1/q - 1} \le M^{1/q-1}. \tag{6.100}$$

By choosing M large enough, this is less than ε. Hence, if f_i is chosen to satisfy (6.99), then $|f(x) - f_i(x)|/V(x) \le \varepsilon$ for every $x \in E$, and thus $\|f - f_i\|_{B(V)} \le \varepsilon$. Hence $\{f_i\}_1^N$ is a finite ε-net in U. Consequently, ((i)) holds.

The property (ii) is a consequence of (6.58), which indeed implies that, for all $f \in B(W)$,

$$\|\mathbf{R}f\|_{B(V)} = \sup_{x \in E} \frac{|\mathbf{R}f(x)|}{V(x)} \le \sup_{x \in E} \frac{\|f/V\|_\infty \mathbf{R}V(x)}{V(x)} \le \|f\|_{B(V)} (\vartheta + C). \tag{6.101}$$

We now prove (iii): Since $\inf_{x \notin K_M} W(x) \ge M \to +\infty$ when $M \to +\infty$, we deduce from (6.96) that, for any $\varepsilon > 0$, there exists $M(\varepsilon) > 0$ and a constant $c_\varepsilon > 0$ such that

$$\mathbf{R}W \le (\vartheta^{1/q} + \varepsilon)W + c_\varepsilon \mathbf{1}_{K_{M(\varepsilon)}}(x),\ \forall x \in E. \tag{6.102}$$

Hence, for all $x \notin K_{M(\varepsilon)}$,

$$|\mathbf{R}f(x)| \le \|f/W\|_\infty \mathbf{R}W(x) \le \|f\|_{B(W)} (\vartheta^{1/q} + \varepsilon)\, W(x). \tag{6.103}$$

But, according to (6.58), for all $x \in K_{M(\varepsilon)}$,

$$|\mathbf{R}f(x)| \le \|f/V\|_\infty \, \mathbf{R}V(x) \le \|f\|_{B(V)} \, (\vartheta + C) \, V(x)$$
$$\le \|f\|_{B(V)} \, (\vartheta + C) \max_{y \in K_{M(\varepsilon)}} \frac{V(y)}{W(y)} \, W(x). \tag{6.104}$$

Setting $C_\varepsilon = (\vartheta + C) \max_{y \in K_{M(\varepsilon)}} \frac{V(y)}{W(y)}$ and using the two previous inequalities, we deduce that, for all $x \in E$,

$$\frac{|\mathbf{R}f(x)|}{W(x)} \le (\vartheta^{1/q} + \varepsilon)\|f\|_{B(W)} + C_\varepsilon \|f\|_{B(V)}, \tag{6.105}$$

which concludes the proof of (iii) and hence of Lemma 6.16. □

Appendix A. Kernels and the definition of the mvpp

We use the notation introduced in Section 1.4.

A.1. Kernels.

Recall that given two measurable spaces $(S, \mathcal{S})$ and $(T, \mathcal{T})$, a *kernel* from S to T is a map $s \mapsto \beta_s$ from S to the set $\mathcal{M}_+(T)$ of positive measures on $(T, \mathcal{T})$ that is measurable; in other words, $s \mapsto \beta_s(B)$ is $\mathcal{S}$-measurable for every fixed set $B \in \mathcal{T}$. See e.g. [22, pp. 20–21] or [23, Section 1.3] for a detailed discussion; we summarize a few facts that we need.

A probability kernel is the special case when each β_s is a probability measure on T.

A signed kernel is defined in the same way, with β_s a signed measure on T.

If $\alpha \in \mathcal{P}(S)$ and β is a probability kernel from S to T, then $\alpha \otimes \beta$ denotes the probability measure on $S \times T$ given by

$$\alpha \otimes \beta(A) := \int_S \mathrm{d}\alpha(s) \int_T \mathbf{1}_A(s, t) \, \mathrm{d}\beta_s(t). \tag{A.1}$$

This means that if (X, Y) is a random variable in $S \times T$ with the distribution $\alpha \otimes \beta$, then X has distribution α, and the conditional distribution of Y given $X = x$ is β_x (for a.e. x); hence (A.1) formalizes the notion of choosing randomly first X with distribution α, and then Y with distribution β_X.

The construction (A.1) generalizes to the case where α is a probability kernel from a third space U to S; then $\alpha \otimes \beta$ is a probability kernel from U to $S \times T$.

A.2. The mvpp.

The definition of the mvpp in Section 2 uses a family $(R_x^{(1)})_{x \in E}$ of random (signed) measures in $\mathcal{M}_\mathbb{R}(E)$. Only their distributions matter, so letting $\mathcal{R}_x := \mathcal{L}(R_x^{(1)})$, the distribution of $R_x^{(1)}$, it is equivalent to start with a family $\mathcal{R} = (\mathcal{R}_x, x \in E)$ of probability distributions in $\mathcal{M}_\mathbb{R}(E)$, or equivalently a map $\mathcal{R} : E \to \mathcal{P}(\mathcal{M}_\mathbb{R}(E))$; we may then define, for each $x \in E$, R_x as a random measure in $\mathcal{M}_\mathbb{R}(E)$ with distribution $\mathcal{R}_x$, and $R_x^{(n)}$ as a sequence of independent copies of $R_x^{(1)}$.

Our basic assumption is that $\mathcal{R} = (\mathcal{R}_x, x \in E)$ is a probability kernel from E to $\mathcal{M}_\mathbb{R}(E)$, which we call the *replacement kernel*. (We abuse notation and use the same name also for the corresponding family $(R_x^{(1)})_x$ of random measures.)

Remark A.1. The assumption that $\mathcal{R}$ is a probability kernel from E to $\mathcal{M}_\mathbb{R}(E)$ implies that its expectation $\overline{R}$ defined in (2.5) is a signed kernel from E to E, provided that (2.4) holds.

It is also easy to see that the assumption that $\mathcal{R}$ is a kernel implies that $\widetilde{\mathbf{B}}_x(f, g)$ in (2.36) is a measurable function of x; hence also $\mathbf{B}_x(f)$ and $\mathbf{C}_x(f)$ in (2.30) are measurable. □

Let us now try to formalize the definition of the mvpp, starting from a given replacement kernel $\mathcal{R}$ and a given deterministic $\mathfrak{m}_0 \in \mathcal{M}_{>0}(E)$. Our aim is to define random variables

$Y_n \in E$ and $R^{(n)}_{Y_n} \in \mathcal{M}_{\mathbb{R}}(E)$ for all $n \geq 1$ satisfying the description in Section 2; then $\mathfrak{m}_n$ is given by

$$\mathfrak{m}_n := \mathfrak{m}_0 + \sum_{i=1}^{n} R^{(i)}_{Y_i}. \tag{A.2}$$

Equivalently, we want to construct the joint distribution of all $(Y_n, R^{(n)}_{Y_n})$, $n \geq 1$, as a probability measure on $(E \times \mathcal{M}_{\mathbb{R}}(E))^\infty$. We will achieve this using the construction (A.1) twice. However, we have (so far) only been able to do so assuming one of the following assumptions (or both).

(i) $R^{(1)}_x$ is always a positive measure, so there are no subtractions in the urn, or
(ii) E is a Borel space (see e.g. [22, Appendix A]).

The reasons for the technical assumption (ii) will be discussed below.

(i): Consider first the simple case when $R^{(1)}_x$ always is a positive measure, i.e., $R^{(1)}_x \in \mathcal{M}_+(E)$. In this case, there is no need to consider signed measures. Write $\mathcal{X} := E \times \mathcal{M}_+(E)$. Let $n \geq 0$ and assume that we have constructed the distribution μ_n of $(Y_i, R^{(i)}_{Y_i})_1^n$, as a probability measure on $\mathcal{X}^n$. (This assumption is void for $n = 0$.) We write an element of $\mathcal{X}^n$ as $(y_i, r_i)_1^n$; then we can realize Y_i and $R^{(i)}_{Y_i}$ for $i \leq n$ as the coordinate functions y_i and r_i on the probability space $(\mathcal{X}^n, \mu_n)$. By (A.2), $\mathfrak{m}_n$ then is given by the function $m_n : \mathcal{X}^n \to \mathcal{M}_{>0}(E)$ defined by

$$m_n\big((y_i, r_i)_1^n\big) := \mathfrak{m}_0 + \sum_{i=1}^{n} r_i. \tag{A.3}$$

Thus, the normalized measure $\widetilde{\mathfrak{m}}_n$ is given by the function $\gamma_n : \mathcal{X}^n \to \mathbb{P}(E)$ defined by

$$\gamma_n(\xi_n) := m_n(\xi_n)/m_n(\xi_n)(E). \tag{A.4}$$

Nota that γ_n is a probability kernel from $\mathcal{X}^n$ to E.

We want Y_{n+1} to be a random element of E such that, conditioned on the history up to time n, Y_{n+1} has the distribution $\widetilde{\mathfrak{m}}_n$. In other words, conditioned on $(Y_i, R^{(i)}_{Y_i})_1^n = \xi_n \in \mathcal{X}^n$, Y_{n+1} has the conditional distribution $\gamma_n(\xi_n)$. This means that

$$\big((Y_i, R^{(i)}_{Y_i})_1^n, Y_{n+1}\big) \sim \mu_n \otimes \gamma_n, \tag{A.5}$$

and we may take this as a formal definition of (the distribution of) Y_{n+1}.

Next, the replacement kernel $\mathcal{R}$ is now assumed to be a probability kernel from E to $\mathcal{M}_+(E)$. We may (trivially) regard it as a kernel from $\mathcal{X}^n \times E$ by letting $\mathcal{R}_{(\xi_n,x)} := \mathcal{R}_x$. Hence, (A.1) defines the probability measure $(\mu_n \otimes \gamma_n) \otimes \mathcal{R}$ on $\mathcal{X}^n \times E \times \mathcal{M}_+(E) = \mathcal{X}^{n+1}$. We want $R^{(n+1)}_{Y_{n+1}}$ to have the conditional distribution, given the previous history, $\mathcal{R}_{Y_{n+1}}$, and thus

$$((Y_i, R^{(i)}_{Y_i})_1^n, Y_{n+1}, R^{(n+1)}_{Y_{n+1}}) \sim \mu_n \otimes \gamma_n \otimes \mathcal{R}. \tag{A.6}$$

(Note that $\otimes$ is associative: $(\mu_n \otimes \gamma_n) \otimes \mathcal{R} = \mu_n \otimes (\gamma_n \otimes \mathcal{R})$, so we may omit the brackets.) We may take (A.6) as a formal definition of $R^{(n+1)}_{Y_{n+1}}$.

In other words, our formal construction is

$$\mu_{n+1} := \mu_n \otimes \gamma_n \otimes \mathcal{R} \in \mathcal{P}\big(\mathcal{X}^{n+1}\big). \tag{A.7}$$

This completes the inductive step, and starting from the trivial probability measure μ_0 on a one-point space, we obtain recursively a probability measure μ_n on $\mathcal{X}^n$ for every $n \geq 1$. Finally, since μ_n are obtained recusively by composing with the probability kernels $\gamma_n \otimes \mathcal{R}$, the Ionescu Tulcea theorem [22, Theorem 6.17] now shows the existence of a probability space and infinite sequences Y_n and $R^{(n)}_{Y_n}$ with the desired distribution; this defines also $\mathfrak{m}_n$ by (A.2). Equivalently, the Ionescu Tulcea theorem shows the existence of a probability measure on $\mathcal{X}^\infty$ with the desired projection μ_n to $\mathcal{X}^n$ for each n. This completes the construction in the special case when $R_x^{(1)} \in \mathcal{M}_+(E)$. It follows from the construction that $\bigl(\mathfrak{m}_n, Y_n, R_{Y_n}^{(n)}\bigr)_{n\geq 1}$ is a Markov chain.

(ii): Consider now the general case, when $R_x^{(1)} \in \mathcal{M}_\mathbb{R}(E)$ is a signed measure, but we assume that the urn is tenable. Assume now also that E is a Borel space. We may now define $\mathcal{X} := E \times \mathcal{M}_\mathbb{R}(E)$, and try to argue as above. The only problem is that γ_n defined by (A.4) is not a probability kernel, since $m_n(\xi_n)$ is not a positive measure for all $\xi_n \in \mathcal{X}^n$. (We will even have $m_n(\xi_n)(E) = 0$ for some ξ_n, and then $\gamma_n(\xi_n)$ is not even defined.) We thus have to modify the definition of γ_n. Consider again some $n \geq 0$ and assume that we have constructed $\mu_n \in \mathcal{P}(\mathcal{X}^n)$. Note that for a Borel space E, $\mathcal{M}_{>0}(E)$ is a measurable subset of $\mathcal{M}_\mathbb{R}(E)$, as may easily be verified. Let $\Upsilon_n := m_n^{-1}(\mathcal{M}_{>0}(E))$, where $m_n : \mathcal{X}^n \to \mathcal{M}_\mathbb{R}(E)$ is the function defined in (A.3); thus Υ_n is a measurable subset of $\mathcal{X}^n$. We assume that the urn is tenable, which means that $\mathfrak{m}_n$ a.s. satisfies $\mathfrak{m}_n \in \mathcal{M}_{>0}(E)$. In other words, $m_n(\xi) \in \mathcal{M}_{>0}(E)$ for μ_n-a.e. ξ; equivalently, $\mu_n(\Upsilon_n) = 1$.

We may now modify (A.4) and define a probability kernel γ_n from $\mathcal{X}^n$ to E by

$$\gamma_n(\xi_n) := \begin{cases} \widetilde{m}_n(\xi_n) := m_n(\xi_n)/m_n(\xi_n)(E), & \text{if } \xi_n \in \Upsilon_n, \\ \nu, & \text{if } \xi_n \in \mathcal{X}^n \setminus \Upsilon_n, \end{cases} \tag{A.8}$$

where ν is an arbitrary, fixed probability measure on E. Then the construction proceeds as above. (The choice of ν does not affect μ_{n+1}, since $\mu_n(\Upsilon) = 1$.) This completes the construction in case (ii), when E is a Borel space.

What happens when E is not a Borel space? In some cases it might be possible to modify the construction above; for example if (for each $n \geq 1$) there exists a measurable subset Υ_n of $m_n^{-1}(\mathcal{M}_{>0}(E))$ such that $\mu_n(\Upsilon_n) = 1$. However, we will see in Example A.2 that in general no such Υ_n exists. In general, unless (i) ar (ii) above holds, we have to assume that the process $\mathfrak{m}_n$ is defined by some external construction. (See Example A.2 for an example where a construction is trivial.)

Example A.2. Let $E := \{0,1\}^\mathcal{A}$ for some uncountable set $\mathcal{A}$. Let Z be a random element of E, with some distribution $\nu_Z \in \mathcal{P}(E)$, and let

$$R_x := -\delta_x + 2\delta_Z, \qquad x \in E; \tag{A.9}$$

also, let $\mathfrak{m}_0 := \delta_{x_0}$ for some $x_0 \in E$. This describes an urn with balls (corresponding to point masses) labelled by elements of E; we start with a single ball x_0, and in each step we remove one randomly chosen ball, and add two new balls with label Z_n, where $(Z_n)_1^\infty$ are i.i.d. This process is obviously well defined and tenable. Nevertheless, we will see that there is no measurable set Υ_1 such that the construction (A.8) works for $n = 1$. (In particular, $\mathcal{M}_+(E)$ is not a measurable subset of $\mathcal{M}_\mathbb{R}(E)$.) Note that necessarily $Y_1 = x_0$, and thus $R_{Y_1}^{(1)} = -\delta_{x_0} + 2\delta_{Z_1}$. Hence, the distribution μ_1 of $(Y_1, R_{Y_1}^{(1)})$ is the product measure $\delta_{x_0} \times \mathcal{L}(R_{Y_1}^{(1)})$. Suppose that $\Upsilon_1 \subseteq \mathcal{X} = E \times \mathcal{M}_\mathbb{R}(E)$ is measurable and such that $\mu_1(\Upsilon_1) = 1$ and $m_1(y, r) = \mathfrak{m}_0 + r \in \mathcal{M}_{>0}(E)$ for every $(y, r) \in \Upsilon_1$. We will show that this leads to a contradiction.

Let $\Lambda \subseteq \mathcal{M}_\mathbb{R}(E)$ be a non-empty measurable set. Recall that the σ-field on $\mathcal{M}_\mathbb{R}(E)$ is generated by the mappings $\mu \mapsto \mu(B)$ for $B \in \mathcal{E}$, where $\mathcal{E}$ is the σ-field on E. It is well known that this implies that there exists a countable family $(B_i) \subset \mathcal{E}$ such that Λ belongs to the σ-field generated by the mappings $\mu \mapsto \mu(B_i)$, $i \in \mathbb{N}$. (Because the union of these σ-field over all countable families (B_i) is a σ-field.)

Similarly, since the product σ-field $\mathcal{E}$ is generated by the coordinate maps $(x_a)_{a\in\mathcal{A}} \mapsto x_a$ for $a \in \mathcal{A}$, for each $B \in \mathcal{E}$ there is a countable subset $\mathcal{A}_B \subset \mathcal{A}$ and a (measurable) set $\tilde{B} \subseteq \{0,1\}^{\mathcal{A}_B}$ such that

$$B = \tilde{B}_i \times \{0,1\}^{\mathcal{A}\setminus\mathcal{A}_B}. \tag{A.10}$$

Fix a coordinate $a' \in \mathcal{A} \setminus \bigcup_i \mathcal{A}_{B_i}$. Define, for $j \in \{0,1\}$, the elements $z^j = (z^j_a) \in E$ by

$$z^j_a := \begin{cases} j, & a = a', \\ 0, & a \neq a'. \end{cases} \tag{A.11}$$

Take a signed measure $\lambda \in \Lambda$, and for $N \geq 0$, let $\lambda_N := \lambda + N(\delta_{z^0} - \delta_{z^1})$. For each B_i, we have $a' \notin \mathcal{A}_{B_i}$, and thus, by (A.10), $z^0 \in B_i \iff z^1 \in B_i$. Consequently, for every $N \geq 0$,

$$\lambda_N(B_i) = \lambda(B_i) + N\big(\mathbf{1}\{z^0 \in \mathcal{A}_{B_i}\} - \mathbf{1}\{z^1 \in \mathcal{A}_{B_i}\}\big) = \lambda(B_i) \tag{A.12}$$

for every B_i. Since Λ is in the σ-field generated by the maps $\mu \mapsto \mu(B_i)$, and $\lambda \in \Lambda$, it follows that $\lambda_N \in \Lambda$. On the other hand, if $B := \{(x_a) \in E : x_{a'} = 1\}$, then $B \in \mathcal{E}$ and $\lambda_N(B) = \lambda(B) - N$; hence, if N is large enough, $\lambda_N(B) < 1$ and thus $\lambda_N + \mathfrak{m}_0 \notin \mathcal{M}_{>0}(E)$.

We have shown that there is no nonempty measurable set $\Lambda \subseteq \mathcal{M}_{\mathbb{R}}(E)$ such that

$$\lambda \in \Lambda \implies \lambda + \mathfrak{m}_0 \in \mathcal{M}_{>0}(E). \tag{A.13}$$

However, if Υ_1 is as above, then the section $\Lambda := \{r \in \Upsilon_1 : (x_0, r) \in \Upsilon_1\}$ is measurable, satisfies (A.13), and also $\mathbb{P}(R^{(1)}_{Y_1} \in \Lambda) = 1$, a contradiction.

Note that the proof shows that $\mathcal{M}_{>0}(E)$ is not a measurable subset of $\mathcal{M}_{\mathbb{R}}(E)$, and, moreover, that it does not contain any non-empty measurable subset. (The same holds for $\mathcal{M}_+(E)$.) ☐

Appendix B. Some functional analysis

In this appendix we state some general results on spectra of operators in Banach spaces; these are used in our examples in Section 6. The results are simple and have presumably been known for a long time, but since we have not found references to the results in the form that we need, we give full proofs for completeness.

Recall that if T is a bounded operator on $\mathcal{X}$, and T^* is its adjoint acting on the dual space $\mathcal{X}^*$, then [5, Proposition VII.6.1]

$$\sigma(T^*) = \sigma(T). \tag{B.1}$$

Our first lemma deals with the situation when we instead consider T^* as acting on a subspace of $\mathcal{X}^*$.

Definition B.1. If K is a compact subset of $\mathbb{C}$, define $K^\frown$ as the union of K and all bounded connected components of $\mathbb{C} \setminus K$; in other words, its complement $\mathbb{C} \setminus K^\frown$ is the unbounded component of $\mathbb{C} \setminus K$. ($K^\frown$ is known as the polynomially convex hull of K, see [5, Definition VII.5.2 and Proposition VII.5.3].) In particular, if T is a bounded operator on a Banach space and $\rho_\infty(T)$ denotes the unbounded component of the resolvent set $\rho(T) = \mathbb{C} \setminus \sigma(T)$, then

$$\sigma(T)^\frown := \mathbb{C} \setminus \rho_\infty(T). \tag{B.2}$$

We let $\langle x^*, x \rangle$ denote the pairing of elements $x^* \in \mathcal{X}^*$ and $x \in \mathcal{X}$, for any Banach space $\mathcal{X}$.

Lemma B.2. *Let T be a bounded operator on a complex Banach space $\mathcal{X}$, and suppose that $\mathcal{Y} \subseteq \mathcal{X}^*$ is a closed subspace of the dual space $\mathcal{X}^*$ such that the adjoint operator T^* maps $\mathcal{Y}$ into itself.*

(i) *Then*

$$\sigma\big(T^*|_{\mathcal{Y}}\big) \subseteq \sigma(T)^\frown. \tag{B.3}$$

(ii) *Suppose further that $\mathcal{Y}$ is norm-determining, i.e., that if $x \in \mathcal{X}$, then*

$$\|x\| = \sup\{\langle x^*, x \rangle : x^* \in \mathcal{Y},\ \|x^*\| = 1\}. \tag{B.4}$$

Then also

$$\sigma(T) \subseteq \sigma\big(T^*|_{\mathcal{Y}}\big)^\frown \tag{B.5}$$

and thus

$$\sigma\bigl(T^*|_{\mathcal{Y}}\bigr)^\frown = \sigma(T)^\frown. \tag{B.6}$$

Proof. (i): As said above, the spectrum $\sigma(T^*)$ of T^* as an operator on $\mathcal{X}^*$ equals $\sigma(T)$, and the resolvent is simply given by $(z-T^*)^{-1} = ((z-T)^{-1})^*$, $z \in \rho(T) = \mathbb{C} \setminus \sigma(T)$.

We first show that this resolvent maps $\mathcal{Y}$ into itself, at least when $z \notin \sigma(T)^\frown$. To do so, we take $y \in \mathcal{Y}$ and let $x^{**} \in \mathcal{X}^{**}$ be such that $x^{**} \perp \mathcal{Y}$, i.e., $\langle x^{**}, y'\rangle = 0$ for every $y' \in \mathcal{Y}$. Consider the function

$$g(z) := \langle x^{**}, (z-T^*)^{-1}y\rangle, \qquad z \in \rho(T) = \rho(T^*). \tag{B.7}$$

The function g is analytic on $\rho(T)$, see [5, Theorem VII.3.6]. Furthermore, if $|z| > \|T\|$, then $(z-T^*)^{-1} = \sum_{k=0}^\infty z^{-k-1}(T^*)^k$ with an absolutely convergent sum, and thus, because $T^*(\mathcal{Y}) \subseteq \mathcal{Y}$,

$$(z-T^*)^{-1}y = \sum_{k=0}^\infty z^{-k-1}(T^*)^k y \in \mathcal{Y}. \tag{B.8}$$

Consequently, (B.7) and (B.8) imply that if $|z| > \|T\|$, then $g(z) = 0$. By analytic continuation, $g(z) = 0$ in the unbounded connected component $\rho_\infty(T)$ of $\rho(T)$.

This holds for any $x^{**} \perp \mathcal{Y}$, and thus, by definition of g in (B.7), it follows that $(z-T^*)^{-1}y \in \mathcal{Y}$ for all $z \in \rho_\infty(T)$. In other words, for all $z \in \rho_\infty(T)$, we have $(z-T^*)^{-1} : \mathcal{Y} \to \mathcal{Y}$, which means that it is the inverse of the restriction $(z-T^*)|_{\mathcal{Y}}$. Hence, for all $z \in \rho_\infty(T)$, z belongs to the resolvent set $\rho(T^*|_{\mathcal{Y}})$; in other words, $\rho_\infty(T) \subseteq \rho(T^*|_{\mathcal{Y}})$, and thus (B.3) holds by (B.2).

(ii): The canonical embedding $\mathcal{X} \to \mathcal{X}^{**}$ induces a linear map $\mathcal{X} \to \mathcal{Y}^*$, which is an isometric embedding by the assumption (B.4). Hence, we may regard $\mathcal{X}$ as a subspace of $\mathcal{Y}^*$. We may thus apply part (i) with $\mathcal{X}$ and $\mathcal{Y}$, and also T and T^*, interchanged. This yields (B.5), and (B.6) then easily follows from (B.3) and (B.5). ☐

Corollary B.3. *Let T be a bounded operator on a complex Banach space $\mathcal{X}$, and suppose that $\mathcal{Y} \subseteq \mathcal{X}^*$ is a closed subspace of the dual space $\mathcal{X}^*$ such that the adjoint operator T^* maps $\mathcal{Y}$ into itself. Suppose further that $\mathcal{Y}$ is norm-determining. Then*

(i) *T is an slqc operator on $\mathcal{X}$ if and only if T^* is an slqc operator on $\mathcal{Y}$.*
(ii) *T is a small operator on $\mathcal{X}$ if and only if T^* is a small operator on $\mathcal{Y}$.*

Proof. (i): Suppose that T is an slqc operator. Let $\theta := \sup \operatorname{Re}\bigl(\sigma(T) \setminus \{1\}\bigr)$ and note that $\theta < 1$ as in Lemma 2.9(i). We then have

$$U := \bigl\{\lambda : \operatorname{Re}\lambda > \theta\bigr\} \setminus \{1\} \subset \rho(T), \tag{B.9}$$

which implies, since the set U is connected and unbounded,

$$\bigl\{\lambda : \operatorname{Re}\lambda > \theta\bigr\} \setminus \{1\} \subset \rho_\infty(T), \tag{B.10}$$

and thus

$$\sigma(T)^\frown \subset \bigl\{\lambda : \operatorname{Re}\lambda \le \theta\bigr\} \cup \{1\}. \tag{B.11}$$

Hence Lemma B.2 yields

$$\sigma(T^*|_{\mathcal{Y}})^\frown = \sigma(T)^\frown \subset \{\lambda : \operatorname{Re}\lambda \le \theta\} \cup \{1\}, \tag{B.12}$$

which implies (QC2) for $T^*|_{\mathcal{Y}}$, and also that 1 is isolated in $\sigma(T^*|_{\mathcal{Y}})$ if 1 belongs to this spectrum at all. It remains to show only that 1 is an eigenvalue of $T^*|_{\mathcal{Y}}$ and that the corresponding spectral projection $\Pi_1(T^*|_{\mathcal{Y}})$ has rank 1.

We can regard T^* as an operator on $\mathcal{X}^*$ or on its subspace $\mathcal{Y}$. In both cases we have, see [5, Equation VII.6.9],

$$\Pi_1(T^*) = \frac{1}{2\pi \mathrm{i}} \oint_\Gamma (z-T^*)^{-1}\,\mathrm{d}z \tag{B.13}$$

where we choose Γ to be a small circle around 1 inside $\rho_\infty(T)$, cf. (B.10). By the proof of Lemma B.2, if $z \in \Gamma$, then $(z - T^*)^{-1}$ maps $\mathcal{Y}$ into itself, and its restriction to $\mathcal{Y}$ is $(z - T^*|_{\mathcal{Y}})^{-1}$. Hence, (B.13) shows that $\Pi_1(T^*)$ maps $\mathcal{Y}$ into itself, and that its restriction to $\mathcal{Y}$ is $\Pi_1(T^*|_{\mathcal{Y}})$.

Moreover, $(z - T^*)^{-1} = \big((z - T)^{-1}\big)^*$ for $z \in \Gamma$, and thus by (B.13) and the same formula for T, we have $\Pi_1(T^*) = \Pi_1(T)^*$. By Assumption (QC1), $\Pi_1(T)$ has rank 1, and is thus given by

$$\Pi_1(T)x = \langle x_0^*, x\rangle x_0 \tag{B.14}$$

for some non-zero $x_0 \in \mathcal{X}$ and $x_0^* \in \mathcal{X}^*$ with $\langle x_0^*, x_0\rangle = 1$. It follows that, for any $x^* \in \mathcal{X}^*$,

$$\Pi_1(T^*)(x^*) = \Pi_1(T)^*(x^*) = \langle x^*, x_0\rangle x_0^*. \tag{B.15}$$

Since $\mathcal{Y}$ is norm-determining, there exists $y \in \mathcal{Y}$ such that $\langle y, x_0\rangle \neq 0$. Since $\Pi_1(T^*) : \mathcal{Y} \to \mathcal{Y}$, we have $\Pi_1(T^*)(y) \in \mathcal{Y}$, and (B.15) then shows that $x_0^* \in \mathcal{Y}$.

We have shown that $\Pi_1(T^*|_{\mathcal{Y}})$ is the rank 1 operator defined by (B.15) restricted to $x^* \in \mathcal{Y}$. In particular, $x_0^* \in \mathcal{Y}$ is an eigenvector with $T^*x_0^* = x_0^*$. Hence, (QC1) in Definition 2.6 holds for $T^*|_{\mathcal{Y}}$, which concludes the proof that $T^*|_{\mathcal{Y}}$ is slqc if T is.

The converse follows, as in the proof of Lemma B.2(ii), by interchanging the roles of $\mathcal{X}$ and $\mathcal{Y}$, noting that $\mathcal{X}$ always is norm-determining as a subspace of $\mathcal{Y}^*$.

(ii): Now suppose that T is small. This means that in the proof of (i), we have $\theta < 1/2$. Hence, (B.12) shows that T^* is small. The converse follows as above. ☐

In the following lemma, we compare the spectra of the "same" operator in two different spaces. When necessary, we use subscripts such as $T_{\mathcal{X}}$ to denote the space where we consider the operator.

Lemma B.4. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two complex Banach spaces and suppose that $\mathcal{Y} \subseteq \mathcal{X}$ with a continuous, but not necessarily isometric, inclusion. Suppose that T is a bounded operator on $\mathcal{X}$ such that $T(\mathcal{X}) \subseteq \mathcal{Y}$.*

(i) *Then*

$$\sigma(T_{\mathcal{X}}) = \begin{cases} \sigma(T_{\mathcal{Y}}), & \mathcal{Y} = \mathcal{X}, \\ \sigma(T_{\mathcal{Y}}) \cup \{0\}, & \mathcal{Y} \subsetneq \mathcal{X}. \end{cases} \tag{B.16}$$

(We do not make any claims on whether $0 \in \sigma(T_{\mathcal{Y}})$ or not.)

(ii) *If $\lambda \neq 0$ is an isolated point in $\sigma(T_{\mathcal{X}})$, then $\Pi_\lambda(T_{\mathcal{Y}})$ equals the restriction of $\Pi_\lambda(T_{\mathcal{X}})$ to $\mathcal{Y}$. (Thus we can use the notation Π_λ for both without confusion.) Moreover, $\Pi_\lambda\mathcal{X} = \Pi_\lambda\mathcal{Y} \subseteq \mathcal{Y}$.*

(iii) *$T_{\mathcal{X}}$ is slqc if and only if $T_{\mathcal{Y}}$ is slqc. $T_{\mathcal{X}}$ is small if and only if $T_{\mathcal{Y}}$ is small.*

Proof. (i): Note first that by the closed graph theorem, $T : \mathcal{X} \to \mathcal{Y}$ is a bounded operator. Hence, the restriction $T_{\mathcal{Y}}$ to $\mathcal{Y}$ is a bounded operator on $\mathcal{Y}$, and the spectra $\sigma(T_{\mathcal{X}})$ and $\sigma(T_{\mathcal{Y}})$ are both defined.

If $\mathcal{Y} = \mathcal{X}$, then the norms on $\mathcal{X}$ and $\mathcal{Y}$ are equivalent, again by the closed graph theorem, and thus $\sigma(T_{\mathcal{X}}) = \sigma(T_{\mathcal{Y}})$.

Assume in the sequel that $\mathcal{Y} \neq \mathcal{X}$. In particular, since $T(\mathcal{X}) \subseteq \mathcal{Y}$, T is not onto $\mathcal{X}$, and thus $T_{\mathcal{X}}$ is not invertible; hence $0 \in \sigma(T_{\mathcal{X}})$.

Suppose that $\lambda \in \rho(T_{\mathcal{X}})$. This means that the resolvent $R_\lambda := (\lambda - T)^{-1}$ exists as a bounded operator on $\mathcal{X}$. We have

$$I = (\lambda - T)R_\lambda = \lambda R_\lambda - TR_\lambda. \tag{B.17}$$

Hence, if $y \in \mathcal{Y}$, then, using again $T(\mathcal{X}) \subseteq \mathcal{Y}$,

$$\lambda R_\lambda y = y + TR_\lambda y \in \mathcal{Y}. \tag{B.18}$$

Since $0 \notin \rho(T_{\mathcal{X}})$, as remarked above, we have $\lambda \neq 0$. Hence (B.18) implies $R_\lambda y \in \mathcal{Y}$, and thus $R_\lambda : \mathcal{Y} \to \mathcal{Y}$. It follows immediately that the restriction of R_λ to $\mathcal{Y}$ is an inverse to

$\lambda - T_{\mathcal{Y}}$, and thus $\lambda \in \rho(T_{\mathcal{Y}})$. We have shown that

$$\rho(T_{\mathcal{X}}) \subseteq \rho(T_{\mathcal{Y}}) \setminus \{0\}. \tag{B.19}$$

Conversely, suppose that $\lambda \in \rho(T_{\mathcal{Y}})$, and let $R'_\lambda := (\lambda - T_{\mathcal{Y}})^{-1} : \mathcal{Y} \to \mathcal{Y}$ denote the corresponding resolvent. Since $T : \mathcal{X} \to \mathcal{Y}$, we may define the operator $Q := I + R'_\lambda T$ on $\mathcal{X}$. For any $x \in \mathcal{X}$, we then have, since $Tx \in \mathcal{Y}$,

$$(\lambda - T)Qx = (\lambda - T)x + (\lambda - T)R'_\lambda Tx = \lambda x - Tx + Tx = \lambda x \tag{B.20}$$

and

$$\begin{aligned} Q(\lambda - T)x &= (\lambda - T)x + R'_\lambda T(\lambda - T)x = (\lambda - T)x + R'_\lambda(\lambda - T)Tx = \lambda x - Tx + Tx \\ &= \lambda x. \end{aligned} \tag{B.21}$$

Hence, if also $\lambda \neq 0$, then $\lambda^{-1}Q$ is an inverse of $\lambda - T$ on $\mathcal{X}$, and thus $\lambda \in \rho(T_{\mathcal{X}})$. Consequently, $\rho(T_{\mathcal{Y}}) \setminus \{0\} \subseteq \rho(\mathcal{X})$. Thus equality holds in (B.19), and thus (B.16) holds.

(ii): Let Γ be a sufficiently small circle around λ, such that Γ and its interior are disjoint from $\sigma(T_{\mathcal{X}}) \setminus \{\lambda\}$. Then the spectral projections $\Pi_\lambda(T_{\mathcal{X}})$ and $\Pi_\lambda(T_{\mathcal{Y}})$ are both obtained by integrating the respective resolvents along Γ, as in (B.13). If $\lambda' \in \Gamma$, then, as shown in the proof of (i), $(\lambda' - T_{\mathcal{Y}})^{-1}$ is the restriction of $(\lambda' - T_{\mathcal{X}})^{-1}$ to $\mathcal{Y}$; hence it follows that $\Pi_\lambda(T_{\mathcal{Y}})$ is the restriction of the projection $\Pi_\lambda(T_{\mathcal{X}})$ to $\mathcal{Y}$. Consequently,

$$\Pi_\lambda \mathcal{Y} = (\Pi_\lambda \mathcal{X}) \cap \mathcal{Y}. \tag{B.22}$$

Moreover, T maps $\Pi_\lambda \mathcal{X}$ into itself, and the restriction of T to $\Pi_\lambda \mathcal{X}$ is invertible (since its spectrum is $\{\lambda\}$, and $\lambda \neq 0$), and thus onto. Since $T : \mathcal{X} \to \mathcal{Y}$, it follows that $\Pi_\lambda \mathcal{X} \subseteq \mathcal{Y}$. Combined with (B.22), this yields $\Pi_\lambda \mathcal{X} = \Pi_\lambda \mathcal{Y}$ as asserted.

(iii): An immediate consequence of (i) and (ii). ☐

Lemma B.5. *Let $\mathcal{N}$ be a closed subspace of a complex Banach space $\mathcal{X}$, and let $\mathcal{Z} := \mathcal{X}/\mathcal{N}$. Suppose that T is a bounded operator on $\mathcal{X}$ such that $Tn = 0$ for every $n \in \mathcal{N}$. Then T can also be regarded as an operator on $\mathcal{Z}$, and the following holds.*

- (i) *If $\mathcal{N} \neq \{0\}$, then $\sigma(T_{\mathcal{X}}) = \sigma(T_{\mathcal{Z}}) \cup \{0\}$. (If $\mathcal{N} = \{0\}$, then trivially $\sigma(T_{\mathcal{X}}) = \sigma(T_{\mathcal{Z}})$.)*
- (ii) *If $\lambda \neq 0$ is an isolated point in $\sigma(T_{\mathcal{X}})$, then $\Pi_\lambda \mathcal{N} = \{0\}$, and thus $\Pi_\lambda(T_{\mathcal{X}})$ induces an operator on $\mathcal{Z} = \mathcal{X}/\mathcal{N}$; this induced operator equals $\Pi_\lambda(T_{\mathcal{Z}})$. Moreover, the quotient map $\mathcal{X} \to \mathcal{Z}$ is a bijection $\Pi_\lambda(T_{\mathcal{X}})\mathcal{X} \to \Pi_\lambda(T_{\mathcal{Z}})\mathcal{Z}$.*
- (iii) *$T_{\mathcal{X}}$ is slqc if and only if $T_{\mathcal{Z}}$ is slqc. $T_{\mathcal{X}}$ is small if and only if $T_{\mathcal{Z}}$ is small.*

Proof. That T can be regarded as an operator on the quotient space $\mathcal{Z}$ is well known. Moreover, $\mathcal{Z}^*$ is identified with the closed subspace $\{x^* \in X^* : x^*(\mathcal{N}) = 0\}$ of $\mathcal{X}^*$.

If $x^* \in X^*$ and $n \in N$, then $\langle T^*x^*, n\rangle = \langle x^*, Tn\rangle = 0$; thus $T^*x^* \in \mathcal{Z}^*$ for every $x^* \in \mathcal{X}^*$. Hence, we can apply Lemma B.4 to T^* on the spaces $\mathcal{X}^*$ and $\mathcal{Z}^* \subseteq \mathcal{X}^*$.

(i): If $\mathcal{N} \neq \{0\}$, then (B.16) yields $\sigma(T^*_{\mathcal{X}^*}) = \sigma(T^*_{\mathcal{Z}^*}) \cup \{0\}$, and thus $\sigma(T_{\mathcal{X}}) = \sigma(T_{\mathcal{Z}}) \cup \{0\}$ by (B.1).

(ii): By (B.1), λ is an isolated point of $\sigma(T^*_{X^*})$. Recall also that (by the argument in the proof of Corollary B.3) $\Pi_\lambda(T)^* = \Pi_\lambda(T^*)$, for any of the spaces $\mathcal{X}$ and $\mathcal{Z}$. Lemma B.4(ii) thus shows that $\Pi_\lambda(T_{\mathcal{X}})^* : \mathcal{X}^* \to \mathcal{Z}^*$. Hence, if $n \in \mathcal{N}$, then for any $x^* \in \mathcal{X}^*$ we have $\langle x^*, \Pi_\lambda n\rangle = \langle \Pi^*_\lambda x^*, n\rangle = 0$, and thus $\Pi_\lambda n = 0$. Hence $\Pi_\lambda \mathcal{N} = \{0\}$ as claimed.

Moreover, if $\pi : \mathcal{X} \to \mathcal{Z}$ is the quotient mapping, then $\pi^* : \mathcal{Z}^* \to \mathcal{X}^*$ is the inclusion mapping, and Lemma B.4(ii) shows also that $\Pi_\lambda(T_{\mathcal{X}})^*\pi^* = \pi^*\Pi_\lambda(T_{\mathcal{Z}})^*$. Hence, by taking adjoints,

$$\pi\Pi_\lambda(T_{\mathcal{X}}) = \Pi_\lambda(T_{\mathcal{Z}})\pi, \tag{B.23}$$

which shows that $\Pi_\lambda(T_{\mathcal{X}})$ induces $\Pi_\lambda(T_{\mathcal{Z}})$ on $\mathcal{Z}$. Furthermore, (B.23) also implies

$$\pi\Pi_\lambda(T_{\mathcal{X}})\mathcal{X} = \Pi_\lambda(T_{\mathcal{Z}})\pi\mathcal{X} = \Pi_\lambda(T_{\mathcal{Z}})\mathcal{Z}, \tag{B.24}$$

and thus π maps $\Pi_\lambda(T_{\mathcal{X}})\mathcal{X}$ onto $\Pi_\lambda(T_{\mathcal{Z}})\mathcal{Z}$. Moreover, π is injective on $\Pi_\lambda(T_{\mathcal{X}})\mathcal{X}$, since $\pi x = 0$ for some $x \in \Pi_\lambda(T_{\mathcal{X}})\mathcal{X}$ means that $x \in \mathcal{N}$, and thus $x = \Pi_\lambda(T_{\mathcal{X}})x = 0$ as shown above.

(iii): An immediate consequence of (i) and (ii). □

We end this appendix with a standard definition.

Definition B.6. Let T be a bounded operator in a complex Banach space B. Let $r(T)$ denote the spectral radius of T. Furthermore, consider all decompositions $B = F \oplus H$ as a direct sum of two closed T-invariant subspaces such that $\dim(F) < \infty$, and define the *essential spectral radius* of T by

$$r_e(T) := \inf\bigl\{r(T|_H) : B = F \oplus H \text{ as above}\bigr\}. \tag{B.25}$$

Remark B.7. It is easily seen that the definition (B.25) is equivalent to [15, Definition XIV.1]. There are several other, equivalent, definitions; for example, $r_e(T)$ equals the spectral radius of T in the Banach algebra $\mathcal{B}(B)/\mathcal{K}(B)$, where $\mathcal{B}(B)$ is the Banach algebra of bounded linear operators and $\mathcal{K}(B)$ is the ideal of compact operators. For this, and the relation to the *essential spectrum* (which has several, non-equivalent, versions), see e.g. [8, §1.4] and [24, p. 243]. □

Remark B.8. Taking $F = \{0\}$ and $H = B$ in (B.25) shows that $r_e(T) \le r(T)$ for every T. An operator T is *quasi-compact* if $r_e(T) < r(T)$. (See [15, Definition II.1] for another, equivalent, definition.) □

Appendix C. A technical lemma

We state an elementary lemma that is used in the proof of Theorem 2.13.

Lemma C.1. *Let $\alpha \in \mathbb{R}$ and $k \ge 0$. Then, as $n \to \infty$,*

$$\sum_{j=1}^{n} j^{-1-\mathrm{i}\alpha} \log^k(n/j) = \begin{cases} (1+o(1))\dfrac{\log^{k+1} n}{k+1} & \text{if } \alpha = 0, \\ O(\log^k n) & \text{if } \alpha \neq 0. \end{cases} \tag{C.1}$$

Proof. We first approximate the sum by an integral. Let $g_n(x) := x^{-1-\mathrm{i}\alpha}\log^k(n/x)$, $x \ge 1$. Then, assuming in the sequel $n \ge 2$,

$$g_n'(x) = (-1-\mathrm{i}\alpha)x^{-2-\mathrm{i}\alpha}\log^k(n/x) - kx^{-2-\mathrm{i}\alpha}\log^{k-1}(n/x) = O\bigl((\log^k n)x^{-2}\bigr), \qquad x \ge 1. \tag{C.2}$$

Hence, for $j \ge 1$,

$$g_n(j) - \int_j^{j+1} g_n(x)\,\mathrm{d}x = \int_j^{j+1}\bigl(g_n(j) - g_n(x)\bigr)\,\mathrm{d}x = O\bigl((\log^k n)j^{-2}\bigr). \tag{C.3}$$

Consequently, with the change of variables $x = n/y$,

$$\begin{aligned} \sum_{j=1}^{n} j^{-1-\mathrm{i}\alpha}\log^k(n/j) &= \sum_{j=1}^{n} g_n(j) = g_n(n) + \sum_{j=1}^{n-1}\Bigl(\int_j^{j+1} g_n(x)\,\mathrm{d}x + O\bigl((\log^k n)j^{-2}\bigr)\Bigr) \\ &= \int_1^n g_n(x)\,\mathrm{d}x + O\bigl(\log^k n\bigr) = \int_1^n \frac{\log^k(n/x)}{x^{1+\mathrm{i}\alpha}}\,\mathrm{d}x + O\bigl(\log^k n\bigr) \\ &= n^{-\mathrm{i}\alpha}\int_1^n \frac{\log^k y}{y^{1-\mathrm{i}\alpha}}\,\mathrm{d}y + O\bigl(\log^k n\bigr). \end{aligned} \tag{C.4}$$

It thus suffices to consider the final integral in (C.4).

If $\alpha = 0$, then

$$\int_1^n \frac{\log^k y}{y^{1-\mathrm{i}\alpha}}\,\mathrm{d}y = \int_1^n \frac{\log^k y}{y}\,\mathrm{d}y = \left[\frac{\log^{k+1} y}{k+1}\right]_1^n = \frac{\log^{k+1} n}{k+1}, \tag{C.5}$$

and thus (C.1) follows in this case.

If $\alpha \neq 0$, we use integration by parts and get

$$\int_1^n \log^k(y) y^{\mathrm{i}\alpha-1}\,\mathrm{d}y = \Bigl[\log^k(y)\frac{y^{\mathrm{i}\alpha}}{\mathrm{i}\alpha}\Bigr]_1^n - \frac{k}{\mathrm{i}\alpha}\int_1^n \log^{k-1}(y) y^{\mathrm{i}\alpha-1}\,\mathrm{d}y$$

$$= O\big(\log^k(n)\big) + \int_1^n O\big(\log^{k-1}(n)\big)\frac{\mathrm{d}y}{y} = O\big(\log^k(n)\big). \tag{C.6}$$

Hence, (C.1) follows from (C.4) in this case too. □

Remark C.2. It is possible to show that for $\alpha \neq 0$, the sum in (C.1) is asymptotic to $\zeta(1+\mathrm{i}\alpha)\log^k n$. Moreover, for any α, an asymptotic expansion with an arbitrary number of terms may be obtain by singularity analysis as in similar examples in [10, Section 3.1]. □

References

[1] Krishna B. Athreya and Samuel Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Annals of Mathematical Statistics* **39** (1968), 1801–1817.

[2] Antar Bandyopadhyay, Svante Janson and Debleena Thacker. Strong convergence of infinite color balanced urns under uniform ergodicity. *Journal of Applied Probability* **57** (2020), no. 3, 853–865.

[3] Antar Bandyopadhyay and Debleena Thacker. Pólya urn schemes with infinitely many colors. *Bernoulli* **23** (2017), no. 4B, 3243–3267.

[4] David Blackwell and James B. MacQueen. Ferguson distributions via Pólya urn schemes. *Annals of Statistics* **1** (1973), no. 2, 353–355.

[5] John B. Conway. *A Course in Functional Analysis*, Second edition. Springer-Verlag, New York, 1990.

[6] Randal Douc, Eric Moulines, Pierre Priouret and Philippe Soulier. *Markov Chains.* Springer, 2018.

[7] Nelson Dunford and Jacob T. Schwartz, *Linear Operators. I. General Theory.* Interscience Publishers, New York, 1958.

[8] D. E. Edmunds and W. D. Evans *Spectral Theory and Differential Operators.* Oxford University Press, Oxford, 1987.

[9] F. Eggenberger and G. Pólya. Über die Statistik verketetter Vorgänge. *Zeitschrift für Angewandte Mathematik und Mechanik* **3**(4) (1923), 279–289.

[10] James Allen Fill, Philippe Flajolet and Nevin Kapur. Singularity analysis, Hadamard products, and tree recurrences. *J. Comput. Appl. Math.* **174** (2005), no. 2, 271–313.

[11] Denis F. Grebenkov and Binh T. Nguyen. Geometrical structure of Laplacian eigenfunctions. *SIAM Review*, **55**(4) (2013), 601–667.

[12] Allan Gut. *Probability: A Graduate Course*, 2nd ed., Springer, New York, 2013.

[13] P. Hall and C. C. Heyde. *Martingale Limit Theory and its Application.* Academic Press, New York, 1980.

[14] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 4th ed. Oxford Univ. Press, Oxford, UK, 1960.

[15] Hubert Hennion and Loïc Hervé. *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness.* Lecture Notes in Mathematics, 1766. Springer-Verlag, Berlin, 2001.

[16] A. Henrot. *Shape Optimization and Spectral Theory*, De Gruyter, 2017.

[17] Svante Janson. Limit theorems for certain branching random walks on compact groups and homogeneous spaces. *Ann. Probab.* **11** (1983), 909–930.

[18] Svante Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and Applications* **110** (2004), no. 2, 177–245.

[19] Svante Janson. Asymptotic degree distribution in random recursive trees. *Random Structures Algorithms* **26** (2005), no. 1-2, 69–83.

[20] Svante Janson. Limit theorems for triangular urn schemes. *Probab. Theory Related Fields* **134** (2006), 417–452.

[21] Svante Janson. Random replacements in Pólya urns with infinitely many colours. *Electronic Communications in Probability* **24** (2019), Paper no. 23, 1–11.

[22] Olav Kallenberg. *Foundations of Modern Probability.* 2nd ed., Springer, New York, 2002.

[23] Olav Kallenberg. *Random Measures, Theory and Applications.* Springer, Cham, 2017.

[24] Tosio Kato. *Perturbation Theory for Linear Operators.* Second edition. Springer-Verlag, Berlin-New York, 1976.

[25] James R. Kuttler and Vincent G. Sigillito. Eigenvalues of the Laplacian in two dimensions. *SIAM Review* **26**(2) (1984), 163–193.

[26] Cécile Mailler and Jean-François Marckert. Measure-valued Pólya processes. *Electronic Journal of Probability* **22** (2017), Paper no. 26, 33 pp.

[27] Cécile Mailler and Denis Villemonais. Stochastic approximation on non-compact measure spaces and application to measure-valued Pólya processes. Preprint, 2018–2020. arXiv:1809.01461v4

[28] A. A. Markov. Sur quelques formules limites du calcul des probabilités (Russian). *Bulletin de lAcadémie Impériale des Sciences* **11** (1917), no. 3, pp 177–186.

[29] *NIST Handbook of Mathematical Functions.* Edited by Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert and Charles W. Clark. Cambridge Univ. Press, 2010. Also available as *NIST Digital Library of Mathematical Functions*, http://dlmf.nist.gov/

[30] Noela Müller. Central limit theorem analogues for multicolour urn models. Preprint, 2019. arXiv:1604.02964v5

[31] Nicolas Pouyanne. An algebraic approach to Pólya processes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **44** (2008), no. 2, 293–323.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
Email address: svante.janson@math.uu.se

University of Bath, Claverton Down, Bath BA2 7AY, UK
Email address: c.mailler@bath.ac.uk

Universit de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France
Email address: denis.villemonais@univ-lorraine.fr