

Continuum graph dynamics via population dynamics: well-posedness, duality and equilibria

Andreas Greven¹, Frank den Hollander², Anton Klimovsky³, Anita Winter⁴

October 31, 2023

Abstract

In this paper we consider stochastic processes taking values in a set of continuum graphs we call *graphemes*, defined as equivalence classes of sequences of vertices labelled by \mathbb{N} embedded in an uncountable Polish space (with the cardinality of the continuum), together with an $\mathbb{N} \times \mathbb{N}$ connection matrix with entries 0 or 1 specifying the absence or presence of edges between pairs of vertices. In particular, we construct a Markov process on a Polish state space \mathbb{G} of graphemes suitable to describe the time-space path of countable graphs effectively. It is expedient to allow the dynamics to depend on types associated with the vertices, drawn from a Polish type-set V , which leads to a Markov process on a Polish state space \mathbb{G}^V of *marked graphemes*.

The class of dynamics we propose arises by specifying simple rules for the evolution of finite graphs and passing to the limit of infinite graphs. The evolution of graphemes is characterised by *well-posed martingale problems*, and leads to strong Markov processes with the Feller property. In particular, we obtain *diffusions* with continuous paths driven by a second-order operator playing the role of a generator. Our dynamics generalises and modifies the *graphon* dynamics introduced in [AdHR21] based on models from population dynamics. We exploit the *duality* as well as the *genealogy* inherent in these models to build up the grapheme dynamics. The *equilibria* are identified in terms of certain key distributions arising in population dynamics, are supported on a small subset of the Polish space, and reflect decompositions of the Polish space into random components corresponding to completely connected components.

A key tool in the construction is the connection between evolutions of graphs and evolutions of genealogies based on classical models from population dynamics, such as the Fleming-Viot diffusion and the Feller diffusion, and their genealogy-valued analogues. The genealogy process drives the evolution of an *environment* of Polish spaces in which the countable graphs are embedded, as well as the evolution of the connection-matrix distribution. The key observation is that the genealogy of the embedded process contains all the information on the space-time structure of the vertices and the edges up to the current time. This is why the genealogy is an *intrinsic structure* for the evolution of graphemes, besides being mathematically convenient.

Keywords: Graph-valued Markov processes, graphemes, marked graphemes, duality, martingale problem, genealogy, population genetics.

MSC 2020: 05C80, 60J68, 60J70, 92D25.

Acknowledgement: AG was supported by the Deutsche Forschungsgemeinschaft (through grant DFG-GR 876/16-2 of SPP-1590). FdH was supported by the Netherlands Organisation for Scientific Research (through NWO Gravitation Grant NETWORKS-024.002.003), and by the Alexander

von Humboldt Foundation. AK was supported by the Deutsche Forschungsgemeinschaft (through Project-ID 443891315 within SPP 2265 and through Project-ID 412848929). AW was supported by the Deutsche Forschungsgemeinschaft (through grant DFG-SPP-2265).

¹⁾ Department Mathematik, Universität Erlangen-Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany
greven@mi.uni-erlangen.de

²⁾ Mathematisch Instituut, Universiteit Leiden, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands
denholla@math.leidenuniv.nl

³⁾ Institut für Mathematik, Emil-Fischer-Strasse 30, 97074 Würzburg, Germany
anton.klymovskiy@mathematik.uni-wuerzburg.de

⁴⁾ Universität Essen, Fakultät für Mathematik, Thea-Leymann-Strasse, 45141 Essen, Germany
anita.winter@uni-due.de

Contents

1	Background, motivation and main theorems	3
1.1	Literature on graphon dynamics	4
1.2	Framework	7
1.2.1	Graphemes	7
1.2.2	Extensions of graphemes	10
1.2.3	Preview on general grapheme dynamics	11
1.2.4	Main goals for general grapheme dynamics	12
1.3	Examples of finite-graph evolution rules and main theorems	13
1.3.1	Two classes of evolution rules from population dynamics	14
1.3.2	Main goals for grapheme dynamics arising from the evolution rules	16
1.3.3	From population dynamics to grapheme dynamics	17
1.3.4	Main theorems	19
1.4	Consequences and extensions of main theorems	26
1.4.1	Equilibrium processes	26
1.4.2	Extension: grapheme diffusions with mutation and/or selection of types	29
1.4.3	Grapheme diffusions with equilibria allowing for non-completely connected components	30
1.4.4	Explicit representations of laws of equilibrium statistics	32
1.5	Relation to the graphon literature	34
1.6	Outline of the remainder of the paper	36
2	What have we achieved? What remains to be done?	36
2.1	Conclusions for the issue of embedding	37
2.2	Choice of state space, challenges and perspectives: intrinsic role of genealogies	40
3	Grapheme evolution 1: State space, topology and test functions	41

3.1	Test functions, topology and tightness for grapheme spaces	41
3.1.1	Test functions on grapheme spaces and topology	41
3.1.2	Tightness in grapheme spaces	46
3.2	Extension to varying grapheme sizes	48
3.3	Extension to marked graphemes	48
4	Grapheme evolution 2: Examples and martingale problems	50
4.1	From finite-graph processes to grapheme processes: two classes of examples	51
4.1.1	Fleming-Viot and Dawson-Watanabe: finite grapheme processes	51
4.1.2	Finite grapheme processes and their generators	54
4.2	Construction: Fleming-Viot and Dawson-Watanabe grapheme operators	56
4.2.1	Fleming-Viot, Dawson-Watanabe operators on functions of connection matrix	56
4.2.2	Operators for functionals of the genealogy	57
4.2.3	Dawson-Watanabe	58
4.3	Extension to non-completely connected components	59
5	Proofs for state spaces and topologies	60
5.1	Algebra of test functions	60
5.2	Polish spaces	61
6	A prelude on duality	61
7	Proofs for grapheme diffusions with completely connected components	62
7.1	Main idea for the proof	62
7.1.1	The process: from genealogies to graphemes	63
7.1.2	Proofs: functionals of connection-matrix distributions and joint laws	66
7.2	Proof of the two corollaries	71
7.3	Adding mutation	72
7.4	Completion of the proof of the main theorems	73
8	Proofs for grapheme diffusions with non-completely connected components	73
8.1	Completion of the proof: martingale problem	74
8.2	Completion of the proof: equilibrium	75

1 Background, motivation and main theorems

We are interested in the stochastic evolution of large random finite *graphs* viewed as Markov processes following simple evolution rules for vertices and edges. We study both the *dynamics* of such graphs and their *equilibria* when the number of vertices n tends to infinity. As limits we obtain *diffusion processes* on a space of *continuum random graphs*

embedded in a Polish space, i.e., the vertices of the graph are the points in the Polish space. The finite-graph evolution rules have a counterpart in individual-based models from population dynamics, a connection exploited in [AdHR21]. As new tool we use the *genealogy* of the underlying population models to build up the dynamics of the random graph together with the dynamics of the random embedding, in such a way that we capture the full information on the *space-time process* of the *evolving countable graph*. The limit $n \rightarrow \infty$ requires us to introduce an abstract concept of a graph, which we call *grapheme* and which already for finite n is a generalisation of the standard concept of a graph. We start from Markov pure-jump processes for finite graphs and in the limit $n \rightarrow \infty$ end up with *grapheme diffusions*.

The present section contains the main ideas and the main theorems of the paper, and is organised as follows:

- Section 1.1 recalls the notion of equivalence classes of *graphons* and their role in describing limits of dense graphs, and summarises recent results on *graphon dynamics*.
- Section 1.2 defines equivalence classes of *graphemes* (of fixed size or variable size, unmarked or marked), which are the key objects in the present paper and are *different* from graphons. It also gives a quick preview on the dynamics of graphemes we are after, and formulates the main goals for general grapheme dynamics.
- Section 1.3 introduces several classes of *evolution rules* for finite graphs based on classical models from *population dynamics*, states our main goals for grapheme dynamics arising from these evolution rules, and formulates our main theorems for this grapheme dynamics, which involve characterisation in terms of *well-posed martingale problems*, description of *path properties*, identification of *equilibria*, and formulation of *duality* relations.
- Section 1.4 describes consequences of the main theorems, analyses equilibria in some detail, provides extensions of the evolution rules in different directions, and explores the difference between *completely connected components* (so-called cliques) and *non-completely connected components*.
- Section 1.5 relates the main theorems to the work in [AdHR21], and makes the *link* between graphemes and graphons.
- Section 1.6 gives the outline of the remainder of the paper.

Section 1 states all the main theorems and lists all the ideas that are needed to build up the grapheme dynamics. Sections 2–8 contain the technicalities of the construction and the proofs of the main theorems. Each covers about one half of the paper.

1.1 Literature on graphon dynamics

Static graphons. Let \mathcal{W} be the space of measurable symmetric functions $h: [0, 1]^2 \rightarrow [0, 1]$. A finite simple graph G with n vertices can be represented, after taking care of

equivalences, as a *graphon*. Define $h^G \in \mathcal{W}$ in a natural way as (see Fig. 1)

$$(1.1) \quad h^G(x, y) = \begin{cases} 1 & \text{if there is an edge between vertex } \lceil nx \rceil \text{ and vertex } \lceil ny \rceil, \\ 0 & \text{otherwise,} \end{cases}$$

which leads to an $n \times n$ block graphon. The space of functions \mathcal{W} is endowed with the *cut distance*

$$(1.2) \quad d_{\square}(h_1, h_2) = \sup_{S, T \subset [0, 1]} \left| \int_{S \times T} dx dy [h_1(x, y) - h_2(x, y)] \right|, \quad h_1, h_2 \in \mathcal{W}.$$

On \mathcal{W} there is a natural equivalence relation \sim . Namely, let Σ be the space of measure-preserving bijections $\sigma: [0, 1] \rightarrow [0, 1]$. Then $h_1, h_2 \in \mathcal{W}$ are said to be equivalent, written $h_1 \sim h_2$, when $\delta_{\square}(h_1, h_2) = 0$, where δ_{\square} is the *cut metric* defined by

$$(1.3) \quad \delta_{\square}(\tilde{h}_1, \tilde{h}_2) = \inf_{\sigma_1, \sigma_2 \in \Sigma} d_{\square}(h_1^{\sigma_1}, h_2^{\sigma_2}), \quad \tilde{h}_1, \tilde{h}_2 \in \widetilde{\mathcal{W}},$$

with $h^{\sigma}(x, y) = h(\sigma x, \sigma y)$, where \tilde{h} denotes the equivalence class of which h is a representative. The equivalence relation yields the *quotient space* $(\widetilde{\mathcal{W}}, \delta_{\square})$, which is compact (see Theorem 5.1. in [LS07]) and which is the proper object to describe countable limits of finite graphs. This is the *state space* in which random equivalence classes of graphons take their values.

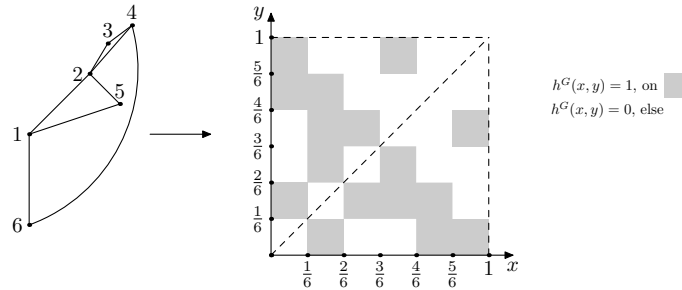


Figure 1: Example of a finite graph (with 6 vertices and 8 edges) and its empirical graphon.

For $h \in \mathcal{W}$ and F a finite simple graph with m vertices and edge set $E(F)$, define

$$(1.4) \quad t(F, h) = \int_{[0, 1]^m} dx_1 \dots dx_m \prod_{\{i, j\} \in E(F)} h(x_i, x_j).$$

Then the *homomorphism density* of F in G equals $t(F, h^G)$, where h^G is the empirical function defined in (1.1). Note that $t(F, h)$ is the same for all representatives h in the equivalence class \tilde{h} . Any graphon $\tilde{h} \in \widetilde{\mathcal{W}}$ is uniquely determined by its subgraph densities, and can be approximated by block graphons. The important mathematical fact is that convergence in $(\widetilde{\mathcal{W}}, \delta_{\square})$ is *equivalent* to convergence of all the subgraph densities (see Fig. 2) so that this space *equals* the compactification of the topology introduced by subgraph counts on $\widetilde{\mathcal{W}}$. For a more detailed description of the structure of the space $(\widetilde{\mathcal{W}}, \delta_{\square})$ we refer the reader to [BCL⁺08, BCL⁺12, DGKR15].

Graphons are suited to describe limits of *dense* graphs, for which the number of edges is of the order of the square of the number of vertices. For *non-dense* graphs we get a single equivalence class of graphemes only, and a multi-scale analysis is required, leading to new objects, for example, a *graphex* (see e.g. [VR15]).

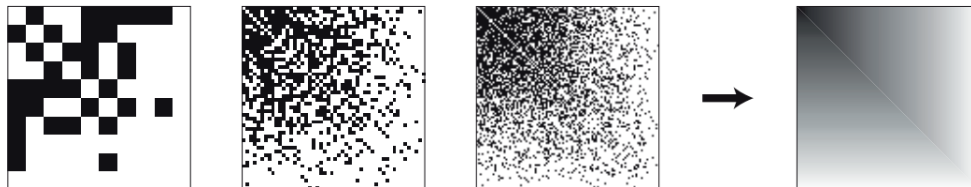


Figure 2: Example of convergence of an empirical graphon to a graphon.

Graphon-valued stochastic processes. The first attempts to construct *dynamics* on the space $\widetilde{\mathcal{W}}$ of graphons were made in [Cra16, Cra17], using the Aldous-Hoover theory for exchangeable infinite arrays. This work led to $\widetilde{\mathcal{W}}$ -valued Markov processes with càdlàg paths of locally bounded variation, i.e., a mixture of a stochastic pure-jump process and a deterministic flow. A further generalisation of this construction, beyond infinite random exchangeable zero-one-valued adjacency matrices, was obtained in [CK18]. Specifically, when it comes to agent-based modelling, the agents typically interact in a way that depends on the current interaction strength between them, or on other features that can be associated with an edge in the graph linking them. It is natural to allow not only for discrete edge features, such as interaction versus no interaction, but also for continuous edge features. In fact, the latter makes the theory of Markovian infinite exchangeable arrays more transparent. Finally, [CK18] highlights an issue that arises in the description of the finite projections of exchangeable Markovian arrays as conditional Markov processes.

Subsequently, building on models from population dynamics, in [AdHR21] *diffusion-like* graphon-valued processes were constructed. Even though [AdHR21] provided a large class of examples, intended as a *proof of concept*, no general theory was developed and consequently many structural questions remained open. For instance, it was not clear whether or not the examples constructed in [AdHR21] are strong Markov processes that can be described by a generator acting on a dense class of test functions on the graphon space. Moreover, all the examples had *trivial* equilibria concentrated on constant graphons.

Further steps were taken in [BdHM23, BdHM22], where *sample-path large deviation principles* were derived for dynamic Erdős-Rényi random graphs (generalising the static large deviation principle in [CV11]; see also [Cha15]), and for random graphs in which each vertex has a type that fluctuates randomly over time, in such a way that *collectively* the paths of the edges and the vertex types up to a given time determine the probabilities that the edges are present or absent at that time.

In the present paper we take a different route towards building a limiting dynamics, based on a new notion we call *grapheme*, which views a countable graph as an object embedded in a measure space built from a *Polish space* (and therefore also looks at finite

embedded graphs). This embedding is necessary to build a *stochastic process* of graphemes via *martingale problems*. Like in [AdHR21], we exploit the machinery developed for models from population dynamics, but we add the idea to also exploit explicitly the *tree structure* behind the *genealogy* of populations, which we review as we go along and which is both natural and mathematically convenient, as we will see. This approach allows for storing information on the history of the *edge structure in the current state*, and hence for studying the *time-space path process* of countable graphs.

1.2 Framework

In Section 1.2.1 we define what we mean by *graphemes*. In Section 1.2.2 we extend this notion to *marked graphemes* and *random-size graphemes* (where the size of the graph evolves randomly). In Section 1.2.3 we take a preview on general grapheme dynamics. In Section 1.2.4 we formulate relevant goals for general grapheme dynamics.

1.2.1 Graphemes

Let $\mathcal{I}^* = (\mathcal{I}, \tau)$ be a *Polish space*, with $\mathcal{I} \neq \emptyset$ a set and τ a topology, $\mathcal{B}_{\mathcal{I}}$ the Borel- σ -algebra of subsets of \mathcal{I} , h a $\mathcal{B}_{\mathcal{I}}$ -measurable symmetric $\{0, 1\}$ -valued *connection function* on $(\mathcal{I} \times \mathcal{I}) \setminus D_{\mathcal{I} \times \mathcal{I}}$, with $D_{\mathcal{I} \times \mathcal{I}}$ the diagonal of $\mathcal{I} \times \mathcal{I}$, and μ a *sampling measure* on \mathcal{I}^* . (For graphemes it is crucial that the connection function h is $\{0, 1\}$ -valued, which is *different* from graphons where $\tilde{h} \in \tilde{\mathcal{W}}$ is based on a function h that is $[0, 1]$ -valued.)

If we draw an \mathbb{N} -sample by using a *non-atomic* sampling measure μ , setting

$$(1.5) \quad X = (x_i)_{i \in \mathbb{N}}$$

drawn from \mathcal{I} according to $\mu^{\otimes \mathbb{N}}$, and consider the matrix

$$(1.6) \quad H = (h(x_i, x_j))_{(i,j) \in (\mathbb{N} \times \mathbb{N}) \setminus D_{\mathbb{N} \times \mathbb{N}}},$$

then we obtain a *countable graph* \mathcal{G} embedded in \mathcal{I} , of which the vertices are represented by X and the edges are represented by H , in the sense that the vertices are labelled by \mathbb{N} while vertices i and j are connected by an edge if and only if $h(x_i, x_j) = 1$ (see Fig. 3).

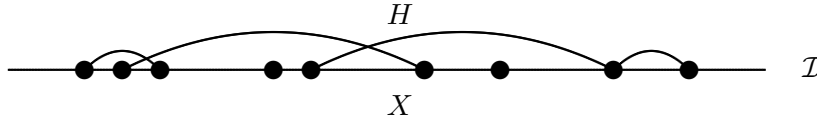


Figure 3: Representation of a graph as a grapheme: \mathcal{I} space, X vertices, H edges.

Note that without loss of generality we may assume that $\text{supp}(\mu) = \mathcal{I}$. Also note that graphs are automatically simple: no self-loops and no multiple edges are present. If we allow μ to have atoms, then we have to sample *without replacement* in order to preserve this property. The connection function can be trivially extended to $\mathcal{I} \times \mathcal{I}$ by setting $h = 0$ on $D_{\mathcal{I} \times \mathcal{I}}$, which we do henceforth.

Two cases are of interest: the set of vertices generated by the sampled infinite sequence for \mathcal{G} is

- (I) *deterministic* when μ is atomic;
- (II) *random* when μ is non-atomic and all sample points are different.

In case (I), the cardinality of the vertex set of \mathcal{G} equals $|\text{supp}(\mu)|$, which can be finite or countably infinite. In case (II), the cardinality is countably infinite. In case (II), a sampled sequence contains only different points and all sequences that we can sample are $\mu^{\otimes \mathbb{N}}$ -a.s. equivalent, in the sense that they are statistically indistinguishable: all finite samples have the same distribution and so the *equivalence class is deterministic*. The mixed case is of no interest, as we will see below.

For $n \in \mathbb{N}$, the n -sample $(x_i)_{i \in [n]}$ is drawn from \mathcal{I} according to $\mu^{\otimes n, \downarrow}$, i.e., without replacement, which generates a random finite graph \mathcal{G}_n . For $m \in [n]$, draw m vertices *uniformly at random without replacement* from the n -sample. Denote the distribution of its connection matrix by $\nu^{(m), n}$. As $n \rightarrow \infty$, $\nu^{(m), n}$ converges weakly $\mu^{\otimes \mathbb{N}}$ -a.s. to a connection-matrix distribution ν_m . Moreover, as $m \rightarrow \infty$, ν_m converges weakly $\mu^{\otimes \mathbb{N}}$ -a.s. (as a projective limit) to a *connection-matrix distribution*

$$(1.7) \quad \nu \in \mathcal{M}_1(\{0, 1\}^{(\mathbb{N} \times \mathbb{N}) \setminus D}),$$

which is the *characteristic object* describing the equivalence class of statistically indistinguishable countable graphs. Therefore, instead of $(\mathcal{I}, \tau), h, \mu$ labelling a grapheme, we might have taken $(\mathcal{I}, \tau), \nu, \mu$, because h is the random variable on $\mathcal{I} \times \mathcal{I}$ realising ν . Thus, in case (II), even though the graph generated by \mathcal{G} is random, its connection-matrix distribution ν is *deterministic*, contains all the information on the edges, and does not use properties of the embedding other than its existence. Note that if we have a diffusive and an atomic component, then the latter is statistically not visible in the countable graph, and ν just depends on the sample we draw from the diffusive part of the measure, because the graph is *simple*.

Note that ν determines all the subgraph densities of \mathcal{G} (recall (1.4)), as these are normalised expectations under the sampling distribution for samples of size $k \in [n]$, arising from the law of large numbers as the sample size tends to infinity. They also determine the graphon corresponding to the graph (see Figs. 1 and 4).

Definition 1.1 (Graphemes).

A grapheme is an equivalence class of triples (\mathcal{I}^*, h, μ) . ♠

Three concepts of equivalence classes will be considered in the sequel (recall (1.7)):

- $\langle \mathcal{I}^*, h, \mu \rangle$: Two triples (\mathcal{I}^*, h, μ) and $(\mathcal{I}'^*, h', \mu')$ are said to be equivalent when $\nu = \nu'$.
- $\{ \mathcal{I}^*, h, \mu \}$: Two triples (\mathcal{I}^*, h, μ) and $(\mathcal{I}'^*, h', \mu')$ are said to be equivalent when there is a measure-preserving *homeomorphism* $\varphi: \mathcal{I} \rightarrow \mathcal{I}'$ such that $\mu' = \mu \circ \varphi^{-1}$ and $h' = h \circ \varphi^{-1}$ μ -a.s.
- $[\mathcal{I}^*, h, \mu]$: The same as above, but for \mathcal{I}^* a Polish *metric* space and φ an *isometry*.

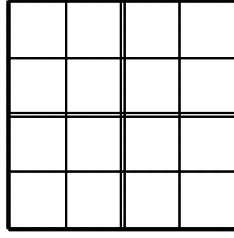


Figure 4: Refinements of equivalence classes: a set divided into subsets, each divided into further subsets.

Clearly, within the set of all admissible triples we have the following inclusions for the representatives:

$$(1.8) \quad [\mathcal{I}^*, h, \mu] \subsetneq \{\mathcal{I}^*, h, \mu\} \subsetneq \langle \mathcal{I}^*, h, \mu \rangle.$$

Consequently, the associated sets of equivalence classes $\mathbb{G}^{\square}, \mathbb{G}^{\{\}}, \mathbb{G}^{\langle \rangle}$ satisfy

$$(1.9) \quad \mathbb{G}^{\square} \ll \mathbb{G}^{\{\}} \ll \mathbb{G}^{\langle \rangle},$$

where \ll denotes *refinement*, i.e., an element of \mathbb{G}^{\square} induces a unique element of $\mathbb{G}^{\{\}}$, which in turn induces a unique element of $\mathbb{G}^{\langle \rangle}$ (see Fig. 4). On the other hand, properties of the spaces in which we embed are taken seriously only in $\mathbb{G}^{\{\}}, \mathbb{G}^{\square}$, and hence the inclusion is strict and so is the refinement. We use \mathbb{G} as a generic symbol when we just speak about graphemes, and add $\square, \{\}, \langle \rangle$ to specify the choice we make for the grapheme.

Remark 1.2 (Choice of spaces). Having both \mathbb{G}^{\square} and $\mathbb{G}^{\{\}}$ to work with resembles the situation in population models, where spaces of tree-like metric measure spaces and algebraic tree-measure spaces appear, which build on weaker equivalences than measure spaces [GPW09, LW21]. The space $\mathbb{G}^{\langle \rangle}$, which is the most general object, still corresponds to a subset of *countable graphs*, just like objects in the *completion* of finite *graphs* w.r.t. the topology of subgraph counts give rise to the space of graphons (see [Jan13]). ♠

- \mathbb{G}^{\square} is a Polish space and is the state space on which we want to construct our grapheme dynamics, because it allows us to obtain *strong Markov processes* with *regular path* properties solving standard martingale problems. The key point is that the information on the *space-time process* with values in $\mathbb{G}^{\langle \rangle}$, which arises as genealogies (as we explain in Section 1.3.3), provide the most natural and minimal choice of embedding in a Polish space necessary to control the space-time path processes of edges via a Markov process. The state space \mathbb{G}^{\square} has the advantage that it applies the machinery of stochastic analysis most efficiently and is fully understood, whereas the two other state spaces have some structural deficiencies and raise a number of *open questions* that remain to be resolved.
- On $\mathbb{G}^{\{\}}$ the nice process properties mentioned above may in principle fail (although we show that they do not for the classes of evolution rules we treat). This state space is conceptually important because it satisfies the *minimal structural requirements*

that we need when we want to embed our countable graph in a Polish space \mathcal{I} , and want to have a stochastic process for which the paths can be embedded in $C([0, \infty), \mathcal{I})$ or $D([0, \infty), \mathcal{I})$.

- The role of \mathbb{G}^\diamond is to link grapheme dynamics to graphon dynamics on $\widetilde{\mathcal{W}}$, because an element of $\widetilde{\mathcal{W}}$ is uniquely characterised by its subgraph densities of a representative (recall Section 1.1) and is an element of a Polish space. We will see that we can associate with our grapheme process a graphon process by taking a \mathbb{G}^\diamond -valued functional as long as the graphon can be interpreted as a countable graph, i.e., our graphemes always have $\{0, 1\}$ -valued connection-matrix distributions. This can result in block graphons, but also in graphons with not necessarily completely connected components. A stochastic process with values in \mathbb{G}^\diamond , which is part of a Polish space (the graphons) but is itself *not* a Polish space (only in a stronger topology), must therefore have paths of embeddings where the associated graphon is $\{0, 1\}$ -valued without exceptional points.

1.2.2 Extensions of graphemes

There are two important enrichments of the grapheme structure: vertices may have types that may influence the presence of edges, and the (scaled) number of edges may be random and changing in time.

Extension 1: Marked graphemes. The structure described above can be enriched by *types* drawn from a Polish space V , to obtain *V-marked graphemes* defined as equivalence classes

$$(1.10) \quad [(\mathcal{I} \times V)^*, h, \mu_{\mathcal{I} \times V}].$$

Here, $\mu_{\mathcal{I} \times V} = \mu \otimes \kappa$ is a sampling measure on $(\mathcal{I} \times V, \mathcal{B}_{\mathcal{I} \times V})$ with $\mathcal{B}_{\mathcal{I} \times V} = \mathcal{B}_{\mathcal{I}} \otimes \mathcal{B}_V$, where \mathcal{B}_V is the Borel- σ -algebra of subsets of V , and κ is a kernel from $(\mathcal{I}, \mathcal{B}_{\mathcal{I}})$ to (V, \mathcal{B}_V) . The type-space V is fixed and therefore equivalences have to be κ -preserving. The connection function now becomes a *connection-type function* on $(\mathcal{I} \times V)^2$ (the set of V -marked edges). This possibility is crucial in order to be able to exhibit certain properties of the graph evolution, as we will demonstrate. The state spaces $\mathbb{G}^\diamond, \mathbb{G}^\square, \mathbb{G}^\{\}$ in the *V-marked* case we denote by

$$(1.11) \quad \mathbb{G}^{\diamond, V}, \mathbb{G}^{\square, V}, \mathbb{G}^{\{\}, V}.$$

Extension 2: Random-size graphemes. To include the possibility of *random* vertex numbers instead of a fixed number n , we consider a new object as state, namely,

$$(1.12) \quad [\mathcal{I}^*, h, \mu], \quad \mu \in \mathcal{M}_{\text{fin}}(\mathcal{I}^*),$$

with $\mathcal{M}_{\text{fin}}(\mathcal{I}^*)$ the set of finite non-negative measures on \mathcal{I}^* , and write

$$(1.13) \quad \mu = \bar{\mu} \hat{\mu}, \quad \bar{\mu} = \mu(\mathcal{I}) \in (0, \infty), \quad \hat{\mu} \in \mathcal{M}_1(\mathcal{I}^*),$$

with $\mathcal{M}_1(\mathcal{I}^*)$ the set of probability measures on \mathcal{I}^* , to obtain a quadruple

$$(1.14) \quad (\bar{\mu}, [\mathcal{I}^*, h, \hat{\mu}]).$$

The state space of these quadruples is denoted by $\bar{\mathbb{G}}^\square$ (and, similarly, $\bar{\mathbb{G}}^\{\}$, $\bar{\mathbb{G}}^\diamond$).

Varying the vertex numbers per type, we consider V -marked graphemes and, for V finite or countable, get the quadruple

$$(1.15) \quad (\bar{\mu}_{\mathcal{I} \times V}, [(\mathcal{I} \times V)^*, h, \hat{\mu}_{\mathcal{I} \times V}])$$

with

$$(1.16) \quad \bar{\mu}_{\mathcal{I} \times V}(\cdot) = \mu_{\mathcal{I} \times V}(\mathcal{I} \times \cdot) \in \mathcal{M}(V), \quad \hat{\mu}_{\mathcal{I} \times V} = (\bar{\mu}_{\mathcal{I} \times V})^{-1} \mu_{\mathcal{I} \times V}, \quad \hat{\mu}_{\mathcal{I} \times V} \in \mathcal{M}_1((\mathcal{I} \times V)^*).$$

The state space of these quadruples is denoted by $\bar{\mathbb{G}}^{\square, V}$ (and, similarly, $\bar{\mathbb{G}}^{\{\}, V}$, $\bar{\mathbb{G}}^{\diamond, V}$).

1.2.3 Preview on general grapheme dynamics

We want to construct stochastic processes on $\mathbb{G}^\square, \mathbb{G}^\{\}, \mathbb{G}^\diamond$ that arise as limits of processes of embedded finite graphs when the number of vertices tends to infinity. We focus on \mathbb{G}^\square , which is the only candidate for a Polish space so that we can use martingale problems to characterise processes. The construction will be carried out in detail in Section 1.3 and worked out further in Sections 3–4. In view of the way we defined graphemes, the proper notion of convergence must be built on *convergence of sampled embedded subgraphs*. Finite-graph dynamics are *Markov pure-jump processes* that are strong Markov and Feller (i.e., continuous in the initial state). In order to get *limiting* processes that are also strong Markov and Feller, with *regular paths* (i.e., continuous or càdlàg) on Polish state spaces (but typically not σ -compact or locally compact state spaces), many standard techniques from stochastic analysis are not readily applicable. In fact, we need to work with a dense class of *equicontinuous test functions* on our state space, and this class needs to be countable, separating and measure determining (see Section 3).

A $\mathbb{G}^\{\}$ -valued process induced by a \mathbb{G}^\square -valued Markov process need *not* be Markov (recall Fig. 4). However, in what follows we will build \mathbb{G}^\square -valued processes with good properties such that, for all initial states in \mathbb{G}^\square , the induced process in $\mathbb{G}^\{\}$ is the unique Markov process solving a martingale problem on a subspace of $\mathbb{G}^\{\}$. A similar fact holds for the \mathbb{G}^\diamond -valued induced process. Note that the latter have to be realisable via countable graphs embedded in some Polish space, which is why we prefer to view them as a *functional* of the \mathbb{G}^\square -valued process that *automatically generates existence* of a concrete embedding. Otherwise we need to verify that the limiting subgraph counts correspond to a connection-matrix distribution arising in *some* arbitrary Polish space embedding. Note, however, that for martingale problems we need Polish state spaces, which is guaranteed on \mathbb{G}^\square for the environment component without further restrictions, but on $\mathbb{G}^\{\}$ by restricting to tree-like spaces, and on \mathbb{G}^\diamond by restricting to states that decompose into completely connected components (also dust would do, but that is too restrictive). We leave this open (see Section 3 for further details).

We want to push the view of the grapheme being represented by *equivalence classes* of countable *partitions of a continuum* via an embedding in a Polish space, and stochastic

processes in these equivalence classes. To pass to *equivalence classes of paths* is a topic of ongoing research. We will demonstrate that, for the class of finite-graph dynamics introduced in Section 1.3, for $t > 0$ we have completely connected components, which provides a concise description of the grapheme dynamics on \mathbb{G}^{\square} , and therefore is preferable to knowledge of the dynamics of the subgraph densities only. We show that a modification of this idea works even for *not* necessarily completely connected components, provided the definition of a component is properly modified. Moreover, the additional information about the embedding space can be used to obtain information about the *history* behind the subgraph densities at a given time as they evolve in time. This means that we get a hold on the space-time process, even the space-time path process, of the subgraph densities up to any given time (see Section 2.2 for further details).

In the weak topology of measures, the state space for the dynamics of the connection-matrix distribution (which reflects the subgraph densities) is compact and hence its law is tight. However, this state space includes as possible limit points the connection matrices $\underline{0}$ and $\underline{1}$, which are *improper* for many of our purposes because they are traps for our dynamics. Furthermore, limit points may lack the property to be a connection-matrix distribution of a sample sequence in a Polish space. We have to *exclude* both, and work on a smaller admissible state space that, unfortunately, no longer is closed. This means that we have to check *compact containment properties* for the paths of our stochastic processes.

Our approach opens up the possibility to choose the embedding in a way that is *adapted* to the dynamics. For dense finite graphs the dynamics typically involves changes that are *not* local. Therefore choosing the embedding is a subtle issue, even though its main role is to allow us to *sample* a countable graph in the $n \rightarrow \infty$ limit. We must therefore choose a stochastic process

$$(1.17) \quad (\mathcal{I}_t^*, h_t, \mu_t)_{t \geq 0}$$

i.e., the spaces we embed in are *random* and *time-dependent*. This will affect the definition of the topology on the state spaces \mathbb{G}^{\square} , $\mathbb{G}^{\{\}}$, since convergence also requires convergence of the spaces in which we embed now. This allows us to store enough information to compensate for the fact that the subgraph densities at a fixed time only reveal part of the grapheme: first, they only provide information that is *averaged* over the sample sequence, and second, they give no information on the embedding of the vertices. Note that, due to the form of the equivalence classes, without loss of generality we may assume that $\text{supp}(\mu_t) = \mathcal{I}_t$.

Remark 1.3 (Continuum limits). When dealing with countable limit objects we must assume that \mathcal{I} is a *continuum*. Our grapheme processes may start in any subset of \mathbb{G}^{\square} , including sets of non-dense graphs, but after any positive time they fall in the set of dense graphs. In the present paper we focus on *dense* graphs. For *non-dense* graphs, not treated in the present paper, an adaptation of the concept of a grapheme would be needed. ♠

1.2.4 Main goals for general grapheme dynamics

Our task is to verify the following (which we will work out in Sections 3–6 for the simple finite-graph dynamics given in Section 1.3):

- Let $\mathcal{G}^{(n)} = (\mathcal{G}_t^{(n)})_{t \geq 0}$ be a Markov pure-jump process based on a triple $([n], h, \mu)$ with $\text{supp}(\mu) = [n]$. As $n \rightarrow \infty$, after the finite graphemes have been suitably *embedded* in a continuum Polish space, this process converges to a limit process $\mathcal{G} = (\mathcal{G}_t)_{t \geq 0}$ in \mathbb{G}^\square .
- The limit process is described by a

$$(1.18) \quad (\widehat{\mathcal{L}}^*, \widehat{\Pi}^*, \widehat{\Gamma}^*)\text{-well-posed martingale problem,}$$

with $\widehat{\mathcal{L}}^*$ a *linear operator* on a measure-determining and convergence-determining subalgebra $\widehat{\Pi}^*$ of $C_b(\widetilde{\mathbb{G}}^\square, \mathbb{R})$ playing the role of a generator, with $\widehat{\Pi}^*$ an *algebra of polynomials* (to build in *sampling of finite substructures*) on \mathbb{G}^\square playing the role of a domain of test functions for $\widehat{\mathcal{L}}^*$, and $\widehat{\Gamma}^*$ an *initial law* on \mathbb{G}^\square . The solution is unique and defines a strong Markov process on \mathbb{G}^\square with the Feller property.

- The same holds for the functional $\nu = (\nu_t)_{t \geq 0}$, viewed as process on $\mathbb{G}^{\langle \rangle}$, via its own

$$(1.19) \quad (\mathcal{L}^*, \Pi^*, \Gamma^*)\text{-well-posed "martingale problem",}$$

with \mathcal{L}^* a linear operator on $C_b(\mathbb{G}^{\langle \rangle}, \mathbb{R})$, Π^* an algebra of polynomials on $\mathbb{G}^{\langle \rangle}$, and Γ^* an initial law on $\mathbb{G}^{\langle \rangle}$. The solution is unique and defines a Markov process on a subset of $\mathbb{G}^{\langle \rangle}$. However, proving existence of the solution requires getting a solution of (1.18) for some natural choice of \mathcal{I}^* and μ . Note here that $\mathbb{G}^{\langle \rangle}$ is only a measurable subset of a Polish space $\widetilde{\mathcal{W}}$ and is not closed in $\widetilde{\mathcal{W}}$. Therefore the "*martingale problem*" in (1.19) allows for many nice calculations, but is *not* of the standard form. This is only the case when we consider it on \mathbb{G}^\square , as a martingale problem of a functional, and after passing to a subspace. Otherwise, the tools available from stochastic analysis are severely restricted. To resolve this problem we must use the property that the dynamics immediately moves into a smaller space.

- The process $\nu = (\nu_t)_{t \geq 0}$ associated with \mathcal{G} , which arises also from the process $\nu^{(n)} = (\nu_t^{(n)})_{t \geq 0}$ associated with $\mathcal{G}^{(n)} = (\mathcal{G}_t^{(n)})_{t \geq 0}$ in the limit $n \rightarrow \infty$, may be either *autonomous* (i.e., not feel the refinement of the $\langle \rangle$ -equivalence class) or may evolve in the *random environment* given by (1.17) (i.e., feel the refinement). In the autonomous case, the *equilibria* of $\mathcal{G} = (\mathcal{G}_t)_{t \geq 0}$ are described by quantities arising from the distribution of the connection-matrix of sampled subgraphs of the finite graph dynamics.

1.3 Examples of finite-graph evolution rules and main theorems

The finite graphs evolve as *Markov pure-jump processes* inducing a grapheme $([n], h, \mu)$, where h is the connection matrix and μ is the sampling measure, with $\text{supp}(\mu) = [n]$. Write \mathbb{G}_n to denote the set of associated equivalence classes, and put $\mathbb{G}_\infty = \cup_{n \in \mathbb{N}} \mathbb{G}_n$ (which is needed when the number of vertices may vary). We refer to the elements as *finite graphemes*. The evolution rules below describe transitions for vertices and edges occurring at certain rates. In Section 1.3.1 we give the evolution rules on which we focus. In Section 1.3.2 we list our specified goals for these dynamics. In Section 1.3.3 we make

the link with the graphon dynamics presented in [AdHR21], define our grapheme process \mathcal{G} and lift our graph dynamics specified by the simple rules to a \mathbb{G}_n^{\square} -valued stochastic grapheme process. In Section 1.3.4 we state our main theorems with which we achieve the above listed goals.

1.3.1 Two classes of evolution rules from population dynamics

(I) We first focus on evolution rules for graphs with *completely connected components* (= all clusters or cliques are complete subgraphs). These have a version with a fixed number of vertices (Fleming-Viot) and a version with a randomly evolving number of vertices (Dawson-Watanabe), and each has a corresponding rule of immigration and emigration (McKean-Vlasov rule), both for fixed size and variable size. The evolution rules carry these names because for a population sizes tending to infinity these processes are the *diffusion limits* of the Moran dynamics and Galton-Watson dynamics, respectively.

We next specify the *rules* and the *initial states*.

► **The Fleming-Viot evolution rule.** At rate $d \in (0, \infty)$ pick a pair of vertices and perform the following transition:

- If the vertices belong to the same component, then do nothing.
- If the vertices belong to different components, then throw a fair coin to decide which vertex is the winner, respectively, the looser, remove all the edges of the looser and replace them by edges to all the vertices in the connected component of the winner.

► **The Dawson-Watanabe evolution rule.** At rate $b \in (0, \infty)$ pick a vertex and perform the following transition:

- Throw a fair coin. If the outcome is 1, then add a new vertex and add edges between the new vertex and all the vertices in the component of the chosen vertex. If the outcome is 0, then delete the chosen vertex and all its edges.

► **The McKean-Vlasov evolution rule.**

- *Fixed size:* At rate $c \in [0, \infty)$ pick a vertex, *remove* the chosen vertex and all its edges, and *add* a new vertex drawn from a source S of labels according to a probability measure θ on S (which is a Polish space), and *connect* it by edges to all the vertices with the same label. Note that for θ diffusive, the new vertex enters without edges, which leads to a simple rule: the new vertex becomes a new connected component and forgets its label. For atomic θ , on the other hand, a vertex is potentially added to an existing completely connected component, and the weights of the atoms in θ become relevant parameters.

- *Variable size:* Perform the addition and removal of vertices independently at rate c .

For FV (= resampling) plus McKV (= immigration and emigration) the number of vertices is preserved, while for DW (= birth or death) plus McKV the number of vertices varies. All four rules have the property that they preserve *complete connectedness*.

► **Possible extensions.** We can add further evolution rules with interesting new features. To do so, we give vertices a *type* drawn from a fixed type-set V . Then we can reformulate the rules in (I) by requiring edges to be present if and only if the two vertices share the same type, and in (II) by adding new rules for the addition and removal of edges.

It is possible to add a *bias of types* in the resampling mechanism. For instance, the fair coin in the FV-rule is replaced by a biased coin, with a bias that depends on the types of the pair of vertices chosen (= *selection*). Another option is to add to resampling in the marked case a change of type: a vertex receives a type drawn independently according to a probability measure $\theta_v \in \mathcal{M}_1(V)$ that depends on the type v of the chosen vertex (= *mutation*). It is also possible to evolve via resampling in larger sweeps. For instance, in the FV-rule a positive fraction of the vertices is chosen, rather than a pair of vertices, and all are connected to the winner (= *Cannings* resampling).

(II) We next focus on evolution rules for graphs with *non-completely connected components* (= not all clusters or cliques are complete subgraphs). Here we use the marked version of (I) with some new rules.

► **Reformulation of (I).**

- Grow the graph according to the FV-rule, with the modification that the loser adopts the type of the winner.
- Grow the graph according to the DW-rule, with the modification that the new vertex adopts the type of the chosen vertex.
- Grow the graph according to the McKV-rule, with the modification that the new vertex receives a type that is drawn independently according to a probability measure $\theta \in \mathcal{M}_1(V)$.

► **Addition of edges.** Apply the same rules as above. In addition, for each pair of vertices carrying the same type but having no edge between them, at rate $a^+ \in [0, \infty)$ add an edge.

► **Removal of edges.** Apply the same rules as above. In addition, for each pair of vertices carrying the same type and having an edge between them, at rate $a^- \in [0, \infty)$ remove the edge.

The latter two evolution rules do *not* preserve complete connectedness. We can run the evolution as in (I), but with the two rules above added (which are switched off when $a^\pm = 0$). The new Markov pure-jump processes allow for equilibria in which the components of a single type are random in number and in size, and in which edges evolve randomly.

Generalizations of the addition and removal mechanism, non-Markovian \mathbb{G}^\diamond -valued process. An interesting extension is to consider non-Markovian processes with values in \mathbb{G}^\diamond obtained by constructing a Markov process with values in \mathbb{G}^\square and letting a^+, a^- be *bounded measurable functions* of the state that depend on the age and on the current size of the subpopulation. This is treatable with the same methods, but generates a non-Markovian process on \mathbb{G}^\diamond . Note that age and subfamily size are observables in the current state of the genealogy, namely, distance and corresponding measure charged to a ball with a radius equal to the distance.

Initial conditions in (I) and (II). For finite graphs we can work with all the elements of \mathbb{G}_n or \mathbb{G}_∞ as initial condition. It is only in the $n \rightarrow \infty$ limit that we need to impose certain restrictions.

Remark 1.4 (Connection to [AdHR21]). Adding types in $V = [0, 1]$ to the underlying Fleming-Viot process and setting $h((i, v), (j, v)) = 1$ for $i, j \in U, v \in V$ and 0 elsewhere, we obtain a model with types and values in $\mathbb{G}^{\square, V}$. Now, compared to Fleming-Viot, further vertices are connected in the case of the Fisher-Wright rule, since they have different ancestors at time 0 but have the *same type* (recall that descendants have the type of the ancestor). This gives the *Fisher-Wright* process of [AdHR21], as opposed to the Fleming-Viot process we focus on here. \spadesuit

1.3.2 Main goals for grapheme dynamics arising from the evolution rules

Our goal is to use the evolution rules specified above for finite graphs to pass to finite grapheme evolutions that we use to build classes of (infinite) grapheme processes with the following features (sharpening the general goals formulated in Section 1.2.4):

- Choose a finite-graph dynamics based on the above evolution rules, enrich it to a *finite grapheme* evolution and, in the limit as $n \rightarrow \infty$, obtain a *strong* Markov grapheme process with the *Feller property*

$$(1.20) \quad \mathcal{G} = (\mathcal{G}_t)_{t \geq 0}$$

on $\tilde{\mathbb{G}}$ and $\tilde{\mathbb{G}}^{\square, V}$ (with $\tilde{\mathbb{G}} = \mathbb{G}, \bar{\mathbb{G}}$) with *continuous paths*, i.e., a *grapheme diffusion* driven by a *second-order operator* $\hat{\mathcal{L}}^*$ (compare with Remark 1.7). Project \mathcal{G} from $\tilde{\mathbb{G}}^\square, \tilde{\mathbb{G}}^{\square, V}$ to $\tilde{\mathbb{G}}, \tilde{\mathbb{G}}^{\diamond, V}$, to obtain a strong Markov *diffusion* on $\tilde{\mathbb{G}}^\diamond, \tilde{\mathbb{G}}^{\diamond, V}$ via the functional (we focus on \mathbb{G}^\diamond for simplicity)

$$(1.21) \quad \nu = (\nu_t)_{t \geq 0}$$

of connection-matrix distributions (equivalent to subgraph densities), driven by a *second-order operator* \mathcal{L}^* . In order to avoid confusion between a process that is a functional on \mathbb{G}^\square and a process that takes values in \mathbb{G}^\diamond obtained from a \mathbb{G}^\diamond -valued process \mathcal{G} , we write \mathcal{G}^\downarrow .

- Identify the *dual* process of \mathcal{G} and \mathcal{G}^\downarrow as a *coalescent* process.

- Identify the *equilibria* of \mathcal{G} and ν , respectively, \mathcal{G}^\downarrow as specific laws appearing in population genetics and statistics. Show, in particular, that these equilibria are non-trivial.

For the dynamics specified in Section 1.3, the existence of a *dual process* with *moment duality relations* is the key instrument towards establishing unique solutions of the martingale problems, convergence of finite-graph approximations, as well as existence and properties of equilibria.

1.3.3 From population dynamics to grapheme dynamics

In [AdHR21], stochastic processes in *population dynamics* were used to build a dynamics on the space of graphons, in particular, the Fisher-Wright diffusion and the measure-valued Fleming-Viot diffusion. The graph evolution rules given above also stem from population dynamics and lead to Markov processes that are already used in population genetics. The new idea is to take into account additional information, namely, the *genealogical structure* of the population, for which graphemes are the key tool. We will use graphemes to resolve two sets of issues left open in [AdHR21]: martingale problem characterisation, proof of the *strong Markov property* and the *Feller property*, and construction of dynamics with *non-trivial equilibria*.

To achieve this, we will exploit the framework of *duality* and *well-posed martingale problems*. For the latter, the key idea is to obtain *functions of the elements* in \mathbb{G}^* of the form $((\mathcal{I}^*, h, \mu))$, with $*$ and the outer bracket standing for $[], \{\}, \langle \rangle$, by taking a sample of size n according to μ , considering functions φ of graphs with n vertices and their edges, and choosing equal weights to get elements of \mathbb{G}_n . These φ are evaluated on the sample of size n from \mathcal{I} . We define an n -monomial on \mathbb{G} by taking as value of the function the n -sample expectation of φ (see Section 3.1). Such functions are called *polynomials* and will appear as test functions specifying the domain of the operator in the martingale problem.

In this section we define the grapheme process \mathcal{G} as a *functional* of a *genealogy-valued diffusion* U . In Section 1.3.4 we provide a characterisation via a well-posed martingale problem and via a limit of finite grapheme evolutions. The evolution rules described above (Fleming-Viot, Dawson-Watanabe, McKean-Vlasov) look at the *vertex population* exactly like the transitions for the population of individuals in the Moran, respectively, Galton-Watson branching model with immigration/emigration from a source, and are well known in population genetics. In particular, their $n \rightarrow \infty$ limit *measure-valued* diffusions have been established a long time ago see [Daw93, EK86]. The corresponding processes of evolving *genealogies* have been studied more recently. The evolution of these genealogies can be described as a Markov process and can be shown to converge in the limit as $n \rightarrow \infty$, as follows.

As explained in [GPW13], the *encoding of genealogies* of a stochastically evolving population (in the $n \rightarrow \infty$ limit) runs via

$$(1.22) \quad \text{equivalence classes } [U^*, \mu]$$

with $U^* = (U, r)$, U a set, r an ultrametric on U , μ a probability measure on U^* , and *equivalence* taken with respect to *measure-preserving isometries* of the *supports of the*

measure μ . This leads to a state space equipped with a *Polish topology* known as

$$(1.23) \quad \mathbb{U}_1 \text{ of genealogies, } \mathbb{U}_1^V \text{ of } V\text{-marked genealogies,}$$

respectively, with state spaces

$$(1.24) \quad \mathbb{U}_{\text{fin}}, \mathbb{U}_{\text{fin}}^V,$$

when μ is a *finite* measure, so that we have a *population size* $\bar{\mu} = \mu(U)$ and a *sampling measure* $\hat{\mu} = \mu/\bar{\mu}$.

The interpretation is as follows: U is the set of individuals, the metric is given by the genealogical distance between two individuals, i.e., twice the time back to the most recent common ancestor or $2t$ if there is none by time t (in particular, we have an *ultrametric*), while μ , respectively, $\hat{\mu}$ is the so-called *sampling measure*, allowing to draw samples of individuals for the vertices and hence also for the edges. Convergence in the topology of \mathbb{U}_1, \mathbb{U} means that the distributions of sampled *finite* subspaces converge (as finite ultrametric measure spaces). Recall that an ultrametric space has the geometric property that it can be decomposed in disjoint balls with an arbitrary radius that are open and closed.

We observe that we can turn our finite graph dynamics from Section 1.3.1, first into a \mathbb{G}_n - or \mathbb{G}_∞ -valued process and then into a \mathbb{G}_n^\square - or $\mathbb{G}_\infty^\square$ -valued *grapheme dynamics*, by defining an ultrametric r_t on $\{1, \dots, n\}$, by setting the distance equal to twice the time back to the most recent common ancestor of the two individuals and equal to $2t$ if there is none, thereby turning the graph into a grapheme. Here, \mathbb{G}_n^\square denotes the subspace of graphemes from \mathbb{G}^\square for which the sampling measure is concentrated on n points.

Basic idea. In the literature, on the spaces in (1.23) and (1.24), *diffusions* $(\mathcal{U}_t)_{t \geq 0}$ have been constructed describing genealogies for the classical measure-valued population models, called Fleming-Viot or Dawson-Watanabe diffusions, arising as limits of the genealogies in the Moran and Galton-Watson model, (see [GPW13, DGP12, DG19, GRG21]), which we will use later on.

The idea to study grapheme processes \mathcal{G} with values in \mathbb{G}^\square (as specified in Section 1.3.1) is to consider \mathbb{U}_1 -valued processes

$$(1.25) \quad \mathcal{U} = (\mathcal{U}_t)_{t \geq 0} \text{ with } \mathcal{U}_t = [U_t^*, \mu_t] = [U_t, r_t, \mu_t]$$

to get a space (U, r) to embed in our countable graph from the sample sequence.

Definition 1.5 (Grapheme processes associated with \mathbb{U} -valued diffusions \mathcal{U}).

(a) Consider the case where $\mathcal{U}_0 = [\{1\}, \underline{0}, \delta_1]$, the *single root case* (i.e., every vertex is on its own at time 0), from which we construct the general case by a simple operation, namely, by picking an operation \vdash on \mathbb{U} and taking $(\mathcal{G}_0 \vdash \mathcal{G}_t^*)_{t \geq 0}$, with \mathcal{G}^* the solution of the single root initial state (see the paragraph above (1.39)).

(b) At every time $t \geq 0$, choose a representative (U_t, r_t, μ_t) of \mathcal{U}_t . Define the triplet (viewed as a pre-grapheme)

$$(1.26) \quad (U_t^*, h_t, \mu_t) \text{ with } h_t(x, y) = 1 \text{ if and only if } r_t(x, y) \leq 2t,$$

called the grapheme process \mathcal{G} associated with \mathcal{U} , i.e., the connection-matrix is chosen so as to incorporate the genealogy. (The last restriction means that at time t there are edges exactly for the pairs of vertices with the same common ancestor at time 0.)

(c) The equivalence class $[(U_t, r_t, h_t, \mu_t)]$ depends on the equivalence class in \mathbb{G}^\square only via the equivalence class of the triple (U_t, r_t, μ_t) in \mathbb{U}_1 or \mathbb{U} , i.e., on \mathcal{U}_t , and defines a grapheme \mathcal{G}_t in \mathbb{G}^\square , and therefore defines a

$$(1.27) \quad \mathbb{G}^\square\text{-valued process } (\mathcal{G}_t)_{t \geq 0} \text{ for each process } \mathbb{U}_1\text{-valued } (\mathcal{U}_t)_{t \geq 0} \text{ prescribed.}$$

We may think of (1.27) as associating with every \mathbb{U}_1 -valued, respectively, \mathbb{U} -valued path a new path with values in \mathbb{G}^\square , respectively, $\bar{\mathbb{G}}^\square$. ♠

Remark 1.6 (\mathbb{G}_n - or G_∞^\square -grapheme dynamics). Note that if we take the version of the finite- n genealogical process, say \mathcal{U}^n , then *via the finite-graph dynamics introduced in Section 1.3.1* we obtain a version of the process \mathcal{G}^n by applying (1.26) and generating a process \mathcal{G}^n with values in \mathbb{G}_n^\square . ♠

The connection via genealogies with the \mathbb{U}_1 -valued dynamics allows us to exploit the following:

- The grapheme process $\mathcal{G} = (\mathcal{G}_t)_{t \geq 0}$ arises as a functional of the \mathbb{U}_1 -valued process $\mathcal{U} = (\mathcal{U}_t)_{t \geq 0}$, respectively, the \mathbb{U} -valued process, and can be characterised by a martingale problem.
- Standard tools can be used to show that \mathcal{G} has the *strong Markov* and the *Feller property*, and has *continuous* paths.
- *Convergence* of the finite grapheme processes \mathcal{G}^n to the continuum grapheme processes \mathcal{G} on \mathbb{G}^\square can be proved.
- The identification of the *dual* process as coalescent driven graphemes can be obtained as well.
- The *genealogical information* encoded in the space in which the graph is embedded can be used to get information on the process $\nu = (\nu_t)_{t \geq 0}$, inducing a process \mathcal{G}^\uparrow with values in $\mathbb{G}^{\langle \uparrow \rangle}$.
- Explicit formulas can be derived for *equilibria* in terms of certain classical key distributions in population dynamics and statistics.

1.3.4 Main theorems

The three theorems stated below concern the evolution rules in class (I) in Section 1.3.1. They are formulated in such a way that we can easily modify them to get the statements for *all* the models we discuss in this paper. Bringing the *genealogies* explicitly in focus, we can extend and generalise the population-dynamic approach advocated in [AdHR21].

The key objects appearing in the theorems will be defined, justified and explained in Sections 3–6, and rely on the theory of *genealogy-valued diffusions* built up in [GPW13,

DGP12, DG19], which we explain as we go along. The proofs of the theorems are given in Sections 5–8.

We need the concept of *order* for operators that are not classical differential operators.

Remark 1.7 (Order of operators). An operator Ω is called first order if $\Omega(\Phi^2) - 2\Phi\Omega\Phi = 0$ for all $\Phi \in D(\Omega)$ (= the domain of Ω), and is called second order if it is not first order and $\Omega(\Phi^3) - 3\Phi\Omega(\Phi^2) - 3\Phi^2\Omega(\Phi) = 0$ for all $\Phi \in D(\Omega)$. (This is the same algebraic characterisation as for first-order or second-order differential operators acting on twice-differentiable functions on \mathbb{R} .) See [DGP12] for more detail and examples. ♠

We consider the $n \rightarrow \infty$ limit dynamics arising from the evolution rules (I) for the $\mathbb{G}^\square, \mathbb{G}^\diamond$ -valued versions and for the V -marked $\mathbb{G}^{\square,V}, \mathbb{G}^{\diamond,V}$ -valued versions. We show below that this dynamics is the process \mathcal{G} alluded to earlier.

Initial states. We next introduce sets of *admissible initial states* for our grapheme evolutions, and *admissible initial laws* supported on these initial states. We introduce the sets of graphemes

$$(1.28) \quad \tilde{\mathbb{G}}_{\text{comp}}^*, \tilde{\mathbb{G}}_{\text{ultr}}^* \text{ with } * = \square, \{\}, \diamond \text{ and } \sim \text{ standing for 'with bar' or 'without bar',}$$

i.e., all graphemes with *non-degenerate* (= positive measure) *completely connected components* that form a subset where $h(i, j)$ is always 1 or 0 when i or j are not in the subset, respectively, with possible embedding in a specific class of Polish metric spaces, namely, those with an ultrametric, i.e., in an element of \mathbb{U}_1 or \mathbb{U} depending on \sim . Note that $\tilde{\mathbb{G}}_{\text{comp}}^*$ is an invariant set for our evolution mechanisms (except the one where we add insertion/deletion), but is not closed topologically. For $\tilde{\mathbb{G}}_{\text{ultr}}^\square$ we require h to be a *coarser ultrametric than r* . We note that $\tilde{\mathbb{G}}_{\text{ultra}}^\square$ is a Polish space that is dynamically closed and

$$(1.29) \quad \tilde{\mathbb{G}}_{\text{comp}}^* \subseteq \tilde{\mathbb{G}}_{\text{ultr}}^*.$$

Remark 1.8 (General initial states). The restriction of the initial state is needed but is harmless, at least for the classes of dynamics we consider (as we will show in Section 7 but not formulate here). From every $\mathcal{G} \in \mathbb{G}^\square$ that can occur as limit of our finite-graph dynamics the process enters at *infinite* rate into the subset of restricted states, jumping from $\tilde{\mathbb{G}} \setminus \mathbb{G}_{\text{comp}}$ to the dust state for h (preserving its $\tilde{\mathbb{G}}_{\text{comp}}^*$ component in the transition), and remains there. Indeed, only finitely many initial vertices have descendants at time $t > 0$, all initial edges have been replaced at time $t > 0$, and the process is in $\tilde{\mathbb{G}}_{\text{comp}}^*$. Only in such a state can edges be preserved for the future, since they appear for macroscopic sets of vertices (see Remark 1.12 below), as can be seen from the approximation with finite grapheme dynamics, or directly from the martingale problem by duality. Hence, we have here exactly the right state space for our dynamics. ♠

In Sections 3–6 we will explain at length the various ingredients that appear below in the statements on the grapheme diffusion.

Theorem 1.9 (Grapheme diffusions).

(a) The process \mathcal{G} defined in (1.27) and $\mathcal{G}_0 \in \widetilde{\mathbb{G}}_{\text{ultr}}^*$ is the unique solution of the

$$(1.30) \quad (\widehat{\mathcal{L}}^*, \widehat{\Pi}^*, \widehat{\Gamma}^*)\text{-martingale problem,}$$

where $\widehat{\mathcal{L}}^*$ is the extension of the operator \mathcal{L} corresponding to the \mathbb{U}_1 -valued processes, respectively, the \mathbb{U} -valued process lifted to \mathbb{G}^\square , $\widehat{\Pi}^*$ is the algebra of test functions on \mathbb{G}^\square , and $\widehat{\Gamma}^* \in \mathcal{M}_1(\mathbb{G}_{\text{ultr}}^\square)$ is the initial law. Furthermore:

- The process \mathcal{G} a.s. has paths with values in $\widetilde{\mathbb{G}}_{\text{comp}}^\square$ for $t > 0$ and $\widetilde{\mathbb{G}}_{\text{ultr}}^\square$ for $t \geq 0$.
- For positive times the grapheme connection function of a representative of \mathcal{G} a.s. is continuous and has completely connected components.
- The grapheme process \mathcal{G} has a dual (see Theorem 1.20).

(b) The $\mathbb{G}^{\langle \cdot \rangle}$ -valued process denoted \mathcal{G}^\uparrow , induced by the process $(\nu_t)_{t \geq 0}$ via the functional ν arising from the \mathbb{G}^\square -valued process \mathcal{G} , is the unique solution of the (compare here Remark 1.13)

$$(1.31) \quad (\mathcal{L}^*, \Pi^*, \Gamma^*)\text{-martingale problem}$$

on the closure of $\mathbb{G}_{\text{comp}}^{\langle \cdot \rangle}$, where \mathcal{L}^* is the operator induced by $\widehat{\mathcal{L}}^*$ on Π^* , which is a subalgebra of $\widehat{\Pi}^*$ of test functions identifying the equivalence class in $\mathbb{G}^{\langle \cdot \rangle}$ of \mathcal{G} , and Γ^* is the initial law induced by Γ .

(c) \mathcal{G} and \mathcal{G}^\uparrow are strong Markov processes with the Feller property (both w.r.t. their natural filtration) and with continuous paths, and $\widehat{\mathcal{L}}^*$ and \mathcal{L}^* are second order operators (i.e., \mathcal{G} and \mathcal{G}^\uparrow are diffusions). \square

Corollary 1.10 (State properties). $\mathcal{G}^{a,c,\theta}$ moves in positive time from every initial state (including sparse states) to a state that has finitely many ($c = 0$) or infinitely countably many ($c > 0$) completely connected components. In the former case, the number of components is random with a distribution that can be identified with the help of duality. \square

To understand the structure of ν_t we note the following. The distribution of the stochastic process of the *size-ordered* vectors $((\mu_t(A_i))_{i \in \mathbb{N}})$, with $(A_i)_{i \in \mathbb{M}}$ the completely connected components, only depends on the equivalence class of (U_t, r_t, h_t, μ_t) and can be characterised by the Fleming-Viot *measure-valued* diffusion on $\mathcal{M}_1([0, 1])$ at rate d , with emigration/immigration at rate c from a source $\theta \in \mathcal{M}_1([0, 1])$ (see [DGV95] and Section 1.4.1). To make this connection, it is convenient to think of the marked setting: each component is given a mark, drawn independently from $[0, 1]$ according to θ , and this type is inheritable and assigned at the immigration time. The weights of the types at time t uniquely specify an atomic measure, and hence an equivalence class of vectors corresponding to the unmarked case.

Remark 1.11 (Fisher-Wright, N -type Feller). If we modify the dynamic such that vertices have one of N fixed inherited types and are connected whenever two vertices have the same type, then we get N -type Fisher-Wright models or the N -type Feller branching model,

and we get models described by N -dimensional diffusions, to which our theory applies. However, these cases can be also treated in a simpler way, based on N -dimensional diffusion theory. Our approach is tailored for the cases with $N = \infty$. ♠

The process \mathcal{G}^\uparrow has a somewhat delicate mathematical structure and does not fit so well into the theory of general Markov processes. We explain why in two remarks.

Remark 1.12 (Initial states). The initial values appearing for \mathcal{G} on \mathbb{G}^\uparrow for ν_0 are the elements where we have finitely many, respectively, countably many components characterised by a probability measure on \mathbb{N} with a finite, respectively, countable support (without loss of generality chosen to be $\{1, 2, \dots, n\}$ or \mathbb{N}), or we have the grapheme given by $([0, 1], H \equiv 0, \text{Lebesgue measure})$, i.e., all vertices are on their own. In these cases we have *continuous paths*. If the process starts outside $\mathbb{G}_{\text{comp}}^\square$ and not in the single-root case, then it instantaneously jumps into the latter state at time 0. ♠

Remark 1.13 (\mathbb{G}^\diamond -martingale problem).

(1) The martingale problem, (a) on all of \mathbb{G}^\square , (b) on *all* of $\widetilde{\mathbb{G}}^\diamond$ or \mathbb{G}^\square , would be *non-standard*, since the space \mathbb{G}^\diamond is not closed and hence is not Polish. However, it is a subset of $\widetilde{\mathcal{W}}$, the space of graphons, which is a Polish space (even a compact Polish space) in the chosen topology. Therefore we have a legitimate stochastic process, taking values in a Polish space, but the martingale problem has to be extended to general graphons (which form a Polish space) in order to become a martingale problem of the standard form. We would need to show that the paths stay in an invariant subspace of \mathbb{G}^\diamond , which all the solutions of the martingale problem enter. Still, not all the standard conclusions from the theory of Markov processes need to hold. Through the h -component, we get similar effects in all of $\mathbb{G}^\square, \mathbb{G}^\{\}$, and the closure would be a space where h becomes a graphon, say, $\widehat{\mathcal{W}}^\square, \widehat{\mathcal{W}}^\{\}$.

(2) For our dynamics the peculiarity arises that, in any positive time, the path moves from points in \mathbb{G}^\diamond outside of $\mathbb{G}_{\text{comp}}^\diamond$ to points inside the *closure* of $\mathbb{G}_{\text{comp}}^\diamond$. The latter is only a small subset of the state space, but it turns out that this subset is entered because the cliques in finite graphs take over on *smaller* time scales. There is only one entrance law from states outside, namely, those with non-degenerate completely connected components with continuous paths, or starting in a point mass, and $h \equiv 0$ when the process starts outside \mathbb{G}_{comp} . In particular, we can conclude from our results that there is no extension of the entrance laws to all points in the graphon space that are not in \mathbb{G}^\diamond or not in $\mathbb{G}_{\text{comp}}^\square$, nor to the dust case $[\{1\}, \mathbf{0}, \mathbf{0}, \delta_1]$ with continuous paths, only càdlàg paths and instantaneous jumps at $t = 0$ into the closure of $\widetilde{\mathbb{G}}_{\text{comp}}$. Furthermore, we cannot get a standard martingale problem that for $t > 0$ has values in \mathbb{G}^\diamond , since the latter is not closed and the rate to jump is infinity off $\widetilde{\mathbb{G}}_{\text{comp}}$ or for the single-root case.

(3) The only option would be to introduce a stronger topology that makes \mathbb{G}^\diamond closed without interfering with the equivalence relation. This is done by requiring the limit to have a $\{0, 1\}$ -valued h , but then we still have to allow for *infinite* jump rates on part of the space in the martingale problem. One way to construct a solution despite these deficiencies would be to use the graphon space $\widehat{\mathcal{W}}$, which is compact, verify that the path of the process stays in \mathbb{G}^\diamond , and use the duality to construct the process (see [DGP23a]) by verifying that for $t \geq 0$ the path is in the closure of $\widetilde{\mathbb{G}}_{\text{comp}}^\diamond$. ♠

The process \mathcal{G} arises as a limit of finite-grapheme evolutions induced by the finite-graph evolutions of Section 1.3.1, which justifies its operator $\widehat{\mathcal{L}}^*$ as follows.

Theorem 1.14 (Grapheme approximations).

For $n \in \mathbb{N}$, let $\mathcal{G}^{a,c,\theta;(n)}$ (with $a = d$ or $a = bn$ for FV, respectively, DW), be the finite-grapheme process starting with n vertices and with a uniform sampling measure, and evolving according to the rules specified above. If, for $\mathcal{G}_0^{a,c,\theta} \in \mathcal{M}_1(\widetilde{\mathbb{G}}_{\text{ultr}}^{\square})$,

$$(1.32) \quad \text{LAW} \left[\mathcal{G}_0^{a,c,\theta;(n)} \right] \xrightarrow{n \rightarrow \infty} \text{LAW} \left[\mathcal{G}_0^{a,c,\theta} \right],$$

then

$$(1.33) \quad \text{LAW} \left[(\mathcal{G}_t^{a,c,\theta;(n)})_{t \geq 0} \right] \xrightarrow{n \rightarrow \infty} \text{LAW} \left[(\mathcal{G}_t^{a,c,\theta})_{t \geq 0} \right] \text{ in } C([0, \infty), \widetilde{\mathbb{G}}^{\square})$$

with $\widetilde{\mathbb{G}} = \mathbb{G}$, respectively, $\widetilde{\mathbb{G}}$, where $\xrightarrow{n \rightarrow \infty}$ denotes weak convergence in $D([a, \infty), \widetilde{\mathbb{G}}^{\square})$, the space of càdlàg paths. \square

Corollary 1.15 (General initial states). *If the process starts in \mathbb{G}^{\square} outside the closure of $\mathbb{G}_{\text{comp}}^{\square}$, then convergence still holds with an initial value given by the completely connected component part and given by the h -dust case (i.e., $h \equiv 0$) with the remaining weight, so that an instantaneous jump in h at time $t = 0$ occurs.*

Connection to the literature. We next address the question how Theorems 1.9 and 1.14 and Corollary 1.10 relate to earlier work in the *graphon literature*.

Remark 1.16 (Connection to graphons and [AdHR21]).

(1) For each $t > 0$ we can represent the state in \mathbb{G}^{\diamond} as a partition of $[0, 1]$, namely, the interval lengths under the law μ_t of the completely connected components ordered by size (up to equivalence). For $c = 0$ and for positive times, these evolve like a multi-dimensional diffusion. However, for $c > 0$, any attempt to get this as an $\ell_1(\mathbb{N})$ -valued diffusion fails due to the fact that there are *countably* many intervals, with countably many intervals either disappearing or being newly inserted during every positive time interval.

(2) The above embedding in $[0, 1]$ can be used to represent the induced $\mathbb{G}^{\{\}}\text{-valued}$ process via the space $([0, 1], \text{Euclidean distance, uniform distribution})$. However, this does *not* work for the \mathbb{G}^{\square} -process, because the embedding is not necessarily isometric.

(3) From the representation on $[0, 1]$ we can obtain the $\widehat{\mathcal{W}}$ -valued process of *graphons* as a *functional* of $(\mathcal{G}_t^{a,c,\theta})_{t \geq 0}$, written

$$(1.34) \quad (\mathcal{G}_t^{\text{graph},(a,c,\theta)})_{t \geq 0},$$

as a process with continuous paths and with a stationary distribution, arising as the limit of finite-grapheme processes. This process on $\widehat{\mathcal{W}}$ does *not* arise as the solution of a standard well-posed martingale problem with the strong Markov property and the Feller property and with continuous paths. Whether or not such a martingale problem exists is questionable for $c > 0$, while for $c = 0$ existence can be shown with the help of the so-called Petrov equations (see [Pet09]), generating the existence of a process of interval partitions. On a subspace of $\widehat{\mathcal{W}}$, for the graphons that are in $\mathbb{G}_{\text{comp}}^{\diamond}$, this can be done (see

also [GdHKW23]), but it cannot be extended to $\widehat{\mathcal{W}}$ in the standard form, as we argued above.

(4) For $c = 0$, the subclass of those \mathbb{G}^{\square} -valued processes corresponding to Fisher-Wright diffusions with immigration and emigration contains the processes given in [AdHR21]. ♠

Remark 1.17 (Dynamics of subgraph-counts for \mathbb{G}^{\diamond} -valued process). Even though the \mathbb{G}^{\diamond} -valued process $(\mathcal{G}_t^{a,c,\theta})_{t \geq 0}$ is uniquely determined by the equations following from the martingale problem, we can only conclude that it is a \mathcal{G}^{\diamond} -valued process because it arises as a functional of $(\mathcal{G}_t^{a,c,\theta})_{t \geq 0}$. Indeed, this guarantees that we have a $\langle \rangle$ -grapheme, i.e., there exists an embedding of the countable graph in some Polish space along the *whole path*. The question whether or not we could obtain the latter from properties of the subgraph-counts as paths, without reference to the \mathbb{G}^{\square} -valued process, has to be based on approximation and duality, and on an entropy criterion (see [GdHKW23]). ♠

Long-time behaviour and equilibrium. Next we discuss the *long-time behaviour*. Here the problem arises that the ultrametric spaces (U_t, r_t) degenerate because pairs of individuals without a most recent common ancestor have distance $2t$, which diverges as $t \rightarrow \infty$. This means that the equilibrium state may have distances equal to ∞ , and so we have to extend \mathbb{U}_1 and \mathbb{U} , or we have to transform the ultrametric r_t to $1 - e^{-r_t}$ (which is a *new ultrametric* for which distance ∞ becomes distance 1). Therefore we now pass to transformed states, which we indicate by writing LAW^+ instead of LAW . Note that, viewed in $\mathbb{G}^{\{\}}$ or \mathbb{G}^{\diamond} , our states do *not* change: they remain equivalent to the untransformed states.

Theorem 1.18 (Grapheme equilibria).

(a) The grapheme dynamics with $\mathcal{G}_0^{a,c,\theta} \in \mathbb{G}_{\text{ultr}}^{\square}$ converges to a unique equilibrium in $\mathcal{M}_1(\widetilde{\mathbb{G}}_{\text{ultr}})$, with \sim denoting bar or not bar, i.e.,

$$(1.35) \quad \text{LAW}^+ \left[\mathcal{G}_t^{a,c,\theta} \right] \xrightarrow[t \rightarrow \infty]{} \nu^{a,c,\theta},$$

with $a = d$ or b for FV , respectively, DW . For $c > 0$ this equilibrium is non-trivial and has completely connected components, while for $c = 0$ it is trivial and equal to δ_K with K the complete graph.

(b) The equilibrium on \mathbb{G}^{\square} , respectively, $\bar{\mathbb{G}}^{\square}$ is determined by the entrance law of the dual process, respectively, the conditional dual process, run for an infinite time (see Theorem 1.20 and Remark 1.22).

(c) The equivalence class of the frequency vector (w.r.t. permutations) of the different completely connected components is a random element in $\ell_1(\mathbb{N})$ whose law

$$(1.36) \quad \Gamma^{a,c,\theta}, \quad a = b, d,$$

can be computed (see Section 1.4.1) and uniquely identifies the law of the connection-matrix distribution in the equilibrium of \mathcal{G}^{\diamond} induced by ν on \mathbb{G}^{\diamond} . \square

In the sequel we need the following extension.

Corollary 1.19 (Marked version). The statements of Theorems 1.9, 1.14 and 1.18 also hold for the V -marked version of the processes on $\mathbb{G}^{\square,V}$, $\mathbb{G}^{\diamond,V}$. Here, the trivial equilibrium is $\int_V \theta(dv) \delta_{K_v}$ with K_v the complete graph whose vertices all have the same type $v \in V$. \square

Duality. The *dual process* of \mathcal{G} in the \mathbb{G} -valued case is indispensable to understand the behaviour of \mathcal{G} . This process lives on $\mathbb{G}^{\{\}}$, \mathbb{G}^{\square} , $\mathbb{G}^{\langle \rangle}$ and is *driven* by a *Kingman coalescent process* $C^t = (C_s^t)_{s \in [0,t]}$ with t the time horizon for the duality. This coalescent is a process of partitions of \mathbb{N} , where pairs of partitions merge at rate d , respectively, b , and jump to a cemetery S at rate c (in the cemetery coalescence is suspended), where a location is chosen according to $\theta \in \mathcal{M}_1(S)$. This coalescent leads to a sequence of simple Markov jump processes when \mathbb{N} is replaced by $\{1, \dots, n\}$, which form a *consistent family* in n . By a result of Kingman, we know that the entrance law from the state \mathbb{N} exists as a projective limit. We equip \mathbb{N} with the ultrametric by setting the distance between i and j to twice the time they need to enter the same partition element before time t , respectively, $2t$ if that does not happen by time t . Taking the completion to get an ultrametric space \tilde{U}_t , using the equidistribution on \tilde{U}_t (see [GPW09] for the construction of the \mathbb{U}_1 -valued coalescent), and using (1.26) to define h and taking equivalence classes, we get the *dual grapheme* $(\mathcal{C}_s^t)_{s \in [0,t]}$ process for time-horizon t and the dual grapheme \mathcal{C}_t^t by evaluating it at time t , written shortly as $\mathcal{C}_t \in \mathbb{G}^{\square}$.

In the case of $\bar{\mathbb{G}}$ -valued processes the situation is different. Here the dual has values in $[0, \infty) \times \mathbb{G}^*$, $*$ = \square or $\{\}$, $\langle \rangle$, where a number in $[0, \infty)$ given by an exponential functional of the coalescent path is added. We have the same dual process for the \mathbb{G}^* -component as before, and the $[0, \infty)$ -component is an exponential of a path integral of the functional of the coalescent, leading to a Feynman-Kac duality, and the collection of the finite dual processes is *no longer consistent*. At the end of Remark 1.22 we describe a way to circumvent this obstacle.

Furthermore, the *duality functions* H on $E \times E'$, where E, E' are the state spaces of the process, respectively, the dual process, are based on polynomials (and use *finite* coalescents) and on a Feynman-Kac potential $\mathfrak{V} : E' \rightarrow \mathbb{R}^+$, given by the number of dual pairs (see Section 6 for detailed definitions of H and \mathfrak{V}). This is a duality for the underlying genealogy process \mathcal{U} , as well as for the functional $\nu^{(\infty)}$ of connection-matrix distributions.

Theorem 1.20 (Duality).

(a) For $\mathcal{G}^{d,c,\theta}$ the following duality relation with duality function H and dual process \mathcal{C} holds:

$$(1.37) \quad \mathbb{E} \left[H \left(\mathcal{G}_t^{d,c,\theta}, \mathcal{C}_0 \right) \right] = \mathbb{E} \left[H \left(\mathcal{G}_0^{d,c,\theta}, \mathcal{C}_t \right) \right].$$

(b) For $\mathcal{G}^{b,c,\theta}$ the Feynman-Kac duality with potential \mathfrak{V} holds:

$$(1.38) \quad \mathbb{E} \left[H \left(\mathcal{G}_t^{b,c,\theta}, \mathcal{C}_0 \right) \right] = \mathbb{E} \left[H \left(\mathcal{G}_0^{b,c,\theta}, \mathcal{C}_t \right) \exp \left(\int_0^t ds \mathfrak{V}(\mathcal{C}_{t-s}) \right) \right].$$

□

Theorem 1.20 says that we can calculate the expectation in the left-hand side (which determines the law of $\mathcal{G}_t^{b,c,\theta}$ uniquely) in terms of the Markov jump process $(\mathcal{C}_t)_{t \geq 0}$, together with the distance growth that drives the grapheme process in the right-hand side.

In fact, we can strengthen this result to a *strong* duality in the Fleming-Viot case. Here, a subtlety arises because we need the concept of *pasting* of graphemes based on

pasting one ultrametric space to another one (symbolic \vdash) in order to modify $(\tilde{\mathcal{U}}_t, \tilde{r}_t)$ to incorporate *general* $\mathcal{U}_0 = [U_0, r_0, \mu_0]$ that are *different* from the *single-root case*. The pasted $\tilde{r}_t \vdash$ arises by adding r_0 and \tilde{r}_t , and taking $\tilde{r}_t \vdash (\iota, \iota') = r_0(\tau(\iota), \tau(\iota')) + \tilde{r}_t(\iota, \iota')$, where τ is the matching of \mathbb{N} with the μ_0 -sample sequence, respectively, its piece up to the number of partition elements at time t .

Corollary 1.21 (Strong duality).

$$(1.39) \quad \mathcal{L} \left[\mathcal{G}_t^{d,c,\theta} \right] = \mathcal{L} \left[\mathcal{G}_0 \vdash \mathcal{C}_t^{d,c,\theta,*} \right].$$

□

Remark 1.22 (Conditional duality). For $\mathcal{G}^{b,c,\theta}$ the situation is more subtle, but we get a similar structure when we condition on the full size process $(\bar{\mu}_t)_{t \geq 0}$. Conditioned on the latter path, the process of the \mathbb{U}_1 -valued component of the state can be represented by a coalescent with coalescence rate $b(\bar{\mu}_{t-s})^{-1}$ and jump rate $c(\bar{\mu}_{t-s})^{-1}$ at the backward time s of the coalescent (see [DG19] for more details). ♠

1.4 Consequences and extensions of main theorems

In Section 1.3.3 we defined $\mathbb{G}^{\{\},V}$ -valued and $\mathbb{G}^{\langle \rangle,V}$ -valued processes $\mathcal{G}^{c,d,\theta}$ and $\mathcal{G}^{c,b,\theta}$, which we abbreviate as $\mathcal{G}^{\text{FV}}, \mathcal{G}^{\text{DW}}$ in the sequel. We have the characterisation of their path properties and their long-time behaviour as stated in Theorems 1.9, 1.14 and 1.18. We proceed with some supplementary results. In Section 1.4.1 we give a detailed description of the *equilibrium process*, including a *phase transition* in the path behaviour with respect to the parameter c/a , where $a = b, d$, and extract the two key equilibrium statistics of completely *connected component's age and size*. Here the size depends on the functional $\mathcal{G}^{\langle \rangle}$ only, but the age depends explicitly on the $[U_t, r_t, \mu_t]$ -component, in particular, on r_t . In Section 1.4.2 we extend the results to *type-biased resampling* (referred to as *selection*) and *type-mutation*, leading to rewiring of edges. In Section 1.4.3 we turn to the dynamics without completely connected components (due to removal and insertion of edges).

1.4.1 Equilibrium processes

Geometric description of the equilibrium process: diffusing partitions. The embedding in \mathcal{I} helps us to better understand the *equilibrium process* of $\mathcal{G}^{a,c,\theta}$ with $a = b, d$ and $c > 0$, because there exists a decomposition into disjoint balls \mathcal{H}_t^i of *completely connected components* for any representative of \mathcal{G}_t of the equilibrium process $(\mathcal{G}_t)_{t \geq 0}$ such that the induced composition of $\mathcal{I}_t \times \mathcal{I}_t$ has the property

$$(1.40) \quad \mathcal{I}_t \times \mathcal{I}_t = \bigsqcup_{i,j \in \mathbb{N}} \left(\mathcal{H}_t^i \times \mathcal{H}_t^j \right), \quad h(u, u') = \begin{cases} 1, & u, u' \in \mathcal{H}_t^i, \quad i \in \mathbb{N}, \\ 0, & u \in \mathcal{H}_t^i, \quad i, j \in \mathbb{N}, i \neq j. \end{cases}$$

In other words, we have a *diffusion process* of (an equivalence class of) a *countable number of balls* that are completely connected in the grapheme sense and that carry weights charged by μ_t . If we size-order this vector of weights, then we obtain an object that

is a function of the equivalence class. This gives us a block-type representation of the grapheme, diffusing over time, from which we can read off the *graphon process* $(\widehat{\mathcal{W}}_t)_{t \geq 0}$ via the embedding of the grapheme in $[0, 1]$, as described in Remark 1.16.

Summarising, we look at the equilibrium path and its law, or the equilibrium law of the equivalence class of the weight vectors (represented by the *size-ordered version*):

$$(1.41) \quad \text{LAW} \left[\left((\mu_t(\mathcal{H}_t^i))_{i \in \mathbb{N}} \right)_{t \geq 0} \right], \quad \mu_t^\dagger = (\mu_t(\mathcal{H}_t^i))_{i \in \mathbb{N}}, \quad t \geq 0.$$

The *frequencies of edges processes* are given by $((\mu_t(\mathcal{H}_t^i))^2)_{i \in \mathbb{N}, t \geq 0}$.

Note that μ^\dagger determines $\nu^{(\infty)}$ uniquely. We show that in our two model classes the equilibrium $\mathcal{L}[\mu^\dagger]$ is given by the *Poisson-Dirichlet distribution* on the unit ball of $\ell_1(\mathbb{N})$, respectively, by a *standard Moran subordinator* on $\ell_1(\mathbb{N})$ introduced in the next paragraph.

Identification of the equilibrium. We are able to characterise the equilibria $\nu^{c,\theta,d}$ and $\nu^{c,\theta,b}$ in terms of the dual process introduced in Theorem 1.20, run for an *infinite* time (which is induced by a coalescent, as explained in Section 6). This dual process converges to a limit state \mathcal{C}_∞ as $t \rightarrow \infty$, and can be used to construct the equilibrium state of the associated diffusion $(\mathcal{U}_t^{a,c,\theta})_{t \geq 0}$ of genealogies, denoted by

$$(1.42) \quad \mathcal{U}_\infty^{a,c,\theta} \text{ with values in } \mathbb{U}_1, \text{ respectively, } \mathbb{U},$$

depending on whether we consider Fleming-Viot or Dawson-Watanabe dynamics.

In the *Fleming-Viot* case, this can even be done via a Kingman coalescent started in countably many dual individuals and state $(\{n\}, n \in \mathbb{N})$ run for an infinite time, for which we define a weighted ultrametric measure space giving an element of \mathbb{U}_1 via completion (also called strong duality). We can form the grapheme with that space by connecting all elements that are in the same partition element.

In the *Dawson-Watanabe* case, the representation is more subtle: we have to take the total mass and represent the normalised state, which can again be treated as a time-inhomogeneous Fleming-Viot process that has a representation via a time-inhomogeneous coalescent (see [DG19]).

In fact, for both we can identify the *law of* $\nu^{(\infty)}$ of the equilibrium, which determines what we can say about the *equilibrium in* \mathbb{G}^\diamond , and contains information on the equilibrium in $\mathbb{G}^\square, \mathbb{G}^\circ$. This is done by identifying $\mathcal{L}[\mu^\dagger]$ in (1.41): in the Fleming-Viot case with the help of the *Poisson-Dirichlet distribution*, in the Dawson-Watanabe case with the help of the *Moran-gamma-process*, which are both recalled below.

For convenience of notation we work on the state space $\mathbb{G}^{\square,V}$ (recall the paragraph after Remark 1.13). Consider the law of $\widehat{\Gamma}^{\text{FV},d,c,\theta}$ on $\mathcal{M}_1([0, 1])$ given by

$$(1.43) \quad \text{LAW} \left[\sum_{i \in \mathbb{N}} V_i \left(\prod_{j=1}^{i-1} (1 - V_j) \right) \delta_{(U_i)} \right],$$

where $(U_i)_{i \in \mathbb{N}}$ are i.i.d. with distribution θ and $(V_j)_{j \in \mathbb{N}}$ are i.i.d. with distribution $B(1, \beta^{-1})$ with $\beta = \frac{c}{d}$. The law $\widehat{\Gamma}^{\text{FV},d,c,\theta}$ is characterised as the unique equilibrium of a measure-valued process, the $\mathcal{M}_1([0, 1])$ -valued Fleming-Viot process with resampling at rate d

and with immigration and emigration at rate c from the source $\theta \in \mathcal{M}_1([0, 1])$ (for details see [DGV95, Section 2(a) and 2(b)]). The expression in (1.43) induces a vector $(V_i \prod_{j=1}^{i-1} (1 - V_j))_{i \in \mathbb{N}}$ of weights whose law is called the *Poisson-Dirichlet distribution* with parameter $(\theta, \frac{d}{c})$ on $\ell_1(\mathbb{N})$ and is denoted by $\Gamma^{\text{FV}, d, c, \theta}$. This is also the law of μ_t^\dagger in the equilibrium of $\mathcal{G}^{c, \theta, d}$, and characterises the law of the functional ν^∞ of the equilibrium in \mathbb{G}^\square , respectively, characterises the full equilibrium in \mathbb{G}^\diamond .

Corollary 1.23 (Identification of key statistics in the equilibrium $\nu^{d, c, \theta}$).

Assume that $c > 0$. Then, under $\nu^{d, c, \theta}$, the law of $\nu^{(\infty)}$ is given via the distribution $\Gamma^{\text{FV}, d, c, \theta}$ and determines the law μ^\dagger of the size-ordered weights of the completely connected components. \square

There is a similar identification in the Dawson-Watanabe case, based on the equilibrium of a $\mathcal{M}_{\text{fin}}([0, 1])$ -valued Dawson-Watanabe process on $[0, 1]$ with immigration and emigration at rate c from the source $\theta \in \mathcal{M}_{\text{fin}}([0, \infty)^2)$, whose equilibrium is based on the *Moran-gamma-subordinator* (see [DG96]). It is known that the multitype branching process with type space $[0, 1]$ and with immigration from the source $\theta \in \mathcal{M}_{\text{fin}}(\mathbb{R}^2)$ at rate c and emigration at the same rate c has a unique equilibrium $\widehat{\Gamma}^{b, c, \theta}$ of the form (see [DG96])

$$(1.44) \quad \text{LAW} \left[\sum_{i \in \mathbb{N}} W_i \delta_{(U_i)} \right],$$

where the $(W_i, U_i)_{i \in \mathbb{N}}$ are the *jump sizes* and *jump times* of the *Moran-gamma-process* up to “time θ ”, with $\theta = \theta([0, 1])$. This is an *infinitely divisible* $[0, \infty)$ -valued process with $\theta([0, 1])$ as time index, which we can characterise uniquely by specifying its *Lévy-measure* Λ on $[0, \infty]$. Indeed, define

$$(1.45) \quad \Lambda(dx) = x^{-1} e^{-\frac{c}{b}x} dx, \quad x \in [0, \infty),$$

and use this as the *Lévy-measure* of the infinitely divisible process, which is called the *standard Moran-gamma-process* on $[0, \infty)$ with parameter c/b .

Corollary 1.24 (Identification of key statistics in the equilibrium $\nu^{b, c, \theta}$). Under the law $\nu^{b, c, \theta}$ on \mathbb{G}^\square (respectively, $\mathbb{G}^{\square, V}$), the law $\mathcal{L}[\mu_t^\dagger]$ in (1.50) is given by $\Gamma^{b, c, \theta}$. \square

Phase transition for equilibrium paths: hitting complete graphs. The evolution of the frequencies of completely connected components in the equilibrium process has interesting features that can be captured by using the embedding of graphs and graphemes in a specific metric Polish space representing an equivalence class of an ultrametric measure space. In what follows we consider a statistic that describes a *property* that can be read off from the *functional* $\nu^{(\infty)}$ projected from $U \times V$ to U (and hence from the \mathbb{G}^\diamond -valued process), namely, the property that the $\ell_1(\mathbb{N})$ -valued path of $(\mu_t^\dagger)_{t \geq 0}$ (recall (1.46)) hits finitely many (or even a single) completely connected components.

Size-order $(\mathcal{H}_u^i)_{i \in \mathbb{N}}$ at a given time $u = s$, and observe the vector of weights at time $u = s + t$. If θ is diffusive, i.e., non-atomic, then for each entry there is a partition into

continua and their weights follow a diffusion process for which we can define *hitting times* by

$$(1.46) \quad (\mu_\infty(\mathcal{H}_{s+t}^i))_{t \geq 0} \text{ hits 0 at time } t = T_i,$$

and ask whether, for $i \geq 2$, $T_i < \infty$ a.s. or $T_i = \infty$ a.s. The radius of the (i)-most charged ball grows like $2t$ as t moves through $[0, \infty)$. At these exceptional times, if finite then the grapheme hits a state that has only finitely many components, namely, $i - 1$ completely connected components for $i \geq 2$.

We have a *phase transition* in the parameter c/a , with $a = b, d$, for the path property of the equilibrium path *to leave the state of countably many completely connected components and hit the state of a single completely connected component*.

Proposition 1.25 (Phase transition of path property: $\mathcal{G}^{a,c,\theta}$, $a = b, d$).

- (a) Depending on whether $c/d \geq 2$ or $c/d < 2$, $T_i = \infty$ a.s., respectively, $T_i < \infty$ a.s.
- (b) The connected component 1 may or may not hit the frequency 1. If yes, then $T_i < \infty$ a.s. for all $i \geq 2$. If not, then $T_i = \infty$ a.s. for all $i \geq 2$, and there are only states with a countably infinite number of different completely connected components. In the former case there is a positive length random time interval during which hitting the complete graph occurs infinitely often, but the time of countably many components is a union of time intervals (the excursions in the countable regime) with full Lebesgue measure. In the latter case, paths have countably many distinct completely connected components. \square

1.4.2 Extension: grapheme diffusions with mutation and/or selection of types

In this section we investigate what happens when we pass to the processes treated in Section 1.3.4, i.e., to evolution rules containing mutation and selection. In particular, we now consider $\mathbb{G}^{\square, V}, \mathbb{G}^{\diamond, V}$ for a suitable mark space V .

First we consider the case $m = 0$. We use the marked version of the Fleming-Viot evolution rule, i.e., we *connect vertices with the same type*, which characterises the most recent common ancestor of a connected component, and add *selection* at rate s with respect to types based on a fitness function χ (see (4.7) and the sequel for formulas). Then we can use the same formula as before to define $\mathcal{G}^{a,c,\theta,0,s}$ from the \mathbb{U}_1^V -valued process $\mathcal{U}^{a,c,\theta,0,s}$. We add *mutation* of types at rate m and jump probabilities $M(u, dv)$ for a mutation from u to v , where we distinguish between the cases where $M(u, \cdot)$ is diffusive, respectively, atomic. This addition requires serious changes, namely, we have to now define \mathcal{G} from \mathcal{U} by setting h equal to 1 for equal types, and we have to consider $\mathbb{G}_{\text{comp}}^{\square, V}, \mathbb{G}_{\text{ultra}}^{\square, V}$ based on the decomposition of $U \times V$ (rather than just U) in union of balls in the U -component and single values in the V -direction, i.e., a *marked subfamily decomposition*. With these adaptations the structure is similar.

In the cases $m > 0$ and $m = 0$, this gives us *marked grapheme diffusions* $(\mathcal{G}_t^{d,c,\theta,m,s})_{t \geq 0}$ and $(\mathcal{G}_t^{d,c,\theta,s})_{t \geq 0}$, for which results analogous to Theorems 1.9, 1.14 and 1.18 hold, based on the knowledge we have for the underlying \mathbb{U}_1^V -valued diffusions (see [DGP12]).

Theorem 1.26 (Grapheme diffusions with selection and mutation).

All the statements in Theorems 1.9, 1.14 and 1.18 hold for $\mathcal{G}^{d,c,\theta,m,s}$ when $m = 0$. For

$m \neq 0$, Theorems 1.9 and 1.14 hold too, but Theorem 1.18 holds if and only if the type-diffusion on V , arising by projecting $\mu_{\mathcal{I} \times V}$ onto V , converges to a unique equilibrium. \square

Remark 1.27 (Form of equilibria).

(1) No explicit form of the equilibrium law exists for $s, m > 0$ when mutation is not state-independent, in which case Gibbs measures appear as equilibria (see [DG14]). In the state-independent case the equilibrium is the same as immigration/emigration with $c = m$ and $\theta(dv) = M(u, dv)$ for all $u \in V$.

(2) For the case $s = 0, c = 0$ with $m > 0$ and M such that there exists a measure θ on the type space with $\theta M = \theta$, the equilibrium has a functional ν^∞ such that the equilibrium distribution is the Poisson-Dirichlet distribution in Corollary 1.23 when we replace c by m . Of special interest are the case where θ is a diffusive measure, for which we have countably many connected components with random weights, and the case where θ is finite atomic case, for which we have a finite number of such components. In both cases we have the explicit representation of this random structure from (1.43).

(3) For $s = 0, m > 0$ and $c = 0$ with θ having a countable number of atoms and no diffusive part, there are models related to the “infinite allele models” of population genetics, where equilibria for $\mathbb{G}^{\langle \cdot \rangle, V}$ -valued grapheme diffusions appear for the mass of completely connected components based on the Poisson-Dirichlet distribution. \spadesuit

Remark 1.28 (Duality). There is also a duality theory for models with selection and mutation via *function-valued* dual processes. The duality relation for the underlying process \mathcal{U} can be found in [DGP12]. *There is no simple strong duality.* \spadesuit

The grapheme diffusions with mutation and selection provide interesting classes of models because a different type of equilibrium appears, but also from a modelling point of view because building in types with different fitness is important to get different rates of expansions of the associated completely connected components. We do not pursue this in more detail here.

1.4.3 Grapheme diffusions with equilibria allowing for non-completely connected components

In this section we show that insertion and deletion of edges according to switches between virtual and true edges in the dynamics considered in Section 1.3.4 allows for *non-completely connected components in equilibrium*, and even for certain *non-Markovian* evolutions for the $\mathbb{G}^{\langle \cdot \rangle}$ -valued functional $(\nu_t)_{t \geq 0}$. We begin with the simplest case.

We *prune* the edges in the diffusions that we had in the previous model classes with completely connected components described in Section 1.3.4, to which we refer as the *background diffusion*, for which we have a unique associated solution \mathcal{U} at hand. For this candidate we derive a martingale problem. The modified operator of the martingale problem consists of the *second-order terms* considered in the main theorems and their extensions in the marked version, with additional *first-order terms* that account for the addition and removal of edges at rates a^+ , respectively, a^- in the finite graph evolution (recall Section 1.3.1).

We obtain the \mathbb{G}^{\square} -valued process \mathcal{G} as before from the \mathbb{U} -valued, respectively, \mathbb{U}^V -valued processes $\mathcal{U}^{\text{FV}}, \mathcal{U}^{\text{DW}}$, but only after pruning the edge-function h_t as defined before, in an

appropriate way via $\{0, 1\}$ -valued white noise (pruning). In other words, we first solve the martingale problems for $\mathcal{G}^{\text{FV}}, \mathcal{G}^{\text{DW}}$ from Section 1.3.4 of the form $([U_t, r_t, h_t, \mu_t])_{t \geq 0}$, which are processes with completely connected components, and then in the solution replace $(h_t)_{t \geq 0}$ by $(\tilde{h}_t)_{t \geq 0}$, where

$$(1.47) \quad \tilde{h}_t = \mathfrak{h}_t h_t \text{ pointwise}$$

with

$$(1.48) \quad \mathfrak{h}_t = U_t\text{-indexed } \{0, 1\}\text{-valued white noise,}$$

generated independently of the path of the underlying background diffusion. Remember, however, that the object arising in the grapheme is the *law of* h_t under the sampling measure μ_t . The *connection-matrix distribution* is still a *continuous* functional. However, (i) the operator changes; (ii) the properties of the states assumed by the path change as well. Nevertheless, the fact remains that the state decomposes into countably many ($c > 0$) or finitely many ($c = 0$) balls of points, which can potentially be connected and which diffuse. This decomposition is explicitly using \mathcal{U} . But the *effective* \tilde{h} is now a *white-noise pruning* of a *continuous function* h , and instead of completely connected components we now have only *path-connected* components (paths of edges built with \tilde{h} : the skeletons given by sample sequences decompose in that way).

The relation in (1.47) defines a $\mathbb{G}^{\square, V}$ -valued grapheme diffusion

$$(1.49) \quad (a^+, a^-) \mathcal{G}^{a, c, \theta} \text{ with } a = b, d \text{ or with } (d, s, m).$$

The grapheme has the property that the countable graphs arising from a typical sampling sequence of the grapheme have the feature that they are obtained by pruning the edges in a representative of the grapheme, i.e., pruning in the countable graph arising from the background grapheme by realising a $\mu^{\otimes \mathbb{N}}$ -distributed sample sequence. We thus find that, indeed, the evolution rules in class (II) in Section 1.3.1 can be included in our results.

We can obtain Theorems 1.9, 1.14, 1.18 and 1.26 after we adapt the martingale problem as indicated above, obtain a modified dual representation by the dual state of the background diffusion with pruned edges, and the identification of the equilibrium states by the dual of the underlying background diffusion. The grapheme again follows a *diffusion* in \mathbb{G}^{\square} , respectively, $\mathbb{G}^{\square, V}$.

Theorem 1.29 (Grapheme diffusion with non-completely connected components).

(a) *With the above modification (i) and (ii) of the operators and the state properties of paths, Theorems 1.9, 1.14, 1.18 and the extension in Theorem 1.26 hold also for the process \mathcal{G} in (1.49).*

(b) *There exists a process (the background diffusion) with completely connected components that produces the path-connected components of the process $(a^+, a^-) \mathcal{G}^{a, c, \theta}$, which for $a^+ \in (0, 1)$ are not completely connected. In particular, in every neighbourhood of a point u in the support of μ_t , the density of outgoing edges in equilibrium is $\frac{a^+}{a^+ + a^-}$, and for arbitrary states is given by the marginal of the Markov chain of the spin-flip system on $\{0, 1\}$ with rate a^+, a^- .*

(c) In equilibrium the connection-matrix distribution in equilibrium is an i.i.d. $\frac{a^+}{a^++a^-}$ -pruning of a block matrix of 1's with block-frequencies given by the Poisson-Dirichlet distribution given in (1.43) associated with the background diffusion.

(d) The duality relation of Theorem 1.20 holds for the dual process in (1.39) of the associated process with $a^+ = a^- = 0$ by taking an i.i.d. pruning of the edges with probabilities according to the marginal distribution at time t of the $\{0, 1\}^{\mathbb{N}}$ spin-flip Markov chain with rates a^+, a^- . \square

In the proof of Theorem 1.29 we formulate a *conditional* martingale problem for a given path of the background process that is the solution of the martingale problem of Theorem 1.9. Note that the associated graphon $\widehat{\mathcal{W}}_t$ consists of blocks of fluctuating sizes specified by their weights under μ_t , but the 0's and 1's occur with frequencies determined by the a_+, a_- , and by the time the system has evolved into equilibrium equal $\frac{a^+}{a^++a^-}$. Therefore the component of a single type is *not* completely connected, only *path-connected*. Recall that the latter holds also on a *countable Erdős-Rényi random graph*. Nevertheless, the graphon has values in $\{0, 1\}$ and therefore still agrees with the grapheme from our \mathbb{G}^{\square} -valued process projected on \mathbb{G}^{\diamond} .

Theorem 1.29 also holds for the extensions to the generalisations given in Section 1.3.1. This holds literally for part (a). Parts (b)-(d) have to be adapted as follows. In (b) we now have a *time-inhomogeneous* spin-flip system in a *random environment* through the dependence on the underlying genealogy process. In (c) the pruning of the equilibrium is now given by random variables $A^+ = a^+(\mathcal{U}_\infty)$ and $A^- = a^-(\mathcal{U}_\infty)$ with \mathcal{U}_∞ the equilibrium genealogy. Part (d) changes analogously to (b).

Corollary 1.30 (Extension). *The modified theorem above also holds for the more general flip rates defined in Section 1.3.1.*

1.4.4 Explicit representations of laws of equilibrium statistics

There are statistics for graphemes that can be explicitly calculated because they are the laws or important characteristics in models from population dynamics. Return to the models described in Section 1.3.4. We next introduce some quantities that concern the space-time structure of a current completely connected component.

Density of edges in equilibrium. For the equilibrium distribution on \mathbb{G}^{\diamond} or $G^{\diamond, V}$ in the Fleming-Viot case, the equilibrium induced by $\nu^{\diamond, FV, c, \theta, d}$ is concentrated on states that decompose into completely connected components, with vertex-frequencies that in *size-ordered* form are given by

$$(1.50) \quad (W_i^{\geq})_{i \in \mathbb{N}}$$

with W_i given by the Poisson-Dirichlet distribution, i.e., the *size-ordered vector*

$$(1.51) \quad \left(V_i \prod_{j=0}^{i-1} (1 - V_j) \right)_{i \in \mathbb{N}},$$

which is called the $GEM(\frac{c}{d})$ -distribution. A key statistics is the frequency of edges among all possible edges:

$$(1.52) \quad \bar{E} = \sum_{i \in \mathbb{N}} W_i^2.$$

This object describes the *relative density* of edges, and measures the deviation from completeness of the graph.

A similar object arises in the Dawson-Watanabe case, where as additional objects we have the *scaled limiting vertex number* and the *vertex numbers of the connected components*,

$$(1.53) \quad (W_i)_{i \in \mathbb{N}},$$

where $\sum_{i \in \mathbb{N}} W_i < \infty$ is a γ -distributed random variable with parameter $(1, \frac{c}{b})$. Define the analogue of the quantity in (1.52) as

$$(1.54) \quad \bar{E} = \sum_{i \in \mathbb{N}} \left(\frac{W_i}{\sum_{j \in \mathbb{N}} W_j} \right)^2.$$

The challenge is to say more about \bar{E} using (1.51) and (1.53).

Decomposition of edges in equilibrium: space-time structure. We now come to properties of the equilibrium process, which depend on the *embedding in an ultrametric measure space of genealogies*. In that space, we can define clusters (or cliques) of the set of edges arising from the time- t descendants of a single immigrant, define the completely connected corresponding components, their size at time t , and the times of their formation some random time back in the past.

(i) *Decomposition w.r.t. immigration time.* We can decompose the state in the space in which the graph is embedded at time t into the connected components coming from a *specific immigrant* (recall that we are in equilibrium, i.e., in the $t \rightarrow \infty$ limit where all individuals descend from one immigrant only). Each component has a finite diameter, i.e., the supremum of pairwise distances, ordered by size:

$$(1.55) \quad s_1 > s_2 > \dots$$

The vertices of these components are the descendants of the *immigrants* at times $t - s_1, t - s_2, \dots$, which have frequencies $\mu_t^{(1)}, \mu_t^{(2)}, \dots$. Note that s_1, s_2, \dots are realisations of *random variables* S_1, S_2, \dots . In particular, all edges that are present at time t date back to an origin later than

$$(1.56) \quad \text{time } t - S_1, \text{ and at time } t \text{ the frequency of edges is } \sum_{i \in \mathbb{N}} (\mu_t^{(i)})^2.$$

In other words, $t - S_1$ is a *renewal point* for the edge structure and $1 - \sum_{i \in \mathbb{N}} (\mu_t^{(i)})^2$ is the *deviation from completeness*.

We want to determine the law of the *ages* and *sizes* of the *completely connected components* in *equilibrium*:

$$(1.57) \quad \mathcal{L} \left[(t - S_i, \mu^{(i)})_{i \in \mathbb{N}} \right], \quad \mathcal{L} \left[(t - S_i)_{i \in \mathbb{N}} \right], \quad \mathcal{L} \left[\sum_{i \in \mathbb{N}} (\mu_t^{(i)})^2 \right].$$

The best way to analyse these objects is to work with two approaches, each giving explicit expressions as follows. In the *backward view* the coalescent driven dual process of the measure-valued process on \mathbb{N} defined by (1.41) is a coalescent with jumps into a cemetery at rate c and coalescence at rate d , with the cemetery state carrying the measure θ . In the *forward view* we use the excursion theory of this measure-valued diffusion, which is a k -dimensional diffusion with a random k , starting from the zero-state.

(ii) *Decomposition of edges in completely connected components w.r.t. the time of their origin.* We can label the completely connected components $k \in \mathbb{N}$ by using the time of their appearance. In each completely connected component we can decompose the edges according to their *origin in time*. Namely, we decompose the component immigrated at time $t - s_k$ for some $k \in \mathbb{N}$ into disjoint balls as follows. First, we take the ball of maximal radius, which is charged by $\mu_t^{(k)}$, and decompose it into disjoint balls of maximal radius less than s_k , say s_k^1 . After that, we continue with the complement within the time s_u -ball, to get successively s_k^2, s_k^3, \dots . We denote the corresponding random variables for the k -th component by $(S^{k,j})_{j \in \mathbb{N}}$. For every $k \in \mathbb{N}$ and every realisation this gives decompositions of the edges in the component dating back to time $t - s_k^2, t - s_k^3, \dots$. As weights $\mu^{(k)}, \mu^{(k,1)}, \mu^{(k,2)}, \dots$ for the latter random variables we use $M^{k,j}$ as notation. The frequencies of the edges present after $t - s_k^j$ are obtained by connecting vertices between the s_k^j -balls and using weights $\mu^{(k,j)}$, $k \in \mathbb{N}, j \in \mathbb{N}$. The frequencies of the edges are given by the product of their weights, namely, $\mu^{(k,j)}(1 - \mu^{(k,j)})$. Hence the *space-time structure* of the *edges in the k -th component* for $k \in \mathbb{N}$ is described by the collection of random variables with values in \mathbb{R}^+ , respectively, $[0, 1]$, which we can define for the \mathbb{G}^{\square} -valued grapheme:

$$(1.58) \quad (S^{k,j}, M^{(k,j)})_{j \in \mathbb{N}}, \quad (S^{k,j}, M^{(k,j)}(1 - M^{(k,j)}))_{j \in \mathbb{N}}.$$

The second components are also observables for the \mathbb{G}^{\diamond} -valued process, but not the first.

This leaves us with the task to calculate the laws for the objects defined in (1.57) and (1.58) based on duality and excursion theory of diffusions, which is not addressed here.

1.5 Relation to the graphon literature

Since in [AdHR21] the graphon dynamics was based on the Fisher-Wright dynamics, which is closely related to our Fleming-Viot dynamics, it is worthwhile to connect the statements in [AdHR21] to what we obtained via graphemes. In our framework the parallel process arises if initially in our Fleming-Viot dynamic $(\mathcal{G}^{d,0,0})_{t \geq 0}$ we assign inheritable types 0 or 1 to our types and connect all the completely connected components carrying 1 with edges to all the other vertices, so that a random number of completely connected components are *merged* into a single component.

Remark 1.31 (Connection matrices and density representation).

(1) The choice of embedding in a Polish space suggested by the literature on graphons is obtained by putting $(\mathcal{I}, \tau, \mu) = ([0, 1], d, \mu)$ with d the Euclidean distance and μ the uniform distribution. A measurable function H can be read off in \mathcal{W} , the space of h -functions defined in Section 1.1. Namely, we can identify a graphon, which is an element of the space $\widehat{\mathcal{W}}$ of equivalence classes of h -functions, with the property that there exists a measure Γ on $[0, 1]$ such that, for a symmetric $[0, 1]$ -valued function H (representing the graphon function),

$$(1.59) \quad \Phi([\mathcal{I}^*, h, \mu]) = \int_{[0,1]^m} \Gamma^{\otimes m}(d\underline{x}) \prod_{i,j \in [m], i \neq j} H(x_i, x_j),$$

i.e., Φ is the ‘all connected monomial’. This object relates to our concept of grapheme $[\mathcal{I}^*, h, \mu]$ as follows. If the unique H is *not* $\{0, 1\}$ -valued, then $([0, 1], \text{Borel}, H, \Gamma)$ does *not* relate to a grapheme in \mathbb{G}^\diamond , so graphons arising from graphemes can converge to a limit graphon without the graphemes converging themselves, since there is no space in which a *countable graph* can be realized having the limiting subgraph counts.

(2) Conversely, we obtain graphons as *functionals* of graphemes. Indeed, for the cases treated in this paper we can consider the measure $(\mu_t)_{t \geq 0}$, take the weights of the completely connected components, giving the vector $(\bar{\mu}_t(i))_{i \in \mathbb{N}}$, size-order this vector, place it on the interval $[0, 1]$, and follow the separation points in time. If (and only if) $c = 0$, then for the Fleming-Viot evolution rule this gives paths of embedded *nested multi-type Fisher-Wright diffusions*, whose number is finite and random, solving the Petrov equations [Pet09]. This specifies a process of graphemes based on $(([0, 1], d), h_t, \mu_t)_{t \geq 0}$, with d the Euclidean metric, where we can read off $(h_t)_{t \geq 0}$ from the diffusions, as well as a process of graphons with values in $\widehat{\mathcal{W}}$. In this way we can formulate another martingale problem, which produces a grapheme dynamics with the same marginal distributions as our \mathbb{G}^\diamond -valued and \mathbb{G}^\square -valued dynamics, but *different* marginals on \mathbb{G}^\square . Nevertheless, in both cases strong Markov processes result.

(3) If $c > 0$, then we get a *countable* set of partitions, we *cannot* follow the simple procedure of (b) above because there are countably many reshufflings occurring to restall the ranking, and we do *not* get a nice process of graphemes in \mathbb{G}^\square based on $(([0, 1], d), \tilde{\mu}_t)_{t \geq 0}$, as we explain below. We have to bring in our genealogical structure to resolve the open issues, as demonstrated in our theorems. ♠

Alternatively, we could connect our grapheme diffusion to the graphon diffusion constructed in [AdHR21], by noting that we can view H_t in (1.59) as densities giving $\mu_t^{\otimes m}$ when μ is the Lebesgue measure on $[0, 1]$. Then H_t generates the sampling of vertices and hence of edges, and thus produces a *connection matrix* h_t that is based on the embedding in $[0, 1]$. However, the evolution of the resulting h_t is *not* interpretable in terms of the underlying population dynamics. The function H_t is a certain *average* over the state h_t , giving the connections *based on the Euclidean metric*. If we would take the ultrametric induced on a subset of $[0, 1]$ and choose the same embedding of the finite graphs in our construction, then we would get different elements in \mathbb{G}^\square and \mathbb{G}^\square , but the same element in \mathbb{G}^\diamond . To get a grapheme dynamics in \mathbb{G}^\square or \mathbb{G}^\square we would need to prove that, for the embedding in [AdHR21], the convergence takes place in the sense of \mathbb{G}^\square and \mathbb{G}^\square . This is difficult because the spaces for embedding are *not* well adapted to the dynamics.

In [AdHR21], graphons in $\widehat{\mathcal{W}}$ describe the continuum evolution as the evolution $(\mu_t)_{t \geq 0}$, which defines a \mathbb{G}^\square -valued dynamics, and is only necessary due to the choice $\mathcal{I}^* = ([0, 1], \mathcal{B}_{[0,1]})$ for all t , which is deterministic. Both this dynamics and our dynamics have the same \mathbb{G}^\diamond -valued functional. The deterministic choice of \mathcal{I}^* restricts the possibilities of getting a manageable Markovian dynamics. In our approach we choose to work with graphemes instead of graphons, getting proper evolutions of countable graphs.

Another way to get a connection with the strategy in [AdHR21] is via the embedding constructed on [AdHR21, page 47] on $[0, 1]^2$ in Euclidean space. Getting an analogue for our embedding in $(\mathcal{U}_t)_{t \geq 0}$ works whenever we have a *skeleton* and an *ultrametric* r on \mathcal{I} such that we can use small balls to cover the edges, count these small balls, and set the normalised number equal to the value of H_n of the size- n graph approximation of the grapheme. The limiting H we get in this way depends on the choice of (\mathcal{I}, r) due to the implicit local averaging, and the same is true for $\widehat{\mathcal{W}}$ with the Euclidean distance in $[0, 1]^2$. The H thus constructed acts as a graphon and is a block function in our examples, except the one with insertion and deletion, where $H \equiv 1$ with "frequency" $\frac{a^+}{a^+ + a^-}$ on a sequence of blocks, and $H \equiv 0$ on the complement of these blocks, where the block masses are fluctuating in time, which is different from the Erdős-Rényi grapheme. Furthermore, this H is a *functional* of our dynamics and is embedded in a *richer* structure.

1.6 Outline of the remainder of the paper

In Section 2 we review what we have achieved and what remains to be done. In Section 3 we prepare the ingredients that are needed to formally and rigorously define Markovian grapheme evolutions, in particular, the *state space* and the *topology*. In Section 4 we specify the associated *martingale problems*, in particular, the algebra of test functions and the *operator* of the martingale problem playing the role of a generator. Sections 5 and 7–8 are devoted to proofs. In Section 6 we introduce the *dual process* and the *duality function* in a restricted setting that will be expanded in Section 7.

2 What have we achieved? What remains to be done?

In this section we reflect on what we have achieved in Section 1. In Section 2.1 we explain what is behind the choice of the spaces \mathcal{I}^* or (\mathcal{I}, r) , and what are the guiding principles for that choice when we are interested in *stochastic processes* taking values in \mathbb{G}^\square , $\mathbb{G}^\{\}$, \mathbb{G}^\diamond : $\mathbb{G}^\{\}$ is the state space that is needed to obtain an *embedded countable graph* that we can follow as time evolves, \mathbb{G}^\diamond keeps the structure of the embeddable countable graph without fixing an equivalence class of embeddings, while \mathbb{G}^\square provides the most manageable setting that allows for a detailed study of the corresponding process of countable graphs by means of stochastic analysis. We address the question how to choose $[\mathcal{I}^*, h, \mu]$ when we start from a given (finite) graph dynamics on \mathbb{G}_n or \mathbb{G}_∞ . The initial state plays a special role. For $t = 0$ this is an arbitrary state, whereas for $t > 0$ the dynamics typically moves into subsets of $\mathbb{G}^\square, \mathbb{G}^\{\}$ of a special form. In Section 2.2 we emphasise the intrinsic role of genealogies, and list some challenges and perspectives.

2.1 Conclusions for the issue of embedding

Representability as a grapheme and dynamics in random environment. For finite graphs the natural embedding set is $\mathcal{I}_n = [n]$, with a probability measure μ_n to sample points. Given a sequence of such finite graphemes and a corresponding sequence of connection-matrix distributions, we want to let $n \rightarrow \infty$, i.e.,

$$(2.1) \quad [\mathcal{I}_n, h_n, \mu_n] \longrightarrow (\nu_n^{(m)})_{m \in \mathbb{N}}, \quad n \rightarrow \infty,$$

such that the monomials evaluated at both sides converge, and there is a limiting sequence

$$(2.2) \quad (\nu^{(m)})_{m \in \mathbb{N}} \text{ of connection-matrix distributions.}$$

From this object we can sample a connection matrix, but what we really want is to *sample subgraphs*, in particular, *typical countable graphs*. For the latter we need to have an *embedding* of the finite graphs in \mathcal{I}^* or (\mathcal{I}, r) . From the embedding we know that the typical $\mathbb{N} \times \mathbb{N}$ matrix with entries 0 or 1 is the connection matrix of a typical sequence of vertices sampled from $\mu^{\otimes \mathbb{N}}$ when μ is diffusive.

However, we can often make better choices for \mathcal{I}_n than $[n]$ when we want to prove *convergence* of a sequence to an informative *uncountable Polish space as grapheme limit*. Therefore we must use an embedding that fits with the expected limit, meaning a finite version of that limit as far as the space we embed in concerns. We saw how this works for our specific examples. Note that the $\mathbb{N} \times \mathbb{N}$ connection-matrix distribution $\nu^{(m)}$ is a distribution on a compact space. Hence $(\nu_n)_{n \in \mathbb{N}}$ is a *tight sequence* of distributions. However, this is *not* necessarily true for $[\mathcal{I}_n^*, h_n, \mu_n]_{n \in \mathbb{N}}$ (which would imply that we can *sample points* and generate a *typical grapheme* embedded in some Polish space), since it need not necessarily hold for $[\mathcal{I}_n^*, \mu_n]_{n \in \mathbb{N}}$.

If we have a limiting object, say (\mathcal{I}^*, h, μ) , then we can generate a finite sampled sequence, respectively, an infinite sampled sequence. Indeed, take a sequence $(i_k)_{k \in \mathbb{N}}$ by sampling i.i.d. according to μ . Let

$$(2.3) \quad \underline{\hat{h}} = (h(i_k, i_\ell))_{k, \ell \in \mathbb{N}},$$

be the *empirical connection matrix* for the *skeleton* $(i_k)_{k \in \mathbb{N}}$, which is a graph with vertex set \mathbb{N} . If \mathcal{I}^* is a Polish space, then the empirical measures $\hat{\mu}^{(m)}$ satisfy

$$(2.4) \quad \begin{aligned} \hat{\mu}^{(m)} &= \frac{1}{m} \sum_{k \in [m]} \delta_{i_k} \xrightarrow{m \rightarrow \infty} \mu, \\ \frac{1}{m^2} \sum_{k, \ell \in [m]} \delta_{(i_k, i_\ell)} &\xrightarrow{m \rightarrow \infty} \mu \otimes \mu, \\ \frac{1}{m^L} \sum_{k_1, \dots, k_L} \delta_{(i_{k_1}, \dots, i_{k_L})} &\xrightarrow{m \rightarrow \infty} \mu^{\otimes L}, \\ \frac{1}{m} \sum_{i \in [m]} \delta_{(\underline{\hat{h}}_L)(\tau_i)} &\xrightarrow{m \rightarrow \infty} \nu^{(L)}, \quad L \in \mathbb{N}. \end{aligned}$$

Hence we obtain the *continuum grapheme* as a limit of the *sampled finite graphemes*.

Thus, starting from an evolution of finite graphs in \mathbb{G}_n for $n \in \mathbb{N}$, the *basic question* is what spaces \mathcal{I}^* we can take to form the elements $[\mathcal{I}_n^*, h_n, \mu_n]$ of the sequence and its limit

$$(2.5) \quad [\mathcal{I}^*, h, \mu],$$

which has the embedding property (denoted by \hookrightarrow):

$$(2.6) \quad \begin{aligned} (i) \quad & [\mathcal{I}_n^*, h_n, \mu_n] \hookrightarrow [\mathcal{I}^*, h, \mu], \quad n \in \mathbb{N}, \\ & \mathcal{I}_n \hookrightarrow \mathcal{I}, h_n = h_{\mathcal{I}_n}, \mu_n \hookrightarrow \tilde{\mu}_n, \\ & \tilde{\mu}_n \longrightarrow \mu, h_n \rightarrow h, h \text{ measurable and } \{0, 1\}\text{-valued,} \\ (ii) \quad & (\mathcal{I}_n^*, h_n, \mu_n) \xrightarrow[n \rightarrow \infty]{} (\mathcal{I}^*, h, \mu). \end{aligned}$$

Here, (i) holds in the sense that $\nu_n \Rightarrow \nu$ or, equivalently,

$$(2.7) \quad \Phi([\mathcal{I}_n^*, h_n, \mu_n]) \xrightarrow[n \rightarrow \infty]{} \Phi([\mathcal{I}^*, h, \mu]), \quad \Phi \in \Pi^*.$$

For (ii) (i.e., convergence of the environment sequence fitting either \mathbb{G}^\square or $\mathbb{G}^{\{\cdot\}}$), see Section 3.1. Of course, more than one choice is possible for the embedding. However, the procedure is *consistent*.

Proposition 2.1 (Consistency). *Every marked grapheme is the limit of its finite samples ν -a.s.* \square

How to best choose the space in which we embed? From graphon theory we know that we can always choose $([0, 1], \mathcal{B}_{[0,1]}, \mu)$ and we get a grapheme if the entropy of the sampled finite graphs of size n grows slower than n^2 (see [GdHKW23] and [HN12]). But this need *not* be the best choice for an embedding of the sample sequence. The space has to be rich enough to support the partition into completely connected components induced by the connection function at any fixed time. But we want more, namely, a stochastic process with *regular paths*. This would allow us to get convergence in path space for the finite-graph dynamics, which is not yet reflected in (2.6), which we strengthen to

$$(2.8) \quad \mathcal{L} \left[([\mathcal{I}_{n,t}, h_{n,t}, \mu_{n,t}])_{t \geq 0} \right] \xrightarrow[u \rightarrow \infty]{} \mathcal{L} \left[([\mathcal{I}_t, h_t, \mu_t])_{t \geq 0} \right]$$

or the same with $\{\cdot\}$. At this point we can either choose a *universal* space, e.g. $([0, 1], \mathcal{B}_{[0,1]})$, with a fitting μ , or *random time-dependent* triples $(\mathcal{I}_t^*, h_t, \mu_t)$ adapted to the dynamics, and try to prove convergence of the finite-graph dynamics by choosing for $\mathcal{G}^{(n)}$ any of the finite versions of $(\mathcal{I}_t^*, h_t, \mu_t)$. As pointed out in Sections 1.3.4 and 1.4, it is indeed useful to work with the latter choice and pick the *space of genealogies* associated with the graph dynamics, since this is an *effective coding* of the *space-time process*, which we want to understand better through our constructions.

Outlook on questions and further research 1. We have already addressed the following questions for graphemes. When is there a *measurable* (or possibly even *continuous*) h for $[\mathcal{I}^*, h, \mu]$, given a sequence $(i^{(m)})_{m \in \mathbb{N}}$ sampled with μ , such that $(h(i_k, i_\ell))_{k, \ell \in \mathbb{N}}$ yield the connection matrix of the sequence? Do we have tightness in n ? Is there a minimal

or convenient embedding space? We have answered these questions when the states arise from the processes defined in Section 1.3, viewed as processes taking values in \mathbb{G}^{\square} or $\mathbb{G}^{\square, V}$. The issue is whether we can view the evolving grapheme as a *process in random environment* with a well-posed evolution of ν in \mathbb{G}^{\diamond} , and with the environment having its own evolution, either *autonomous* or arising from a *well-posed* Markov process. These properties hold for the examples of Markov processes of graphs given in Section 1.3, since for the underlying dynamics we know from the literature that $[\mathcal{I}_n, r_n, \mu_n]$ converges to $[\mathcal{I}, r, \mu]$. The dynamics of the vertices is exactly as in standard models, and we showed that $[(\mathcal{I}_n, r_n), h_n, \mu_n]$ converges to a limit $[(\mathcal{I}, r), h, \mu]$ and $n \rightarrow \infty$.

What we have to look for are further types of dynamics that can be tackled in the same way. There are two interesting directions:

(1) Models where a *fraction of edges* is removed and reconnected at any given time, giving jump processes. Here our approach can be based on so-called Λ -Cannings processes and their genealogies (see [GKW23]) and leads to the same theorems, without an explicit identification of the equilibrium.

(2) Consider processes where events as we have treated them above do *not* come about in a Markovian way. This can again be tackled by using *population models*, but now with *seed-banks*, which allow us to remove the non-Markovianess by introducing countably many seed-banks (see [GdHO22, GdHO21, GdH23]). Here, the genealogy process has not yet been treated, but would allow for a similar analysis as given here.

Summary of the challenges. We face the following challenges when we start with a specific finite-graph dynamics:

- Tightness, non-degeneracy and existence of h . Are *further conditions* needed to guarantee the *existence of \mathcal{G}* with values in \mathbb{G}^{\square} , i.e., do (2.5)–(2.8) hold for a given dynamics?
- Is (\mathcal{I}^*, h, μ) *non-trivial*?
- We can always find an h on $\mathcal{I} \times \mathcal{I}$ that is measurable as a random variable. Even when we have a sampled sequence $(i_n)_{n \in \mathbb{N}}$ and a connection matrix $(h(i_k, i_\ell))_{k, \ell \in \mathbb{N}}$, a key question is in what cases we can, instead of a random variable, obtain a continuous function on $\mathcal{I} \times \mathcal{I}$ via a continuum decomposition on $\text{supp}(\mu^{\otimes 2})$, so that we have an evolving partition of a continuum for the completely connected components. In our examples we obtain an h that allows for a *continuous* continuation to $\mathcal{I} \times \mathcal{I}$ or, even better, to $(\text{supp}(\mu))^2$. The question is how this construction can be generalised. We saw a possibility in our last theorem on the model with insertion and removal.

We have already resolved these challenges for our earlier classes of dynamics, but we must see how this can be generalised as a recipe, in particular, for the model where edges switch between active and non-active (Fleming-Viot with insertion or removal). In the other examples, we can identify h and μ explicitly, because we have a *genealogy state \mathcal{U}* generating the connection matrix that evolves according to a \mathbb{U} -valued stochastic process, which itself evolves autonomously. The latter yields our \mathbb{G}^{\square} -valued process, which has a natural topology on \mathcal{I} , as well as the property of being *dust-free*.

In Section 1 we have seen, for a collection of finite-graph dynamics, how these challenges can be addressed. We continue by explaining why our approach has a more general nature, even though we cannot yet work this out fully in the present paper.

2.2 Choice of state space, challenges and perspectives: intrinsic role of genealogies

Often we are able to generate $\mathcal{I}^* = (\mathcal{I}, \mathcal{B})$ by considering a convenient Polish space with a Borel- σ -algebra like $[0, 1]$ with the Euclidean distance. We always have a metric r in any Polish space, but the problem is that it might not be easy to obtain a proper dynamics for this metric, in such a way that we get a choice for all $t \geq 0$ and for every realisation of the randomness. Since the measure changes as we follow the evolution, it is appropriate to choose also $(\mathcal{I}_t, \mathcal{B}_t)_{t \geq 0}$ *randomly*, at the cost that we do *not* get a *universal* space with *only* μ depending on time and on the underlying randomness.

For the classes of examples of finite-graph dynamics and grapheme dynamics we look at here, there is a *canonical choice* based on the underlying *ultrametric measure space* (\mathcal{U}_t, r_t) of the \mathbb{U} -valued Fleming-Viot or Dawson-Watanabe process generated by the stochastically evolving *genealogy* \mathcal{U}_t of the population at time t of the underlying population model. We saw that the ultrametricity of the spaces in which we embed is crucial to get a continuous representative of h . Does this mean that the choice had in fact an *intrinsic character*? The answer is *yes*, the reason being that the *genealogy* of the population stores all the information about the *space-time path process* up to time t of the \mathbb{G}^\diamond -valued version of our process in a *natural and minimal way* in the state of the \mathbb{G}^\square -valued process, as we next explain.

If $(X_t)_{t \geq 0}$ is a Markov process, then we can consider the *space-time path process* $(\mathcal{X}_t)_{t \geq 0}$ with

$$(2.9) \quad \mathcal{X}_t = (s, X_s)_{s \in [0, t]}.$$

In our context and for finite graphs, if at time t we ask at what earlier time a vertex got its connection to its current connected component from (by looking backwards in time), then we find an ancestral path and an imbedded search tree. This defines an ultrametric space (\mathcal{I}_t, r_t) that has this search tree as a skeleton. This tree codes the pivotal event that two vertices are in the same connected component. In particular, this space is rich enough to encode the information on the edges and their *history*, and therefore so is the canonical space (\mathcal{I}, r) for the state at time t for every dynamics that arises from a finite-graph dynamics where edges are added or removed by connecting or disconnecting vertices from their connected components. Here, we simply use the *space-time process* to see the *intrinsic tree structure*. The observation above also indicates that the path in the state space \mathbb{G}^\square is the minimal one that unravels the information on the *space-time path process* in \mathbb{G}^\diamond if we choose the space given by the underlying \mathbb{U} -valued genealogy process. These are also the right spaces to get limiting dynamics as the number of vertices diverges.

Outlook on questions and further research 2. In fact, ongoing work on population models shows that we can embed realisations of the grapheme as countable graphs at

different times in a single grounded \mathbb{R} -tree equipped with a collection of sampling measures. In this way we get a representation of the *whole space-time path process* of the *countable graphs*, in a single tree-like metric measure space for all times simultaneously.

3 Grapheme evolution 1: State space, topology and test functions

In this section we build up the topology to get a proper *state space* for our grapheme dynamics and state basic properties of the *topology*. This is done by specifying *algebras of test functions* on \mathbb{G}^\sim , with $\sim = \langle \rangle, \{\}, \square$, that we use to define *convergence of sequences* in these spaces and thus a topology. Later, in Section 4.2, these functions will also be the basis for the *domain of the operator* of our martingale problems. In Section 3.1 we focus on the topology, on the test functions, and on tightness issues. In Section 3.2 we extend the setting to varying size, i.e., $\bar{\mathbb{G}}^\sim$. In Section 3.3 we extend further to V -marked versions $\mathbb{G}^{\sim,V}, \bar{\mathbb{G}}^{\sim,V}$.

3.1 Test functions, topology and tightness for grapheme spaces

3.1.1 Test functions on grapheme spaces and topology

We introduce a *topology* on the sets $\mathbb{G}^\square, \mathbb{G}^{\{\}}, \mathbb{G}^{\langle \rangle}$ defined in Section 1.2, in order to specify Borel σ -algebras $\mathcal{B}_{\mathbb{G}^\square}, \mathcal{B}_{\mathbb{G}^{\{\}}}, \mathcal{B}_{\mathbb{G}^{\langle \rangle}}$ and, if possible, turn them into a Polish space. We do so by defining *convergence of sequences* of elements of the space $\mathbb{G}^\square, \mathbb{G}^{\{\}}, \mathbb{G}^{\langle \rangle}$. This in turn is based on convergence of the sequences of *evaluations* of a specific class of test functions on the spaces $(\mathcal{I}, \mathcal{B}_{\mathcal{I}})$ for $\mathbb{G}^{\{\}}, (\mathcal{I}, r)$ for \mathbb{G}^\square , respectively, $\{0, 1\}^{\mathbb{N} \times \mathbb{N}}$ for $\mathbb{G}^{\langle \rangle}$. The proof of the propositions stated in the present section are given in Section 5.

Functions. Our functions are based on evaluations for specific functions of points in sampled *finite* subspaces from \mathcal{I} such that they only depend on the equivalence class of $(\mathcal{I}, \mathcal{B}_{\mathcal{I}})$ and on the connection-matrix distribution via h .

- The algebra of functions, called *polynomials*, arise as functions on *sampled finite subspaces*, with their *edges* obtained by sampling finitely many vertices with μ from \mathcal{I} (without replacement if μ has atoms), forming the *finite version of the space \mathcal{I}^** with the restriction of h defining a *finite subgraph* that is *embedded* in \mathcal{I} . If we put the equidistribution on the vertices, then we get a *sampled finite grapheme* (see (3.7), (3.10) and (3.12) below). On that grapheme we define functions on a representative of the edges and of the vertices as points in the underlying space \mathcal{I} , which are taken from the finite space built as just indicated. Below we deal separately with two cases: (i) corresponding to $\mathbb{G}^{\langle \rangle}$, respectively, (ii) corresponding to \mathbb{G}^\square or $\mathbb{G}^{\{\}}$.

(i) Functions on $\mathbb{G}^{\langle \rangle}$. Consider the functional ν of the connection-matrix distribution in an \mathbb{N} -sample, i.e., look at the space $\mathbb{G}^{\langle \rangle}$ and use test functions on $\mathbb{G}^{\langle \rangle}$ that can be lifted to

functions on $\mathbb{G}^\square, \mathbb{G}^\{\}$. Recall that an element of \mathbb{G}^\diamond is of the form $\langle (\mathcal{I}, \tau), h, \mu \rangle$ for *some* Polish space $\langle \mathcal{I}, \tau \rangle$. Our first observation is that we can study the marginal distributions $\nu^{(m)}$, $m \in \mathbb{N}$, of ν via the *subgraph count functions*, which are the *monomials* Φ of degree m , given as follows.

Fix $C = \{1, \dots, m\}$ and $A, B \subseteq C^2 \setminus \text{Diag}(C^2)$ (denoted by $C^{2\setminus}$) such that $A \cup B = C^{2\setminus}$, $A \cap B = \emptyset$ and $|A| + |B| = \binom{m}{2}$ for some $m = |C| \in \mathbb{N}$ (this induces an m -subgraph with specified vertices from C and its edges). Introduce functions $\tilde{\varphi}_{A,B} = 1_A(1 - 1_B)$ that are indicators of the set of m -graphs with the pairs in A connected and the pairs in B not connected. In this way we get functions on the pair (\mathcal{I}^m, h) :

$$(3.1) \quad \mathcal{F}_m = \left\{ \varphi_{A,B}: A, B \subseteq C^{2\setminus}, A \cap B = \emptyset, |A| + |B| = \binom{m}{2} \right\},$$

where $\varphi_{A,B}$ on (\mathcal{I}^m, h) is defined via $\tilde{\varphi}_{A,B}$ as

$$(3.2) \quad \varphi_{A,B}(\underline{u}) = \tilde{\varphi}_{A,B}(\underline{h}(\underline{u})) = \prod_{(i,j) \in A} h(u_i, u_j) \prod_{(i,j) \in B} (1 - h(u_i, u_j)).$$

Define

$$(3.3) \quad \mathcal{F} = \bigcup_{m \in \mathbb{N}} \mathcal{F}_m,$$

which distinguishes equivalence classes of countable graphs, i.e., graphemes. Set

$$(3.4) \quad \Phi^{m, \varphi_{A,B}}(\mathcal{I}, h, \mu) = \int_{\mathcal{I}^m} \varphi_{A,B}(\underline{h}(\underline{u})) \mu^{\otimes m}(d\underline{u}),$$

where for μ with atoms we use $\mu^{\otimes m, \downarrow}$. Then the test functions on \mathbb{G}^\diamond are given by

$$(3.5) \quad \Pi^{*, \mathcal{F}} = \text{the algebra generated by (= connection polynomials)}.$$

These polynomials determine asymptotically the behaviour of the *normalised subgraphs counts* in sequences sampled according to μ .

For $m \in \mathbb{N}$, consider more generally functions in $\Pi^{*, \mathcal{F}}$ that are specified for some symmetric φ vanishing on the diagonal,

$$(3.6) \quad \varphi: \{0, 1\}^{[m] \times [m]} \rightarrow \mathbb{R}.$$

Call this set $\tilde{\mathcal{F}}_m$, with \underline{h} denoting the connection matrix of an m -sample \underline{u} from \mathcal{I} . Then the test functions are given by

$$(3.7) \quad \Phi^{m, \varphi}([\mathcal{I}, h, \mu]) = \int_{\mathcal{I}^m} \varphi(\underline{h})(\underline{u}) \mu^{\otimes m}(d\underline{u}) = \int_{\{0,1\}^m} \varphi(\underline{h}) \nu^{(m)}(d\underline{h}), \quad \varphi \in \tilde{\mathcal{F}}_m, m \in \mathbb{N}.$$

Note that (3.7) is only defined when h is a *measurable function*. The right-hand side shows that our quantities so far do *not* depend on the choice of representatives of the equivalence class. We define the algebra generated by the monomials in (3.7) on \mathbb{G}^\diamond as

$$(3.8) \quad \Pi^* = \text{algebra of general connection polynomials}.$$

On $\mathbb{G}^{\langle \rangle}$ we want a topology that is based on these functions, requiring for *convergence* of a sequence $(\mathcal{G}_n)_{n \in \mathbb{N}}$ to \mathcal{G} the convergence of all evaluations of \mathcal{G}_n by $\Phi \in \Pi^*$ to the one of \mathcal{G} . This works because of the following properties.

Proposition 3.1 (Π^* law convergence determining for ν).

(a) The algebra generated by Π^* uniquely determines the $m \times m$ connection-matrix distributions $\nu^{(m)}$, $m \in \mathbb{N}_{\geq 2}$, and is by definition convergence determining on $\mathbb{G}^{\langle \rangle}$.

(b) The moments of the random grapheme \mathcal{G} with state space $\mathbb{G}^{\langle \rangle}$,

$$(3.9) \quad \{\mathbb{E}[\Phi(\mathcal{G})], \Phi \in \Pi^*\},$$

determine $\mathcal{L}[\nu]$ uniquely and are convergence determining for laws on $\mathbb{G}^{\langle \rangle}$. \square

We can extend these functions $\mathbb{G}^{\langle \rangle}$ from Π^* to \mathbb{G}^{\square} and $\mathbb{G}^{\{\}}$, since they do not depend on anything other than the functional ν . We misuse notation by still talking about the algebra Π^* , even though strictly speaking we have $\Pi^* \sim$ with $\sim = \langle \rangle, \square$ or $\{\}$. Proposition 3.1 holds for the functional ν on the spaces $\mathbb{G}^{\square}, \mathbb{G}^{\{\}}$ as well.

(ii) **Functions on $\mathbb{G}^{\square}, \mathbb{G}^{\{\}}$.** In these space we have to look also at *convergence of the underlying spaces*, i.e., we consider functions for (\mathcal{I}^*, μ) , respectively, (\mathcal{I}, r, μ) in order to be able to treat $\mathbb{G}^{\square}, \mathbb{G}^{\{\}}$. Note that the equivalence relation leading to \mathbb{G}^{\square} requires isometric measure-preserving bijections between the spaces $(\mathcal{I}', \mathcal{B}_{\mathcal{I}'})$, $(\mathcal{I}'', \mathcal{B}_{\mathcal{I}''})$, under the assumption that $\mathcal{I}' = \text{supp}(\mu')$, $\mathcal{I}'' = \text{supp}(\mu'')$. This *equivalence* is denoted by \square . The topology on \mathbb{G}^{\square} is therefore obtained by defining *convergence of sequences* $\mathcal{G}_n \rightarrow \mathcal{G}$, with $\mathcal{G}_n, \mathcal{G} \in \mathbb{G}^{\square}$ as below in the paragraph on topology. We proceed similarly in $\mathbb{G}^{\{\}}$, replacing isometric by homeomorphic. This means that we need a set of test functions of $[\mathcal{I}, r, \mu]$, respectively, $[\mathcal{I}^*, \mu]$ that are *separating*.

Take \mathbb{G}^{\square} and consider test functions that are averages with respect to the sampling measure μ of functions of matrices of m -samples, namely, the *monomials* $\Phi^{m, \varphi}$ on \mathbb{U}_1 , i.e., functions of $[\mathcal{I}^*, h, \mu]$. More precisely, on \mathbb{G}^{\square} we choose a bounded continuous test function $\varphi(\underline{r})$, with \underline{r} a continuous \mathbb{R} -valued function on $\mathcal{I} \times \mathcal{I}$, given by *distances* $r(\cdot, \cdot)$ between points in \mathcal{I} , which determines the equivalence class of (U, r, μ) in \mathbb{U}_1 and is of the form (see [GPW09]) generating the algebra that we call $\widehat{\Pi}$:

$$(3.10) \quad \Phi^{m, \varphi}([\mathcal{I}, r), h, \mu]) = \int_{\mathcal{I}^m} \varphi(\underline{r}) d\mu^{\otimes m}, \quad m \in \mathbb{N}, \varphi \in C_b(\mathbb{R}^2).$$

In order to obtain the test functions for \mathbb{G}^{\square} , we take the function from Π^* that determines the distribution of h under μ sampling, consider $\varphi(\underline{h})$ in (3.10) instead of $\varphi(\underline{r})$, take the integral of the product $\varphi_r(\underline{r})\varphi_h(\underline{h})$ to get a monomial on \mathbb{G}^{\square} , and consider the generated algebra $\widehat{\Pi}^*$.

For \mathcal{I}^* and $\mathbb{G}^{\{\}}$,

$$(3.11) \quad \text{in (3.10) use as integral a continuous function } \varphi(\underline{u}) \text{ on } \mathcal{I}^m \text{ (instead of } \varphi(\underline{r})),$$

induced by products of functions, say $(\varphi_i)_{i \in [m]}$, that are bounded and continuous on \mathcal{I} . Take the product over $i \in [m]$. This identifies $((\mathcal{I}, \tau), \mu)$ and generates an algebra $\widehat{\Pi}$ on the

equivalence classes of Polish measure spaces. To identify the grapheme including h , this has to be *complemented by a further factor of the form in (3.7)*. We note that, in order to evaluate the function in another representative, the formula in the analogue of (3.10) involves the homeomorphism.

The above construction on $\mathbb{G}^{\square}, \mathbb{G}^{\{ \}$ (we use the same notation for the algebras whenever no confusion can occur) leads to

$$(3.12) \quad \widehat{\Pi}^* = \text{algebra of polynomials of } [(\mathcal{I}, r), h, \mu], \text{ respectively, } \{\mathcal{I}^*, h, \mu\}.$$

Proposition 3.2 (Law determining). *The algebra $\widehat{\Pi}^*$ is separating for \mathbb{G}^{\square} and $\mathbb{G}^{\{ \}$, and, when taking random elements as input expectations, determine the laws on \mathbb{G}^{\square} , respectively, $\mathbb{G}^{\{ \}$. \square*

Topology. We next define *convergence* of sequences to a limit point, which has two ingredients.

(I) The *first* requirement for a sequence of graphemes from $\widetilde{\mathbb{G}}$ with $\sim = \langle \rangle, \square, \{ \}$ to converge to a limit grapheme is that the corresponding *connection-matrix distributions converge*. Formally, the sequence $([\mathcal{I}_n^*, h_n, \mu_n])_{n \in \mathbb{N}}$ from $\mathbb{G}^{\langle \rangle}$ converges to an element $[\mathcal{I}^*, h, \mu] \in \mathbb{G}^{\langle \rangle}$ if and only if

$$(3.13) \quad \Phi([\mathcal{I}_n^*, h_n, \mu_n]) \xrightarrow[n \rightarrow \infty]{} \Phi([\mathcal{I}^*, h, \mu]) \quad \forall \Phi \in \Pi^{*, \mathcal{F}},$$

and hence for all $\Phi \in \Pi^*$.

(II) For $\mathbb{G}^{\langle \rangle}$, the *second* requirement is strengthening the graphon topology as follows. The set of $\{0, 1\}^{\mathbb{N}}$ -valued connection-matrix distributions is compact in the weak topology. But what we need in $\mathbb{G}^{\langle \rangle}$ are connection-matrix distributions of an infinite sample sequence guaranteeing that $(\nu_n)_{n \in \mathbb{N}}$ has limit points ν that arise from a sampling sequence evaluating a $\{0, 1\}$ -valued function on $\mathcal{I}^* \times \mathcal{I}^*$. For $\mathbb{G}^{\{ \}$ and \mathbb{G}^{\square} , the *second* requirement is that the limit points ν from (I) have a suitable (\mathcal{I}^*, μ) in which they can be embedded, and that these triples arise as the *limit* along the sequence as well. In other words, we have to specify a topology for the space of equivalence classes of (\mathcal{I}^*, h, μ) from either $\mathbb{G}^{\{ \}$ or \mathbb{G}^{\square} . For that purpose we set

$$(3.14) \quad \mathbb{T} \text{ is the space of } \textit{equivalence classes } [\mathcal{I}^*, \mu] \text{ of Polish measure spaces,}$$

arising as limit of elements with finitely many points arising from elements of \mathbb{G}_{∞} , and

$$(3.15) \quad \mathbb{M} \text{ is the space of } \textit{equivalence classes } [\mathcal{I}, r, \mu],$$

arising of limits of *finite* metric measure spaces. For convergence of a sequence $[\mathcal{I}^*, h, \mu]$, respectively, $\{\mathcal{I}^*, h, \mu\}$ we also have to require, beyond (3.13), that

$$(3.16) \quad ((\mathcal{I}_n^*, \mu_n))_{n \in \mathbb{N}} \text{ converges to } (\mathcal{I}^*, \mu) \text{ in the topology of } \mathbb{T}, \text{ respectively, } \mathbb{M} \text{ with } (\mathcal{I}, r, \mu).$$

The topologies for \mathbb{M} , respectively, \mathbb{T} are defined by specifying *convergence of sequences*, which is obtained by requiring convergence of evaluations of the chosen class of test functions $\widehat{\Pi}$ given in (3.10), respectively, (3.11) for \mathbb{G}^{\square} , respectively, $\mathbb{G}^{\{ \}$.

To get convergence of the joint structure of h and (\mathcal{I}^*, μ) , we have to *enrich* the algebra of polynomials to $\widehat{\Pi}^*$, so that they separate equivalence classes of Polish measure spaces, and subsequently define the topology based on the evaluations from $\widehat{\Pi}^*$. Hence we define *convergence in $\mathbb{G}^\square, \mathbb{G}^\{\}$* by requiring

$$(3.17) \quad \text{convergence of the evaluations of all functions from (3.12).}$$

Thus, in (I) and (II) we have turned $\mathbb{G}^\square, \mathbb{G}^\{\}, \mathbb{G}^\diamond$ into *topological* spaces and we have Borel- σ -algebras to obtain measure spaces. The important point is to get, whenever possible, *Polish* spaces, as state spaces for random variables and stochastic processes in their standard form.

Proposition 3.3 (State spaces). $\mathbb{G}_{\text{ultra}}^\square$ is a Polish space, and so is the closure of \mathbb{G}^\diamond , called $\widehat{\mathcal{W}}$, and of $\mathbb{G}^\square, \mathbb{G}^\{\}$, called $\widehat{\mathcal{W}}^\square, \widehat{\mathcal{W}}^\{\}$. \square

Remark 3.4. We leave open the issue that if we introduce on $\mathbb{G}^*, * = \diamond, \{\}, \square$ a stronger topology for convergence of the h -component in the grapheme topology, guaranteeing that the functional ν satisfies an entropy condition on its range that is equivalent to having $\{0, 1\}$ -valued graphons, then the closure in this stronger topology gives a Polish space \mathbb{G}^* . \spadesuit

Since the closure of \mathbb{G}^\diamond corresponds to the space of graphons, the set \mathbb{G}^\diamond is strictly smaller than the closure and we have to therefore check compact containment conditions for our processes in order to obtain processes with values only in \mathbb{G}^\diamond . Since in our theorems we obtained processes with values in \mathbb{G}^\square , this guarantees the required property after we pass to \mathbb{G}^\diamond and we know that we truly obtain a process with values in \mathbb{G}^\diamond .

Does the same hold for $\mathbb{G}^\{\}$, or must we restrict to smaller spaces, like spaces of algebraic trees (see [LW21])? What remains is to settle metrisability (separability and completeness follow from the definition), i.e., to *exhibit a metric* generating the topology. There is no obvious reference for this. However, we observe that if we restrict \mathbb{T} to spaces that have a representative that is an *algebraic tree*, then this closed subspace can be equipped with a concrete metric via an embedding, and for our processes we would get a dynamics in this Polish space. We will not pursue this task further in the present paper.

We had been able to do the construction for $\mathbb{G}^\diamond, \mathbb{G}^\square$, since the topology can be based on a countable set of test functions that are *universal* for all representations of an equivalence class, which is not available for $\mathbb{G}^\{\}$. But this comes at the price that we have to check on \mathbb{G}^\diamond whether our stochastic processes have paths with values in \mathbb{G}^\diamond , which is not complete in the metric based on the subgraph counts.

Remark 3.5 (\mathbb{G}^\square versus $\mathbb{G}^\{\}$ or \mathbb{G}^\diamond).

(1) These spaces are problematic for the h -component, namely, the lack of completeness of \mathbb{G}^\diamond , respectively, the fact that its completion does not represent countable graphs (they may be randomised), so that we cannot formulate standard martingale problems on the full space, except on the $\widetilde{\mathbb{G}}_{\text{ultra}}^*$ -versions, as we have demonstrated.

(2) There is a clear advantage to work with (\mathcal{I}, r) and on \mathbb{G}^\square . On $\mathbb{G}^\{\}$ we have the following further drawback. On \mathbb{G}^\square the test functions are natural and are comparable between different representatives, since we deal with φ that are *universal* in the representatives.

This is not the case on $\mathbb{G}^{\{\}}$ with $\varphi(\underline{u})$, where we have to use homeomorphisms between different representatives that are used to define φ on a different representative than a tagged representative from the equivalence class. This problem can only be circumvented on $\mathbb{G}^{\{\}}$, when we embed the space of a representative of the equivalence class homomorphically into $([0, 1], \mathcal{B}_{[0,1]}, du)$ and work with *uniformly* continuous functions on $[0, 1]$. However, this yields bounded continuous functions on \mathcal{I}^* , but as a function of \mathcal{I} we lose uniformity, because the embedding in $[0, 1]$ does *not* necessarily have this uniformity. This in turn means that we can define continuous bounded polynomials as test functions to define our topology, but we are lacking concrete *uniformly continuous subclasses* of such functions that are also measure and convergence determining. This in turn means that we cannot exhibit on $\mathbb{G}^{\{\}}$ a countable measure-determining set of uniformly continuous test functions in $C_b(\mathbb{G}^{\{\}}, \mathbb{R})$, induced by a function on \mathcal{I}^* that is invariant under equivalence, as we can on \mathbb{G}^{\square} . This fact is an obstacle, for example, when we want to prove the strong *Markov property* or a (generalised) *Feller property* for stochastic processes with values in these spaces. ♠

The polynomials on $\mathbb{G}^{\{\}}$ are a *subclass* of the polynomials on \mathbb{G}^{\square} , since fewer Polish spaces are equivalent. In particular, the martingale problems on $\mathbb{G}^{\{\}}$ are more complicated than those on \mathbb{G}^{\square} , since we deal with invariants for larger sets.

3.1.2 Tightness in grapheme spaces

We often have to decide whether sequences of graphemes converge to a *limit element* in the respective spaces \mathbb{G}^{\sim} with $\sim = \square, \{\}, \langle \rangle$ when the evaluations of polynomials in $\widehat{\Pi}^*$ converge to a limit number. In order to get laws and convergence of graphemes with different non-trivial limits of h (i.e., not equal to $\underline{0}$ or $\underline{1}$), we need *extra conditions*.

For convergence we must have that $((\mathcal{I}_n^*, h_n, \mu_n))_{n \in \mathbb{N}}$ is determined by the polynomials as test function and converges to a limit in \mathbb{G}^{\sim} with $\sim = \square, \{\}, \langle \rangle$. Since we have defined the convergence of elements in \mathbb{G}^{\square} as convergence of evaluations of polynomials, i.e., elements of the algebra $\widehat{\Pi}^*$ giving us the natural *topology* on \mathbb{G}^{\square} , it remains to prove that the *limit of the evaluations* is an *evaluation of a limit point* in the state space.

Recall that we require \mathcal{I}^* to be Polish only, so \mathcal{I}^* need not be compact or locally compact, nor does $\mathcal{M}_1(\mathcal{I}^*)$ need to be compact. Therefore we possibly *cannot* sample from a limit object. This poses conditions on $((\mathcal{I}_n^*, h_n, \mu_n))_{n \in \mathbb{N}}$, namely, *tightness properties* of the spaces in which we embed the graphemes of index n . It is here that the *embedding* enters into the game. Furthermore, we have to guarantee that limiting elements equal a $\{0, 1\}$ -valued h , which requires that the connection-matrix distributions viewed as measures on $[0, 1]^{\mathbb{N}}$ are tight in the stronger topology, strengthened by an *additional requirement* guaranteeing $\{0, 1\}$ -valued rather than $[0, 1]$ -valued limits. Such a tightness criterion is the entropy criterion for sequences of finite samples, requiring the entropy of these samples to be $o(n^2)$ (see [GdHKW23] and [HN12]), which restricts the closure of the finite graphemes to a subspace of $\widehat{\mathcal{W}}$. We will see in Section 7.1 that on $\mathbb{G}_{\text{ultra}}^{\square}$ this is immediate, since in that case the h_n are continuous and the metric is an ultrametric, so we need not involve the entropy criterion.

If we consider \mathbb{G}^{\square} , then we have metric measure spaces (\mathcal{I}, r) . We also know that here

we deal with the closed subspaces \mathbb{U}_1, \mathbb{U} , which evolve in $\mathbb{G}_{\text{ultra}}^{\square}$. We can use the theory in [GPW09, GPW13, DGP11] to get Polish spaces, convergence criteria and tightness criteria, the latter amounting to the so-called

$$(3.18) \quad \text{no dust condition.}$$

On \mathbb{U}_1, \mathbb{U} this amounts to requiring that, for every $\varepsilon > 0$, the sampling measure puts weight $\geq 1 - \varepsilon$ on *finitely many ancestors* a time ε back. Indeed, this gives us a condition for the tightness of the spaces in which we embed the elements of our sequence. Furthermore, we need convergence of the connection-matrix distributions. However, if the evaluations using $\Pi^* \subseteq \widehat{\Pi}^*$ converge and the spaces of embedding have a limit in which we can embed, then we have found the limit in \mathbb{G}^{\square} . Therefore with (3.18) we get a tightness condition in \mathbb{G}^{\square} .

Next, we consider random graphemes and their laws. We can derive a tightness criterion because these live in $\mathcal{M}_1(\mathbb{G}^{\square})$, which takes on the form \mathbb{U}_1, \mathbb{U} , namely,

$$(3.19) \quad \begin{array}{l} \text{the number of ancestors time } \varepsilon \text{ back that have descendants} \\ \text{of total mass } \geq 1 - \varepsilon \text{ is stochastically bounded.} \end{array}$$

On \mathbb{U} we need, in addition, that the total masses are stochastically bounded. Finally, we need that the laws of the connection-matrix distributions are concentrated on the set of functions with values in $\{0, 1\}$, for example, when we form local averages over h .

Replacing \mathbb{G}^{\square} by $\mathbb{G}^{\{\}}$, we get into different territory and there is not much available in the literature. By the embedding in $[0, 1]$ mentioned above, we can introduce the topology in the same way and obtain a topological space (in fact, a Polish space), but the algebra of polynomials is lacking nice properties. For \mathbb{G}^{\square} we can work with *distance-matrix distributions* to represent the state (analogously to the connection-matrix distributions $\nu^{(m)}$, $m \in \mathbb{N}$, used on \mathbb{G}^{\diamond}). On $\mathbb{G}^{\{\}}$ we have *sampled tuple distributions*, which via the embedding become sampled tuple distributions in $([0, 1], \mathcal{B}_{[0,1]})$. A formulation independent of the equivalence class is subtle, because we can only say that the polynomial transforms via the integral transformation formula on the equivalence class, hence is a function of that class, and so we can use the representation on a subset of $([0, 1], \mathcal{B}_{[0,1]})$. This means that we have test functions involving the space and the homeomorphisms, a much less canonical setting than functions of distances φ that are the same on every representative, as for \mathbb{G}^{\square} . As pointed out before, we cannot gain much by pursuing this line further.

Turning to \mathbb{G}^{\diamond} , we note that the elements of \mathbb{G}^{\diamond} correspond to a subset of probability measures on $\{0, 1\}^{\binom{\mathbb{N}}{2}}$ that is not closed. Hence, even though $\mathcal{M}(\{0, 1\}^{\binom{\mathbb{N}}{2}})$ is a Polish space, and as a measure space on a compact set is classical, we require restrictions that lead to a *subset*. Namely, the elements of this subset arise from i.i.d. sampling on *some Polish space*, and the *empirical function* h is $\{0, 1\}$ -valued. In other words, we need to check that limit points of sequences obtained from the connection-matrix distributions can indeed be realised, i.e., arise from an embedded sequence, and lead to a $\{0, 1\}$ -valued function h . This need not be easy, as is exemplified by graphons with grey-shades, corresponding to sequences $(h_n)_{n \in \mathbb{N}}$ that do not have a limit h with values in $\{0, 1\}$, and hence do not define a grapheme based on the space $([0, 1], \text{Borel}, \mu)$, because the limiting subgraph counts determine a unique h with values in $[0, 1]$. Therefore \mathbb{G}^{\diamond} is a subtle subset

that is homeomorphic to a subset of the Polish space $\mathcal{M}_1(\{0, 1\}^{\mathbb{N}})$. However, we can use the well-known fact that the space of graphons with the topology induced by the subgraph counts can be generated by the cut metric. Therefore we get with \mathbb{G}^{\diamond} a separable metrisable subset of a Polish space, which, however, is *not* closed. For our dynamics there are topologically closed subspaces $\mathbb{G}_{\text{ultra}}^{\square}$ that are closed under the dynamics and induce subspaces in \mathbb{G}^{\diamond} .

For a sequence of graphemes in \mathbb{G}^{\diamond} to converge, it *suffices* to find a sequence of elements from \mathbb{G}^{\square} , a limit candidate from that space, such that the corresponding *subgraph counts converge*. The limit candidate for the embedding we find by showing that the sequence has a limit in \mathbb{G}^{\square} , including a limiting function h . This means that from tightness criteria in \mathbb{T} we get *sufficient* criteria for tightness in \mathbb{G}^{\diamond} .

We can deal with \mathbb{G}^{\square} in the same way, which gives us accessible sufficient tightness criteria. With (3.19) we get a criterion guaranteeing that we have limit points in \mathbb{G}^{\square} , and convergence when the subgraph counts converge as well, since the sequence of equivalence classes of metric measure spaces has a limit grapheme in \mathbb{G}^{\square} and a limiting $\{0, 1\}$ -valued h . The latter is often immediate on $\mathbb{G}_{\text{ultra}}^{\square}$. However, we know from the general theory of graphons that the topology can be generated by a metric (the *cut distance* from (1.2)), so we get values in a subset of a Polish space.

3.2 Extension to varying grapheme sizes

We have to argue that the previous observations extend to varying grapheme sizes. To do so, we have to introduce polynomials Φ on $[U, r, \mu]$ of the form $\Phi = \bar{\Phi} \hat{\Phi}$ with $\bar{\Phi} \in C_b(\mathbb{R}^+, \mathbb{R})$ and $\hat{\Phi}$ a polynomial of $[U, r, \hat{\mu}]$, in order to be able to identify the pair $\bar{\mu}$ and $[U, r, \hat{\mu}]$. Hence for $[U, r, \mu]$ we work with functions $\bar{\Phi}$ of $\bar{\mu}$, that are *bounded* and continuous. In our models we will be able to work with polynomials based on μ rather than on $\hat{\mu}$, but we have to argue that we have the necessary *integrability properties* for $\bar{\mu}$, and must include growth properties in the degree in order to ensure that they are separating for a subset of \mathbb{U} that occurs as possible values of our random variables. In our examples based on the Galton-Watson process, we can work with $\bar{\Phi}(\bar{\mu}) = \bar{\mu}^n$, $n \in \mathbb{N}$, and consider the subset of μ for which $\bar{\mu}^n$ is integrable and the Laplace transform of $\bar{\mu}$ exists in some neighbourhood of 0. This is described in detail in [DG19] and [GRG21].

3.3 Extension to marked graphemes

For the case of V -marked graphemes, take a Polish space (V, r_V) and consider the *V -marked grapheme*

$$(3.20) \quad [(\mathcal{I} \times V)^*, h, \mu_{\mathcal{I} \times V}]$$

with $\tilde{\kappa}, \kappa$ denoting mark kernels, respectively, mark functions from \mathcal{I} to \mathcal{B}_V , respectively, V and

$$(3.21) \quad \mu_{\mathcal{I} \times V} = \mu \otimes \tilde{\kappa}, \quad \mu \in \mathcal{M}_1(\mathcal{I}^*),$$

where $\tilde{\kappa}$ is a kernel assigning marks to vertices (typically, $\tilde{\kappa}(\cdot, 0) = \delta_{\kappa(\cdot)}(0)$ for a *mark function* κ that measurably maps \mathcal{I} into V), and consider the spaces

$$(3.22) \quad \begin{aligned} \tilde{\mathbb{G}}^V &= \text{space of } V\text{-marked graphemes,} \\ \mathbb{G}^V &\text{ with } h(i, i') = 1 \implies \kappa(i) = \kappa(i'). \end{aligned}$$

Recall that the corresponding state space

$$(3.23) \quad \mathbb{G}^{\langle \cdot \rangle}, \mathbb{G}^{V, \langle \cdot \rangle}$$

is a coarse-graining of $\mathbb{G}^{\square}, \mathbb{G}^{\square, V}$. In the marked case $\tilde{\mathbb{G}}^V$, ν must be replaced by the

$$(3.24) \quad \text{marked connection-matrix distribution } \nu_V,$$

where we keep track of the edges and of the type of the vertices. Here,

$$(3.25) \quad \begin{aligned} &\text{the entry 1 is replaced by } (1, (v_1, v_2)) \text{ and the entry 0 by } (0, (v_1, v_2)), \text{ with } v_1, v_2 \in V. \end{aligned}$$

Proposition 3.6 (Determining properties).

(a) Any element in $\mathbb{G}^{\square, V}$ or $\mathbb{G}^{\{\cdot\}, V}$ uniquely determines ν_V , and similarly after \mathbb{G} is replaced by $\tilde{\mathbb{G}}$.

(b) Consider the case of a non-atomic sampling measure μ . Then every ν_V , with the sub-matrix independence property for any two sub-matrices connected with two disjoint index sets, determines an element in $\mathbb{G}^{\langle \cdot \rangle, V}$. This element may, however, be trivial, i.e., equal the complete graph or the totally disconnected graph. \square

We continue with the state space

$$(3.26) \quad \begin{aligned} \mathbb{G}^{\square, V} &= [(\mathcal{I} \times V)^*, (h, \kappa), \mu_{\mathcal{I} \times V}], \\ \kappa: \mathcal{I} &\rightarrow V, \quad \mu_{\mathcal{I} \times V} \in \mathcal{M}_1((\mathcal{I} \times V)^*), \quad h: \mathcal{I}^2 \rightarrow \{0, 1\}, \\ \mu_{\mathcal{I} \times V}(\cdot, \cdot) &: \mu(\cdot) \otimes \delta_{\kappa(\cdot)}^{(0)}, \quad \mu \in \mathcal{M}_1(\mathcal{I}^*). \end{aligned}$$

The equivalence class is formed with respect to h -preserving, κ -preserving and measure-preserving bijections between the supports of the measures. The objects that characterise the marked connection-type structure for finite n are

$$(3.27) \quad k_n: \mathcal{I}^n \rightarrow \{0, 1\}^{n(n-1)} \times V^n, \quad k_n(\mathcal{G}) = ((h(\iota_i, \iota_j))_{i, j \in [n]}, (\kappa(\iota_i))_{i \in [n]}),$$

which gives us $(\nu_{*\kappa_N}^{\otimes n})_{n \in \mathbb{N}}$, the sequence of connection-type matrix distributions of order n . We proceed analogously when we work with (\mathcal{I}, r) and \mathbb{G}^{\square} .

We can generalise further by allowing, instead of mark functions, *mark kernels*

$$(3.28) \quad \kappa: (\mathcal{I}, \mathcal{B}_{\mathcal{I}}) \rightarrow (V, \mathcal{B}_V).$$

Then the space $\mathbb{G}^{\langle \cdot \rangle, V}$ is based on the *connection-type matrix distribution* ν_V as invariant. The state spaces depend on whether we choose \mathcal{I}^* or (\mathcal{I}, r) :

$$(3.29) \quad \mathbb{G}^{\square, V}, \quad \mathbb{G}^{\{\cdot\}, V}.$$

We introduce a topology induced by defining convergence of evaluations of functions on $\mathbb{G}^{\square, V}$ taken from an *algebra of polynomials*, denoted by $\widehat{\Pi}^{*, V}$.

Let us first look at the simpler case $\mathbb{G}^{\langle \cdot \rangle, V}$ again. We define $\Pi^{*, V}$ acting on $\mathbb{G}^{\langle \cdot \rangle, V}$. A *monomial* $\Phi \in \Pi^{*, V}$ is given by a function $\varphi \in C_b(\{0, 1\}^{n(n-1)}, \mathbb{R})$, $\Psi \in C_b(V^n, \mathbb{R})$, $n \in \mathbb{N}$, and

$$\begin{aligned}
 \Phi(\mathcal{G}) &= \Phi([\mathcal{I} \times V]^*, h, \kappa, \mu_{\mathcal{I} \times V}) \\
 (3.30) \quad &= \int_{(\mathcal{I} \times V)^n} \varphi((h(\iota_k, v_k)(\iota_\ell, v_\ell))_{k, \ell \in [n]}) \psi((\kappa(\iota_k))_{k \in [n]}) \mu_{\mathcal{I} \times V}^{\otimes n}(\underline{\iota}, \underline{v}) \\
 &= \int_{(\{0, 1\} \times V^2)^{n(n-1)}} (\varphi \otimes \psi) d\nu_V^{(n)}
 \end{aligned}$$

with $\underline{\iota} = (\iota_1, \dots, \iota_n)$, $\underline{v} = (v_1, \dots, v_n)$, and ν_V the connection-type matrix distribution on the edges and their end points.

In order to get to the topology on $\mathbb{G}^{\square, V}$, $\mathbb{G}^{\{\cdot\}, V}$, we need to control the embedding in the space \mathcal{I}^* , and also need finer test functions

$$(3.31) \quad \widehat{\Pi}^{*, V}$$

to fully characterise the elements of $\mathbb{G}^{\square, V}$. We consider (3.30) and replace the test function φ by *products* of the two test functions

$$(3.32) \quad \varphi_h((h((\iota_k, v_k), (i_\ell, v_\ell))_{k, \ell \in [n]})) \varphi_r((r(\iota_k, i_\ell))_{k, \ell \in [n]}),$$

where φ_h denotes the φ in (3.30), and $\varphi_r: [0, \infty)^{n^2} \rightarrow [0, \infty)$ is continuous.

If we turn to $\widetilde{\mathbb{G}}^{\{\cdot\}, V}$, then we replace φ_r by a continuous function on \mathcal{I}^n . We cannot rely on the results in [GPW09, GPW13, DGP11], because for $\mathbb{G}^{\{\cdot\}, V}$ we have no nice subset that is measure-determining and convergence-determining and uniformly continuous. For the dynamics in the present paper this is not really crucial, since we work on $\mathbb{G}^{\square, V}$ anyway.

Proposition 3.7 (The state space: V -marked). $\mathbb{G}_{\text{ultra}}^{\square, V}$ and the closure of $\mathbb{G}^{\langle \cdot \rangle, V}$, $\mathbb{G}^{\square, V}$ are Polish. \square

As before, the question is whether the same holds for $\mathbb{G}^{\{\cdot\}, V}$ and what is the situation for $\mathbb{G}^{\langle \cdot \rangle, V}$. Here we refer to the discussion in the non-marked case.

4 Grapheme evolution 2: Examples and martingale problems

In this section we give the details of the construction behind the Fleming-Viot and the Dawson-Watanabe grapheme diffusion on the state space $\mathbb{G}^{\square, V}$ by using *well-posed martingale problems*, focussing mainly on the specifications of the *operators* acting on the domains of test functions $\widetilde{\Pi}^*$, respectively, $\widehat{\Pi}^{*, V}$, which we specified in Section 3.

Section 4.1 specifies the *finite-grapheme* Markov pure-jump processes, their rates and transitions and their deterministic part (the growth of distances), and gives their *generators*. Section 4.2 defines the *operator* of the martingale problem for the grapheme

diffusions in $\mathbb{G}^{\square}, \mathbb{G}^{\square, V}$ in the case of the dynamics with completely connected components. Section 4.3 proceeds with the case of dynamics with *non*-completely connected components, which arises after we allow independent insertion and deletion of single edges at rates a^+, a^- , or its *generalised version* with \mathcal{U} -dependent a^+, a^- .

4.1 From finite-graph processes to grapheme processes: two classes of examples

4.1.1 Fleming-Viot and Dawson-Watanabe: finite grapheme processes

We use the rules presented in Section 1.3.1 for a jump process of finite graphs, specify all the parameters, extend the description systematically to the *marked* version, and give the graph the structure of a grapheme by using a specific embedding of the graph in a measure space. We do this by constructing an equivalence class of finite ultrametric measure spaces from our evolving graphs. We define the Markov jump processes with deterministic component on \mathbb{G}_{∞} and introduce a notation for the corresponding grapheme processes. Recall that, due to the increase of distances in time, the evolving environment also has a deterministic continuous component and not only jumps. The formulation with types and marked graphemes is needed only when we extend to rules with mutation and selection. Otherwise we can think of the unmarked case. Nonetheless, it is useful to have the marks available to see some parts of the construction more clearly.

Key objects. The finite V -type graph of size n is described as a grapheme, i.e., as the equivalence class

$$(4.1) \quad \mathcal{G} = [\mathcal{I}^* \times V, h, \kappa, \mu_{\mathcal{I} \times V}]$$

of a quadruple $(\mathcal{I}^*, h, \kappa, \mu)$, where \mathcal{I} is a set of vertices ($\mathcal{I}^* = (\mathcal{I}, \mathcal{B}_{\mathcal{I}})$ is a measurable space) and μ is a probability measure on $\mathcal{B}_{\mathcal{I}}$. The mark function $\kappa: \mathcal{I} \rightarrow V$ is the *type-function* that associates with each vertex a type drawn from the set V , where (V, \mathcal{B}_V) is a measurable space based on a Polish space V . (Recall that for elements in $\mathbb{U}_1^V, \mathbb{U}^V$ this can in general be a transition kernel from $(\mathcal{I}, \mathcal{B}_{\mathcal{I}})$ to (V, \mathcal{B}_V) .) The measure $\mu_{\mathcal{I} \times V}$ on $\mathcal{I} \times V$ is defined as $\mu \otimes \kappa$. In the case where we have a mark *function*, the connection-type matrix

$$(4.2) \quad h: \mathcal{I} \times \mathcal{I} \rightarrow \{0, 1\}, \text{ respectively, } (\mathcal{I} \times V)^2 \rightarrow (\{0, 1\} \times V)^2, h \text{ measurable}$$

defines the edges (a, a') , with $a, a' \in \mathcal{I}$, and the types $v, v' \in V$ with $v = \kappa(a), v' = \kappa(a')$, by requiring that

$$(4.3) \quad (a, a') \text{ is an edge if and only if } h(a, a') = 1, \text{ respectively, } h((a, v), (a', v')) = 1.$$

The measure μ is the *sampling measure*, and allows us to draw m -samples from the graph by independently drawing m vertices (a_1, \dots, a_m) according to μ without replacement. Two bijections between $\text{supp}(\mu)$, $(\mathcal{I}, r, h, \kappa, \mu)$ and $(\mathcal{I}', h', \kappa', \mu')$, are equivalent if there exists a bijection φ of $\text{supp}(\mu)$ to $\text{supp}(\mu')$ such that μ -a.s. $h' = h \circ \varphi^{-1}$, $\kappa' = \kappa \circ \varphi^{-1}$ and $\mu' = \mu_{*\varphi}$. The set of all such objects is called

$$(4.4) \quad \tilde{\mathbb{G}}^V, \quad \mathbb{G}^V = \left\{ \tilde{\mathcal{G}} \in \mathbb{G}^V : h(a, a') = 1 \implies \kappa(a) = \kappa(a') \right\}.$$

The embedding in a discrete metric space. We have defined two classes of Markov pure-jump processes of finite graphs. From our evolving *finite* graph at time t we can build a *finite ultrametric measure space* with an h -function, namely, $(\mathcal{I}_t, h_t, \mu_t)$. For example, interpret invasion as a birth event. Then we can define *ancestors* and *descendants* in the obvious way. For the set \mathcal{I} of current vertices we can define the *distance of* (i, i') as twice the time back to the last common ancestor of i and i' . In this way we obtain a *finite ultra-metric space*, giving us (\mathcal{I}_t, r_t) . Hence our *finite graph* gets embedded in a *metric space* and we have

(4.5) constructed $((\mathcal{I}_t, r_t), h_t, \mu_t)$ and defined a (finite) grapheme process.

If we add the types, then we similarly obtain the marked grapheme and its evolution.

Finite-grapheme dynamics. We now turn the graph dynamics into a grapheme jump processes with deterministic part taking values in $\mathbb{G}^{\square, V}$: (i) the *Fleming-Viot process* and the *Dawson-Watanabe process* with *immigration and emigration*; (ii) the same processes as in (i) but with *insertion or deletion* of edges.

As possible initial states we consider the space $\mathbb{G}^{\square, V}$, which requires that edges occur only between vertices marked with the same type. The processes are defined via the following three rules, each specifying at which *rate* which *jump* occurs.

► **Fleming-Viot evolution rule.**

- At rate $d\binom{n}{2}$ sample a pair (a, a') from $(\mathcal{I} \times \mathcal{I}) \setminus D_{\mathcal{I} \times \mathcal{I}}$ based on μ . If $\kappa(a) = \kappa(a')$, then nothing happens, while if $\kappa(a) \neq \kappa(a')$, then with probability $\frac{1}{2}$

$$(4.6) \quad \begin{aligned} & a' \text{ takes over the type of } a, \\ & a \text{ takes over the type of } a', \end{aligned}$$

the edges of the looser are removed, and the looser is connected to all the vertices the winner is connected to, μ is fixed and ν changes to $\nu - \delta_{(a, \kappa(a))} + \delta_{(a, \kappa(a'))}$, respectively, $\nu - \delta_{(a', \kappa(a'))} + \delta_{(a, \kappa(a))}$. Another way of saying this is that individuals $a', a \in \mathcal{I}$ die and are replaced by two children, of individual either a or a' , inheriting the type (in population dynamics this resampling is known as the Moran model). This choice also defines the evolution of $(\mathcal{I}_t, r_t)_{t \geq 0}$. Namely, distances of two vertices grow at rate 2, and at a birth event the distance of the newborn to the ancestor is 0 and is the same for all other vertices. We start from the space where all distances are 0, and all vertices are identified with one point with full mass μ . If we have types, then the n points get different types drawn from the set V according to some diffusive measure κ_0 on V .

- At rate cn a vertex is sampled and removed (together with all its connections), a new vertex a' enters with the type v' chosen according to the measure θ on V , μ is fixed and $\mu_{\mathcal{I} \times V}$ changes to $\mu_{\mathcal{I} \times V} - \delta_{(a, \kappa(a))} + \delta_{(a', v')}$ with probability $\theta(dv')$.

Extensions of the Fleming-Viot rule with types and emigration/immigration.

Consider marked graphemes $[\mathcal{I}, h, \mu, \kappa]$, where $\kappa: \mathcal{I} \rightarrow V$ is a mark function (i.e., a measurable function from \mathcal{I} to V), and consider additional rules for marked populations and their genealogical processes with values in \mathbb{U}_1^V . Of interest are cases where *mutation* of types and/or *selection* w.r.t. types is added, leading to new processes

$$(4.7) \quad \left(\mathcal{G}_t^{d,c,\theta,m,s}\right)_{t \geq 0}, \quad \left(\mathcal{G}_t^{b,c,\theta,m,s}\right)_{t \geq 0},$$

that allow for statements like in Theorems 1.9, 1.14, 1.18 and 1.26 for the Fleming-Viot, Dawson-Watanabe and McKean-Vlasov evolution rules, extended by additional *mutation of types*. The type- b component mutates to type b'

$$(4.8) \quad \text{at rate } m \text{ with transition probability } M(b, db'), b, b' \in K,$$

and edges are connected and disconnected accordingly. To keep the notation simple, we only discuss the case when $M(b, \cdot)$ is *diffusive*, so that always *new* types appear. Furthermore, we have weak *selection* of types with fitness function χ , i.e., at rate s/n , $s \in (0, \infty)$, a vertex pair x, y is picked, and with probability

$$(4.9) \quad \begin{array}{ll} \frac{\chi(x)}{\chi(x) + \chi(y)} : & x \rightarrow y \text{ replacement takes place,} \\ \frac{\chi(y)}{\chi(y) + \chi(x)} : & y \rightarrow x \text{ replacement takes place.} \end{array}$$

► **Dawson-Watanabe evolution rule.** In the Dawson-Watanabe case, consider supercritical or subcritical branching. If $s > 0$, then at rate s a new vertex is born and is connected with every vertex. If $s < 0$, then at rate $-s$ a vertex is removed (including all its edges). Furthermore, a vertex changes type at rate m and with transition probability kernel $M(\cdot, \cdot)$ as above, and all old edges are removed and edges are added between all vertices carrying the new type.

► **Marked Fleming-Viot evolution rule with immigration.**

- Every initial individual gets its own type, drawn from the set $[0, 1]$. All individuals of the same type are connected.
- Pairs are sampled from the population at rate d . If a pair is sampled and the two individuals are of different type, then a coin is thrown to decide which of the two types is adopted by the pair, all connections are updated, one individual is *disconnected* from all the individuals of its component and is *connected* to all the individuals of the new component. Note that this is the same as saying that one individual duplicates the other and dies.
- At rate c a vertex is *removed* from its component and gets a *new type* with law θ on $[0, 1]$. In particular, the measure θ gives rise to different types of mechanisms: if θ is non-atomic, then incoming types are *always new* and create a new component, while if θ is atomic, then this is not so and θ is in fact a parameter.

The accompanying changes in the measure $\mu_{\mathcal{I} \times V}$ are left to the reader. We obtain the same dynamics of connection matrices as before, but now the *inheritable type of a vertex* allows us to impose new evolution rules.

► **Additional insertion and deletion of edges.** We discuss two choices:

- Simple model: At rate $a^+ \binom{n}{2}$ a pair (a, a') is chosen according to μ and the edge (a, a') is connected, at rate $a^- \binom{n}{2}$ a pair (a, a') is chosen according to μ and the edge (a, a') is disconnected, μ and $\mu_{\mathcal{I} \times V}$ are fixed, and h is jumping, as indicated.
- Modify the above by considering rates $a^+(\cdot), a^-(\cdot)$ that are functions of the population size of the type and/or of the age of the population, which is the maximal distance in the subpopulation. (For example, larger and older populations may be less pruned to loose a connection and more pruned to reinstall a lost connection). Here, $a^+(\cdot), a^-$ are bounded continuous functions.

4.1.2 Finite grapheme processes and their generators

The rules specified above via *jump transitions* and their *rates* uniquely define a *Markov pure-jump process* on \mathbb{G}_n or \mathbb{G}_∞ with additional deterministic growth of distances.

To find the operator of the grapheme evolution we use the observation that we can view the dynamics of the finite graphs as driven by a specific process, namely, the Moran process. A priori we have no guidance from the finite jump processes of abstract graphs on how to best embed it in a Polish space. Even for a metric structure on $\mathbb{G}_n^{(\cdot)}$ we have to work with the finite embedded graph (whose embedding we have defined in Section 4.1), which is given via the *genealogical Moran process*, respectively, the *genealogical Galton-Watson process* for the case of Dawson-Watanabe. This gives us an *enrichment* of the *abstract finite-graph dynamics* via a topological structure on each element of the *sequence of finite graphs* $(\mathcal{G}^{(n)})_{n \in \mathbb{N}}$, which is moving in (a subset of) the space \mathbb{U}_1 , respectively, \mathbb{U} of equivalence classes of ultrametric measure spaces, and represents the *genealogy* of the population of vertices evolving according to the Moran model. Namely, the distance of two individuals is twice the time to the most recent common ancestor. The sampling measure is the equidistribution, the weight is 2 at birth and 0 at death.

This gives us our \mathbb{G}_∞ -valued Markov processes, with \sim either bar or no bar, and dropping the parameters d, c, θ we write

$$(4.10) \quad (\mathcal{G}_t^{\text{FV},(n)})_{t \geq 0}, \text{ respectively, } (\mathcal{G}_t^{\text{DW},(n)})_{t \geq 0}.$$

These simple processes can be defined as solution to a *well-posed martingale problem*, with the test functions being the polynomials $\widehat{\Pi}^*$ from Section 3.1.1, as follows.

The *operator* (here a classical generator) of the martingale problem is of the following form, where we split into resampling and emigration/immigration:

$$(4.11) \quad \widehat{\mathcal{L}}^{*,\text{FV},d,c,\theta,(n)} = \widehat{\mathcal{L}}^{*,\text{FV},d,(n)} + \widehat{\mathcal{L}}^{*,\text{FV},c,\theta,(n),\text{em-im}} \text{ on } \mathbb{G}_n^\square,$$

$$(4.12) \quad \widehat{\mathcal{L}}^{*,\text{DW},b,c,\theta,(n)} = \mathcal{L}^{*,\text{FV},b,(n)} + \mathcal{L}^{*,\text{DW},c,\theta,(n),\text{em-im}} \text{ on } \bar{\mathbb{G}}_\infty^\square.$$

Recall that we used graphemes embedded in an ultrametric space \mathbb{U}_1 or \mathbb{U} that is a *continuum*. To that end we had passed to \mathbb{G}_n^\parallel and had given the finite system a structure as a *finite ultrametric probability measure space*. Consequently, we have operators arising from the structure of the ultrametric measure space in which we *embed our graph dynamics*, and obtain a grapheme for each n that in the limit as $n \rightarrow \infty$ converges to a continuum grapheme.

Consider first the case without immigration or emigration. Then the operator (generator in this case) $\widehat{\mathcal{L}}^{*,\text{FV},(n)}$ is given on $\widehat{\Pi}^*$ (recall (3.10)) by the formula

$$(4.13) \quad \widehat{\mathcal{L}}^{*,\text{FV},(n)} = \widehat{\mathcal{L}}^{*,\text{res},d} + \widehat{\mathcal{L}}^{*,\text{grow}}, \text{ respectively, } \mathcal{L}^{*,\text{DW},(n)} = \widehat{\mathcal{L}}^{*,\text{res},b} + \widehat{\mathcal{L}}^{*,\text{grow}}.$$

The first-order term, giving rise to the deterministic increase of distances with time, is given by

$$(4.14) \quad \widehat{\mathcal{L}}^{*,\text{grow}} \Phi^{(n),\varphi} = 2 \langle \nu^{(n)}, \nabla(\varphi) \rangle.$$

The second-order term is given via $\theta_{k,\ell}$, the operator replacing variable ℓ by variable k (see (4.21)),

$$(4.15) \quad \widehat{\mathcal{L}}^{*,\text{res},d} \Phi^{(n),\varphi} = \frac{d}{2} \sum_{k,\ell \in [n]} (\langle \nu^{(n)}, \varphi \circ \theta_{k,\ell} \rangle - \langle \nu^{(n)}, \varphi \rangle).$$

The generator $\mathcal{L}^{*,\text{DW},(n)}$ is similar, except that in (4.15) we drop the term $\langle \nu^{(n)}, \Phi^{(n),\varphi} \rangle$.

In both cases the polynomial for sampling *without* replacement is given by

$$(4.16) \quad \Phi^{(n),\varphi} = \langle \nu^{(n),\varphi}, \varphi \rangle,$$

and φ is a function of $\underline{r} = ((r(u_i, u_j)))_{i,j \in [n]}$ with $(u_i)_{i \in [n]}$ the n -sample from \mathcal{I} . Replacing $\varphi(\underline{r})$ by $\varphi(\underline{h})$, with \underline{h} the connection matrix, we have as operator for the functional ν on Π^* also

$$(4.17) \quad \widehat{\mathcal{L}}^{*,\text{FV},d,(n)} \Phi^{(n),\varphi} = \mathcal{L}^{*,\text{res},d} \Phi^{(n),\varphi}.$$

We also need the operator for *immigration* and *emigration*, which reads

$$(4.18) \quad \mathcal{L}^{*,\text{FV},c,\theta,(n),\text{em-im}} \Phi^{(n),\varphi} = \sum_{k \in [n]} (\langle \nu_{(k)}^{(n)}, \Phi^{(n),\varphi} \rangle - \langle \nu^{(n)}, \Phi^{(n),\varphi} \rangle),$$

where $\nu_{(k)}$ arises from ν by sampling the mark of the k -th member from a source S with measure $\mu \otimes \theta$ on $\mathcal{I}^* \otimes (S)^*$ rather than $\mu \otimes \kappa$.

Note that if in the formulas above we instead sample *with replacement*, i.e., draw from $\mu^{\otimes n}$, then we get

$$(4.19) \quad \widehat{\mathcal{L}}^{*,\text{res},d} \Phi^{(n),\varphi} = \frac{d}{2} \sum_{k,\ell \in [n]} (\langle \nu^{(n)}, \varphi \circ \theta_{k,\ell} \rangle - \langle \nu^{(n)}, \varphi \rangle) + O\left(\frac{1}{n}\right).$$

4.2 Construction: Fleming-Viot and Dawson-Watanabe grapheme operators

In order to write down the martingale problem arising in the $n \rightarrow \infty$ limit for the processes introduced in Section 4.1.1, we define the *operator* for the $\mathbb{G}^\square, \mathbb{G}^{\square,V}$ -valued Fleming-Viot process ($\sim = \text{no bar}$) and Dawson-Watanabe process ($\sim = -$). This dynamics has two components, one from the *abstract graph* (evolution of the connection-matrix distributions) and one from the evolution of the spaces in which we *embed* the abstract graph, here associated with elements from \mathbb{U}_1, \mathbb{U} , respectively, $\mathbb{U}_1^V, \mathbb{U}^V$. For both components of the evolution we have to specify the operators of the martingale problem.

We start with the pure Fleming-Viot and Dawson-Watanabe operator, where the *second-order term* represents the *essence of the grapheme diffusion process*. We first turn to the *dynamics* of the *connection matrix*, which induces a dynamics for the functional ν on $\mathbb{G}^\square, \mathbb{G}^{\square,V}$, and deduce from that the operator on $\mathbb{G}^{\langle \cdot \rangle}, \mathbb{G}^{\langle \cdot \rangle,V}$. Afterwards we turn to the evolution of the space in which we embed the vertices and combine this with the operator on \mathbb{G}^\square . Finally, we turn to the *first-order terms*.

4.2.1 Fleming-Viot, Dawson-Watanabe operators on functions of connection matrix

Consider test functions on \mathbb{G}^\square of the form

$$(4.20) \quad \Phi(G) = \Phi([\mathcal{I}, h, \mu]) = \int_{\mathcal{I}^m} \varphi \left((h(u_i, u_j))_{i,j \in [m]} \right) \mu(du_1) \dots \mu(du_m), \quad \varphi \in C_b([0, \infty)^m),$$

with $G \in \mathbb{G}^\square$. We denoted the set of test functions of this form by Π_{mon}^* and the generated algebra by Π^* . Consider the map $\tilde{\theta}_{i,j}$ acting on (u_1, \dots, u_m) via the replacement of i by j defined by

$$(4.21) \quad \tilde{\theta}_{i,j}(u_1, \dots, u_m) = (u_1, \dots, u_{i-1}, u_j, u_{i+1}, \dots, u_{j-1}, u_j, u_{j+1}, \dots, u_m).$$

Thus, $\tilde{\theta}_{i,j}$ maps an m -sample to an $(m-1)$ -sample, and $\tilde{\theta}_{i,j}$ induces the map $\theta_{i,j}$ on matrices $(h(u_k, u_\ell))_{k,\ell \in [m]}$. Set, with $\underline{u} = (u_1, \dots, u_m)$,

$$(4.22) \quad \mathcal{L}^{*,\text{FV},\text{res,d}}\Phi = \sum_{i,j \in [m]} \int_{\mathcal{I}^m} [\varphi(\theta_{i,j}(h(u_k, u_\ell)_{k,\ell \in [m]})) - \varphi((h(u_k, u_\ell)_{k,\ell \in [m]}))] \mu^{\otimes m}(d\underline{u}),$$

where we note that it suffices to consider φ of the form in (3.2). This extends easily to an operator that acts on Π^* and determines the evolution of $(\nu_t)_{t \geq 0}$.

Remark 4.1 (Lifting from $\mathbb{G}^\square, \mathbb{G}^{\langle \cdot \rangle}$ to $\mathbb{G}^{\langle \cdot \rangle}$). Note that in (4.20) for $G \in \mathbb{G}^\square$ we can write, for a unique $G^{\langle \cdot \rangle} \in \mathbb{G}^{\langle \cdot \rangle}$,

$$(4.23) \quad \Phi(G) = \int_{\{0,1\}^{m^2}} \varphi(\underline{h}) \nu^{(m)}(d\underline{h}) = \Phi^{\langle \cdot \rangle}(G^{\langle \cdot \rangle}),$$

with $\Phi^{\langle \cdot \rangle}$ a polynomial on $\mathbb{G}^{\langle \cdot \rangle}$, i.e., $\Phi^{\langle \cdot \rangle} \in \Pi^{*,\langle \cdot \rangle}$. Therefore in (4.22) we can define an operator $\mathcal{L}^{\langle \cdot \rangle,\text{FV}}$, with the operation $\theta_{i,j}^{\langle \cdot \rangle}$ instead of $\theta_{i,j}$, to obtain the

$$(4.24) \quad (\mathcal{L}^{\langle \cdot \rangle,\text{FV}}, \Pi^{*,\langle \cdot \rangle}, \Gamma^{\langle \cdot \rangle})\text{-martingale problem on } \mathbb{G}^{\langle \cdot \rangle},$$

which is solved by a solution of the $(\mathcal{L}^{*,\text{FV}}, \Pi^*, \Gamma^*)$ -martingale problem *lifted from* \mathbb{G}^\square to \mathbb{G}^\diamond . Recall, however, that \mathbb{G}^\diamond is not a Polish space, so standard theory may not apply, or may apply only after restriction to a subspace, for example, $\mathbb{G}_{\text{ultra}}^\diamond$, the equivalence classes in \mathbb{G}^\diamond induced by elements in $\mathbb{G}_{\text{ultra}}^\square$. \spadesuit

Next, we must *extend* this operator from Π^* to $\widehat{\Pi}^*$ by using the well-known operator on $\widehat{\Pi}^{*,\downarrow}$ for the dynamics $(\mathcal{U}_t)_{t \geq 0}$ on \mathbb{U}_1 , respectively, \mathbb{U} , and lifting these to operators on \mathbb{G}^\square , respectively, $\widehat{\mathbb{G}}^\square$, and also their marked versions on $\mathbb{G}^{\square,V}$, $\mathbb{G}^{\{\},V}$, in order to obtain the evolution of the *embedded grapheme* via the *evolution of the random environment* (specification of the embedding). Of course, we also have to pass to $\widehat{\Pi}^*$, the algebra where the joint law of ν and \mathcal{U} is determined.

4.2.2 Operators for functionals of the genealogy

We next specify the operator $\widehat{\mathcal{L}}^*$ on $\widehat{\Phi}^{*,V} \in \widehat{\Pi}^{*,V}$, i.e., we *extend the operator* to the algebra $\widehat{\Pi}^{*,V}$ generated by $\Pi^{*,V} \cup \Pi^{*,V,\downarrow}$. We recall the operator for the *genealogy-valued Fleming-Viot process* \mathcal{U} associated with our graphemes, which is the sum of four terms, and extend it to $\widehat{\Pi}^*$ or $\widehat{\Pi}^{*,V}$, which will be immediate after we look at the object in the right way, namely, by using that h is given by a specific distance function in our definition.

The limiting genealogical process \mathcal{U} for Fleming-Viot and Dawson-Watanabe population processes have been constructed as follows (see [GPW13, DG19, GRG21]). On the domain $\widehat{\Pi}^{*,V,\downarrow}$ (i.e., the functions in $\widehat{\Pi}^*$ restricted to functions not depending explicitly on h), call this operator $\widehat{\mathcal{L}}^{*,\text{FV}}$. This specifies a $(\widehat{\mathcal{L}}^{*,\text{FV}}, \widehat{\Pi}^{*,V,\downarrow}, \widehat{\Gamma})$ -martingale problem for the process \mathcal{U} on \mathbb{U}_1 or \mathbb{U} , for which a solution has been constructed in the above references and for which a *duality relation* has been checked, ensuring uniqueness of solutions. The dual is a *function-valued* process driven by a *coalescent*. At the same time, we obtain a martingale problem for the process $(\nu_t)_{t \geq 0}$, with ν_t arising as a functional induced by h_t for a given measure space $((\mathcal{I}_t, r_t), \mu_t)$ (see the references above). The four terms of the operator are as follows.

(I) The so-called *growth operator*, with growth referring to the increase of distances with time, is obtained after taking the test function defined earlier, replacing h in (4.20) by r , and acting on the test function as in (4.14).

(II) The *resampling operator* $\mathcal{L}^{*,V,\text{res}}$ defined above describes, in the \mathbb{U}_1^V -valued processes, the change of the genealogy due to the birth of new individuals, at distance 0 from the ancestor, and the sharing with other individuals of the distance to the most recent common ancestor. This is a second-order operator (acting on products of test functions in the same way as differential operators act by the product rule). The extension for the limit process runs well when we work with (\mathcal{I}, r) (hence on \mathbb{G}^\square), since in that case we have a similar form for the extended operator $\widehat{\mathcal{L}}^{*,\text{FV},\text{res}}$, namely,

(4.25)

replace connections h by distances r in (4.20) and (4.22) to get $\mathcal{L}^{*,\text{FV},\text{res}}$ on $\widehat{\Pi}_r^{*,V,\downarrow}$,

where the subscript r indicates that the functions do *not* depend on h . In this way we also get the dynamics of the spaces in which we embed our abstract graph dynamics, which

gives us the extension of the second-order term $\mathcal{L}^{*,V,\text{res}}$. To justify the latter, we use the approximation with finite processes, for which it was proved in [GPW09] that the same formula holds (by using an observation we explain in Section 7.1.2).

(III) The operator of *emigration* and *immigration* is first order, and is defined on a subclass of the polynomials on \mathcal{U}_t by (4.18). This definition works on $\widehat{\Pi}^{*,V}$ and its subsets. As above, we can use [DG19] to obtain the formula on $\widehat{\Pi}^{*,\downarrow}$ on $\widehat{\mathbb{G}}_n, \widehat{\mathbb{G}}_\infty$. We have to first do this on Π^* , and afterwards lift to $\widehat{\Pi}^*$. Here, we use again the approximation by finite- n models explained in the above reference.

(IV) The first-order operators capturing *mutation* and *selection*, denoted by $\mathcal{L}^{*,\text{FV},\text{mut}}$ and $\mathcal{L}^{*,\text{FV},\text{sel}}$, respectively, are defined on $\widehat{\Pi}^{*,\downarrow}$ in the literature (see [DGP12]). We have to extend these operators first to Π^* first, and afterwards to $\widehat{\Pi}^*$. We begin with the action on $\widehat{\Pi}^{*,\downarrow}$.

The ingredients are the rates m and s , the mutation kernel $M(b, db')$ from K to (K, \mathcal{B}_K) in (4.8), and the fitness matrix or fitness function $\chi: K \rightarrow [0, 1]$ from (4.9). Finally, we have to extend from $\widehat{\Pi}_r^{*,V}$ to $\widehat{\Pi}^{*,V}$, where now *products* of functions φ_h and φ_r depending on h , respectively, r appear.

In the sequel we use the polynomials defined in (3.30). For the *selection operator* on $\widehat{\Pi}^{*,\downarrow}$ we have

$$(4.26) \quad (\mathcal{L}^{\text{sel}}\Phi)(G) = s \sum_{k=1}^n \int_{(U \times V)^{n+1}} \varphi(\underline{r}(\underline{l})) [\psi(\underline{v})\chi(v_k) - \psi(\underline{v})\chi(v_{n+1})] \mu_{U \times V}^{\otimes n+1}(d(\underline{l}, \underline{v})).$$

The *mutation operator* on $\widehat{\Pi}^{*,\downarrow}$ reads as

$$(4.27) \quad (\mathcal{L}^{\text{mut}}\Phi)(G) = \sum_{k=1}^n [(M_k\Phi)(G) - \Phi(G)],$$

where, for $k \in \{1, \dots, n\}$, we set

$$(4.28) \quad (M_k\Phi)(G) = m \left\{ \int_K \Phi_k(G^{b'}) M(b_k, db') - \Phi(G) \right\},$$

with

$$(4.29) \quad \Phi_k(G) = \int_{(U \times K)^n} \varphi(\underline{r}(\underline{l})) \psi(\underline{b}^k) \mu_{U \times V}^{\otimes n}(d(\underline{l}, \underline{b}^k)), \quad \underline{b}^k = (b_1, \dots, b_{k-1}, b', b_{k+1}, \dots, b_n).$$

On Π^* the operator acts as above, leaving $\varphi(\underline{h})$ fixed but acting on $\psi(\underline{b})$. We extend this to $\widehat{\Pi}^*$ by following the same principle as before.

4.2.3 Dawson-Watanabe

It remains to specify the Dawson-Watanabe operator. Here, only the second-order term changes, namely, in (4.22) the second part with the minus term is dropped (see [DG19] for details).

4.3 Extension to non-completely connected components

Here the states are of a somewhat different nature, even though in the background there still is an autonomously fluctuating process in the form of a decomposition of the vertices into open and closed (potentially completely connected) subsets such that, *given* the whole (time-)path, the state is given by a *pruning* of the edges of the *potential edges* by the noise, where the potential edges are given by the processes with completely connected components studied previously.

We look at such *edge process* in a *random medium*. This means that, starting from our desired process \mathcal{G} , we now take the process that is associated with the \mathbb{G}^\square -valued process for the case where $a^+ = a^- = 0$, which we call $(\mathcal{G}_t^{(0,0)})_{t \geq 0}$ and which is associated uniquely with a process $(\mathcal{U}_t)_{t \geq 0}$ (see (1.26)). Then we take a field \mathcal{A}_t indexed by $U_t \times U_t$ of i.i.d. $\{0, 1\}$ -valued Markov spin-flip chains with rates a^+, a^- to generate a representative of the state of the desired process $\mathcal{G}^{(a^+, a^-)}$. To make this rigorous, we have to argue on \mathbb{G}^\square and write down the operator of $\mathcal{G}^{(a^+, a^-)}$. We modify the action of this operator on Π^* (the h -dependent functions) and also modify (1.26) by adding a further requirement on the evolution via the field $(\mathcal{A}_t(u_1, u_2))_{t \geq 0}$, namely, that no flip has occurred.

The evolution of the condition is well understood, since it is simply the underlying \mathbb{U}_1 (or \mathbb{U})-valued process $(\mathcal{U}_t)_{t \geq 0}$, which is autonomous since there is no feedback from the status of being an active or a non-active edge to the underlying process $\mathcal{G}^{(0,0)}$. The effect of the pruning is captured in a *conditional* martingale problem involving an induced process for the connection-matrix distributions. Given the condition, this is a distribution that arises as the pruned distribution of $(\nu_t)_{t \geq 0}$ for the solution $(\mathcal{G}_t^{(0,0)})_{t \geq 0}$ of the martingale problem *without* the additional mechanism of insertion and deletion, where the frequency of pruning is given via the value a_t of the frequency of current virtual states among all eligible states. We have to justify this splitting of the dynamics.

For the finite approximations, we can split the process into an autonomous process \mathcal{U}^n , creating the potential edges, and an evolution of the edges by the field \mathcal{A} acting on the potential edges as a spin-flip dynamics, giving \mathcal{G}^n . The evolution of the edges does not influence \mathcal{U}^n . Therefore the splitting of the condition (autonomous) and the flipping of virtual edges works well.

We have to show that this also holds in the limit dynamics, which can be deduced by using *only* the martingale problem. For that we observe that the operator can be split into the action of the operator corresponding to $\mathcal{G}^{(0,0)}$ and the operator acting on functions of the state of the edge, i.e., the specification active or virtual on functions that are products of functions of each of the components, one of which observes the active-virtual information, representing the spin-flip dynamics. On $\widehat{\Pi}^*$ restricted to functions not depending on h , i.e., $\widehat{\Pi}^{*, \downarrow}$, we find the generator of \mathcal{U} . On Π^* we obtain the spin-flip generator acting on h . On $\widehat{\Pi}^*$ it suffices to consider the product of a function from $\widehat{\Pi}^{*, \downarrow}$ and a function from Π^* .

The additional operator acts on functions of the form as in (3.30) like

$$(4.30) \quad (\mathcal{L}_t^{\text{ins, del}} \Phi)(G) = \left(a^+ \sum_{k, \ell=1}^n (\Phi_{k, \ell}^{+, t} - \Phi) + a^- \sum_{k, \ell=1}^n (\Phi_{k, \ell}^{-, t} - \Phi) \right) (G),$$

where, for $G = (U, r, h, \mu)$, we define at time t , with the states at time t ,

$$(4.31) \quad \Phi_h^{+,t}(G) = \int_{U^n} \varphi_h(\underline{h}^{k,\ell,\sim}(\underline{u})) \varphi_r(\underline{r}(\underline{u})) 1(r(u_k, u_\ell) < 2t) \mu^{\otimes n}(d\underline{u}), \quad \sim \in \{+, -\},$$

and

$$(4.32) \quad (\underline{h}^{k,\ell,+}(\underline{u}))_{(i,j)} = 1 \text{ for } k = i, j = \ell, \sim = +,$$

and

$$(4.33) \quad (\underline{h}^{k,\ell,\sim}(\underline{u}))_{(i,j)} = \underline{h}(\underline{u}) \text{ otherwise,}$$

and similarly for $\Phi^{-,t}$.

We can avoid *time-dependent terms* if we introduce *marks* characterising the *founder* of a family surviving at time t for all $t > 0$ and requiring identity of type in the indicator of (4.31) instead of having the same time-0 ancestor.

5 Proofs for state spaces and topologies

In Section 5.1 we derive the properties of the *algebras of test functions* $\widehat{\Pi}^*, \widehat{\Pi}^{*,V}$ on the spaces $\widetilde{\mathbb{G}}^\square, \widetilde{\mathbb{G}}^{\square,V}$, respectively, $\widetilde{\mathbb{G}}^\diamond, \widetilde{\mathbb{G}}^{\diamond,V}$, with \sim being bar or no bar. In Section 5.2 we derive the properties of the topologies that were introduced, in particular, the property of being *Polish*.

5.1 Algebra of test functions

In this section we prove Proposition 3.1–3.2 and Proposition 3.6–3.7. To do so, we exploit the literature for the respective properties on $\mathbb{U}_1, \mathbb{U}, \mathbb{U}_1^V, \mathbb{U}^V$, respectively, the more general version on $\mathbb{M}_1, \mathbb{M}, \mathbb{M}_1^V, \mathbb{M}^V$, of which the former are closed subspaces (see [GPW09, DGP11, DG19]). What remains on $\mathbb{G}^\square, \mathbb{G}^{\square,V}$ (or $\widetilde{\mathbb{G}}^\square, \widetilde{\mathbb{G}}^{\square,V}$), and is not covered by the literature on genealogies of populations for functions on \mathbb{M}_1, \mathbb{M} and their marked versions, is to provide the same information for functions of $[U_t, r_t, \mu_t]$ that are used to study the *connection-function part*, i.e., the functional $(\nu_t)_{t \geq 0}$ induced by h , and to get the complete object $([(\mathcal{I}_t, r_t), h_t, \mu_t])_{t \geq 0}$.

However, on $\mathbb{G}_{\text{ultr}}^\square, \widetilde{\mathbb{G}}_{\text{ultr}}^\square$ the function h can be viewed as an ultrametric, and so (\mathcal{I}, r, μ) and (\mathcal{I}, r, h) are of a similar structure, so that we can get the properties of both $\widehat{\Pi}^{*,\downarrow}$ and Π^* from the literature on \mathbb{U}_1, \mathbb{U} . It remains to get the combined structure from the algebra that is generated by taking products of elements of the two algebras, as studied in [DGP23b] in detail. In our context we can use that the h -component is a continuous functional of the r -component of the state. We have to also extend to $\widetilde{\mathbb{G}}$. However, since $\widetilde{\mathbb{G}}$ consists of objects that are pairs of a total mass component and an element of \mathbb{G} , this can be achieved in the same way as for the space \mathbb{U} from \mathbb{U}_1 (see, for example, [DG19, GRG21]). In the general case $\mathbb{G}^\square, \widetilde{\mathbb{G}}^\square$, we use the results on \mathbb{M}_1, \mathbb{M} (recall (7.4) and the sequel).

5.2 Polish spaces

In this section we prove Propositions 3.3 and 3.7. We have to show that, by adding h to the structure $((U, r), \mu)$ to get $[(U, r, h), \mu]$, we get a Polish space. This fact is based on the knowledge that the equivalence classes of a metric measure space $[U, r, \mu]$ from \mathbb{U}_1 or \mathbb{U} form a Polish space in the Gromov weak topology [GPW09]. We have that also $\mathbb{G}_{\text{ultr}}^{\square}$ is a Polish space in the topology specified in Section 3. Namely, since h is just inducing a second ultra-metric by defining the h -path connected components as the 1-balls, this space is like a *multi-metric* measure space (as studied in [DGP23b]), for which the claim holds. Note, however, that here we are in the simpler setting of $\mathbb{G}_{\text{comp}} \subseteq \mathbb{G}_{\text{ultr}}$, with the h that we used for our models, because one is a coarse-graining of the other, so that convergence of the r -topology implies convergence of the h -topology on those sets of states (because we have ultrametric spaces). The same applies in the more general setup where h can be viewed as a metric but not an ultrametric, where the second metric generates a coarser Borel-algebra.

6 A prelude on duality

In this section we consider the dynamics without additional insertion/deletion and introduce: (i) the dual process \mathcal{C} driven by the coalescent process $C^t = (C_s^t)_{s \in [0, t]}$, defined for every time horizon t , (ii) the duality function H . In Section 7 we will extend the duality to the general setting.

(i) The dual process. The state space for the dual process is $E'_n = S_n \times D_n$, with S_n the set of partitions of $[n] = \{1, \dots, n\}$, D_n the set of $(n \times n)$ - \mathbb{R}_+ -valued distance matrices. The dynamics is such that each pair of partition elements, *coalesces* at rate d , and the distance of each pair of different vertices *grows* at speed 2. This defines an E' -valued Markov jump process, $(\mathcal{C}_s = (c_s, \underline{r}_s, \underline{h}_s))_{s \geq 0}$, with a deterministically evolving part (distance growth). Its generator is given as follows: for $(C, \underline{r}) \in S_n \times D_n$, $g : S_n \times D_n$ bounded and continuously differentiable in distances,

$$(6.1) \quad G^{\text{dual}}g(C, \underline{r}) = d \sum_{\varpi_1 \neq \varpi_2 \in C} \left(g(C \setminus \{\varpi_1, \varpi_2\} \cup \{\varpi_1 \cup \varpi_2\}, \underline{r}) - g(C, \underline{r}) \right) + 2 \sum_{1 \leq i < j \leq n} \frac{\partial}{\partial r_{i,j}} g(C, \underline{r}) \mathbf{1}_{i \sim_c j},$$

where as usual we write $i \sim_c j$ if and only if there is $\varpi \in c$ with $i, j \in \varpi$.

(ii) The duality function. The duality function combines the duality relation for the two components of \mathcal{G}_t , namely, for \mathcal{U}_t in \mathcal{C}_t^t and for the connection function h in \mathcal{C}_t^t (the first is well known in the literature, the second is specific for graphemes). Both have to be integrated into a single duality function, link two \mathbb{G}^{\square} -grapheme processes, and give the strong duality.

First, we define the duality function for the process $(\mathcal{U}_t)_{t \geq 0}$ and the dual $(C_s^t)_{s \in [0, t]}$. This is obtained by putting in the duality function for \mathcal{G} and $(C_s^t)_{s \in [0, t]}$ in (6.2) below the

function $\varphi_h \equiv 1$. Define $H: E \times E' \rightarrow \mathbb{R}$, with $E = \mathbb{G}^\square$ and $E' = \bigcup_{n \in \mathbb{N}} E'_n$, as

$$(6.2) \quad H([U, r, h, \mu], (C, \underline{r}')) := \int_{U^n} \varphi_r(\underline{r}^C(\underline{u}) + \underline{r}') \varphi_h((h^C(i, j))_{1 \leq i < j \leq n}) \mu^{\otimes n}(d\underline{u}),$$

where $n \in \mathbb{N}$ such that $C \in S_n$, r^C is the distance matrix of the sample \underline{u} given by $r_{i,j}^C = r_{C(i), C(j)}$, with $C(i)$ being the partition element containing $i \in [n]$, and $h^C(i, j) = 1$ if and only if $j \in C(i)$.

Next, we turn to h , and hence to the functional $(\nu_t)_{t \geq 0}$, and combine them into a duality for the process to $(\mathcal{G}_t)_{t \geq 0}$. The duality w.r.t. H as well as the strong duality for \mathcal{G} are, for *our* dynamics, in fact consequences of the known *strong* duality relation for \mathcal{U} . Recall that the connection function arises as a *function* of the distances, both for the process and for the dual process. Namely, a decomposition of the sets U_t, \tilde{U}_t (for process, respectively, dual process) gives the completely connected components as $2t$ -balls in both cases. These are given by the disjoint decomposition into open and closed $2t$ -balls, represented by partition elements in the dual, which have a closure that is a *completely connected component*, and their (disjoint) union is \tilde{U}_t . Therefore the strong duality for \mathcal{U} and \mathcal{C} *implies* the *strong duality of the grapheme process* \mathcal{G} and \mathcal{C} in our setting.

However, if we want to get information about the uniqueness of the martingale problem for \mathcal{G} , then we need to obtain the equivalence of $[U_t, r_t, h_t, \mu_t]$ and $[\tilde{U}_t, \tilde{r}_t, \tilde{h}_t, \tilde{\mu}_t]$ by checking that expectations of the functions given by H are the same, *just* using the forward and the backward (dual) operators (see (6.3) below). This has been done in the literature for the \mathcal{U} -part, because we saw that the h -part has the same structure, so that the relation must hold for our operator on \mathbb{G}^\square , because the merging of variables allows us to extend from $\widehat{\Pi}^{*, \downarrow}$, and from Π^* functions to products, and hence to $\widehat{\Pi}^*$. The general criterion we use here is as follows. If the family $\{H(\cdot, c), c \in E'\}$ is law determining, then we have to check the following generator relation (which is done in [GPW13] for our case):

$$(6.3) \quad (G^{a, c, \theta} H)(\cdot, c)(u) = (G^{\text{dual}} H)(u, \cdot)(c), \quad u \in E, c \in E', \quad a = d,$$

while for $a = b$ we must add the term $\mathfrak{V}(\cdot)H(u, c)$ to the right-hand side.

7 Proofs for grapheme diffusions with completely connected components

In Section 7.1 we explain the main ideas, list the tasks that remain, and give the key arguments for Theorems 1.9, 1.14, 1.18, 1.20 and 1.25. In Section 7.2 we prove their corollaries. Theorem 1.25 is extended in Section 7.3, and details on the references are worked out in Section 7.4.

7.1 Main idea for the proof

As in [AdHR21], for the special case where c, θ are both 0 we have chosen simple evolution rules for finite graphs $([n], h, \mu)$ arising from population dynamics of individuals in \mathcal{I} . In

addition to this reference, we now use that we also have complete information about the evolution of the embedding space $([\mathcal{I}_t, r_t, \mu_t])_{t \geq 0}$ process. Namely, these processes have been treated in the literature, as far as the analogues of Theorems 1.9 and 1.20, and Theorem 1.26 for the genealogy \mathcal{U} , are concerned.

We must complete the proof of the claims on \mathbb{G}^\square , where the *connection function* h is an additional part besides the *genealogy* \mathcal{U} . Moreover, the *joint law* of these two components has to be treated also. This has to be done by using that the process \mathcal{G} is given as a *functional* of the *genealogy process* \mathcal{U} associated with our population genetic models, namely, $(\mathcal{U}_t^{a,c,\theta})_{t \geq 0}$ with $a = b, d$ for Fleming-Viot, respectively, Dawson-Watanabe. In the sequel we specify the theorems in the literature that relate to our theorems, and handle the two problems just mentioned. We do so in the last two points in Section 7.1.1, where we exhibit the specific structure of the processes we deal with, which substantially facilitates the task and the details of the argument given in Section 7.1.2. Additional ideas are needed in the presence of mutation, which are postponed to Section 7.3.

7.1.1 The process: from genealogies to graphemes

In this section we look at the process $\mathcal{U}_t = ([\mathcal{U}_t, r_t, \mu_t])_{t \geq 0}$ associated with our models, recall their characterisation as solution to a well-posed martingale problem, and describe path properties, equilibria and convergence to them and converging finite approximations, in order to see how we can include the connection function h into the picture. This involves both initial states and dynamics.

A key property of the dynamics is that a large part of \mathbb{G}^\square is transient and is left instantaneously toward $\mathbb{G}_{\text{comp}}^\square$, and the states are completely connected components built by the time- t descendants of finite or countably many (for $c, \theta > 0$) founders from the time-0 population. We can start in the marked case in *any* initial configuration of types, but at positive times our evolution leads to a state with finitely or countably many types, and consequently to completely connected components only. This means that our evolution takes place in a subset of $\mathbb{G}_{\text{ultra}}^{\square,V} \subseteq \mathbb{G}^{\square,V}$ and, in fact, stays in $\mathbb{G}_{\text{comp}}^{\square,V}$ when started there:

$$(7.1) \quad \mathbb{G}_{\text{comp}}^{\square,V}, \mathbb{G}_{\text{ultra}}^{\square,V} \text{ are } \textit{dynamically closed} \text{ subsets of } \mathbb{G}^{\square,V}.$$

Only the latter is *closed topologically*.

(i) Initial states. First we have to decide what *initial states* $[(\mathcal{I}_0, r_0), h_0, \mu_0]$ are *possible* for the grapheme evolution, since we have defined it as functional of \mathcal{U} . It turns out that we can choose the initial state only with $\mathcal{U}_0 \in \mathbb{U}_1$, respectively, \mathbb{U} for reasons of *consistency*. Namely, in order to obtain *càdlàg paths* of graphemes, we have to take h such that the initial h_0 is determined by \mathcal{U}_0 . In particular, we still have to give a version of the formula relating \mathcal{G} to these initial states, which are different from the single-root case for \mathcal{U}_0 .

First, we consider a special case: start the genealogical part in $[\{1\}, \underline{0}, \delta_1]$, i.e., *all* initial particles form a single root, and we follow the arguments for the dynamics based on (1.26) to get \mathcal{G} from \mathcal{U} . How do we get general initial states? We embed the process \mathcal{U} in a $[0, 1]$ -marked process $\mathcal{U}^{[0,1]}$. We *dissolve* the root by taking as type space $V = ([0, 1], \mathcal{B}_{[0,1]})$ and

starting in $\mu_{U \times V, 0} = \delta \otimes \pi$, with π the Lebesgue measure on $[0, 1]$, so that every initial individual gets a different type. In this way, at time $t > 0$ we have finitely or countably many atoms in $\mu_{U_t \times V, t}$, and this enables us to distinguish all *founders* by their type in $[0, 1]$. We call founders at time $s > 0$ all individuals having descendants at some time $t > s$. In other words, introducing *inheritable types* allows us to simulate all different \mathcal{U}_0 -states by using a marked model and identifying specific marks.

We have to adapt (1.26) when we do *not* start the process \mathcal{U} in a single root. First, if the initial state is a more general $\mathcal{G}_0 = [(U_0, r_0), h_0, \mu_0] \in \mathbb{G}_{\text{comp}} \subseteq \mathbb{G}_{\text{ultr}}^{\square}$, then we add a further condition, namely, that the types of two vertices have to agree if and only if there is an edge between them. We then define the process \mathcal{G} by using the functional \mathcal{U}_t^* instead of \mathcal{U}_t in the analogue of (1.26), where \mathcal{U}_t^* is the t -truncation of \mathcal{U}_t . Therefore h_0 has a form that represents the partition of the set of founders for the later states at time $t > 0$ in completely connected components (i.e., 1 if and only if both arguments are from one element of the partition), so that components of the first initial condition are merged according to \mathcal{G}_0 . This gives us the Fisher-Wright model of [AdHR21], also for non-trivial c, θ .

For general h_0 , we decompose $h_0 = h_0^1 + h_0^2$, with h_0^1 associated with completely connected components and h_0^2 the remainder, and we follow the diffusion corresponding to $h_0^1 + \tilde{h}_0^1$, with \tilde{h}_0^1 the 0-component on the points where h_0^2 was supported (which is an *instantaneous jump* to a state treated above). This gives us a càdlàg path, but one that jumps at an infinite rate at time 0.

(ii) Dynamics. We next formulate our basic grapheme diffusions arising from graph evolutions, as described in Section 1.3 and above. Take the $\mathbb{G}^{\square, V}$ -valued or $\mathbb{G}^{\square, V}$ -valued processes \mathcal{G} , driven by the \mathbb{U}^V -valued processes $(\mathcal{U}_t^{d, c, \theta})_{t \geq 0}$, $(\mathcal{U}_t^{b, c, \theta})_{t \geq 0}$ as in Definition 1.5, and (1.26):

- Fleming-Viot process with rate $c \geq 0$ of immigration and emigration from a source $[[0, 1], \text{Euclidean distance}, \theta]$, $\theta \in \mathcal{M}_1([0, 1])$, and sampling rate d : $\mathcal{G}^{\text{FV}, c, d, \theta}$.
- Add selection to get $\mathcal{G}^{\text{FV}, c, d, \theta, s, m}$ in the case $m = 0$, written $\mathcal{G}^{\text{FV}, c, d, \theta, s}$, which allows for the same reasoning as above despite the marks, since the latter are passed on to the descendants when $m = 0$.
- Dawson-Watanabe process with rate $b > 0$ on $[0, 1]$, with rate $c \geq 0$ of immigration and emigration, source $\theta \in \mathcal{M}_{\text{fin}}([0, 1])$: $\mathcal{G}^{\text{DW}, c, b, \theta}$.

We need information on \mathcal{U} . To cover non-trivial c, θ , we use [GPW13, DGP12, DG19] and [GSW16] to get *existence* and *uniqueness* of the solution to the martingale problem and *continuous paths* of the \mathbb{U}_1 -valued, respectively, \mathbb{U} -valued *driving processes* \mathcal{U} and their *genealogy* underlying \mathcal{G} . Therefore, in our definition of \mathcal{G} with h_t based on \mathcal{U}_t at each time t , we have obtained as grapheme diffusion a *unique stochastic processes* with values in \mathbb{G}^{\square} with *continuous paths* (since h is continuous in that case), and it is known from the references mentioned that $(\mathcal{U}_t)_{t \geq 0}$ has the strong Markov and Feller property, and for $t \rightarrow \infty$ converges to a limiting *equilibrium* state. A further observation is the

following: the process \mathcal{U} is approximated in *path space* by individual-based models (with n individuals) in the limit as $n \rightarrow \infty$. By our formula relating \mathcal{U} and \mathcal{G} , we have, in particular, also defined a *finite approximation* of \mathcal{G} via grapheme processes with values in \mathbb{G}_∞ , called \mathcal{G}^n , in path space. For detailed references, see Section 7.4.

In order to prove Theorems 1.9, 1.14 and 1.18, we want to view the processes induced by \mathbb{U} as *strong Markov processes* solving a *well-posed martingale problem* on state space \mathbb{G}^\square , and get the $t \rightarrow \infty$ limit as an *equilibrium* of a Markov process. Furthermore, this process has to be proven to be second-order, i.e., to be a *diffusion*. Here, the order is defined by following the generalisation of the concept of order of differential operators to operators acting on functions $\mathbb{G}^\square, \mathbb{G}^\{\}, \mathbb{G}^\langle \rangle$ based on \mathbb{U}_1 and \mathbb{U} in [DGP12], which we recalled in Remark 1.7. The issue is whether the criteria that were used hold for the lifted operators as well. This means that, based on our definition of a version of \mathcal{G} for a given genealogy process \mathcal{U} , we need to show that the respective properties on \mathbb{U}_1 or \mathbb{U} indeed *imply* the properties of the process \mathcal{G} on both $\mathbb{G}^\square, \mathbb{G}^\langle \rangle$. In turn, this means that we need to identify the proper quotations of results on \mathcal{U} and formulate the *implication* for the evolution of the *connection-matrix distribution* process and the *joint law* of the process $(h_t)_{t \geq 0}$ with the underlying $(\mathcal{U}_t)_{t \geq 0}$. The transfer arguments are set up as follows.

The key points in the argument for our grapheme processes are the following: (i) lift the results on $\mathcal{U}^{d,c,\theta}, \mathcal{U}^{b,c,\theta}$ from \mathbb{U}_1, \mathbb{U} to \mathcal{G}^\square in order to obtain a solution to the martingale problem for \mathcal{G}^\square , by taking a functional from the underlying \mathbb{U} -valued process, as specified in 1.5 and (1.26), respectively, for general initial conditions; (ii) establish *duality* based on the duality established for \mathcal{U} , which makes the solutions unique; (iii) lift the *approximation results* on \mathbb{U} -valued processes to \mathbb{G}^\square , i.e., lift $\mathcal{U}^{(n)} \rightarrow \mathcal{U}$ as $n \rightarrow \infty$ to $\mathcal{G}^{(n)} \rightarrow \mathcal{G}$ as $n \rightarrow \infty$, i.e., convergence of the grapheme Fleming-Viot process. In particular, establish the statements about the process $(\nu_t)_{t \geq 0}$ (ν_t is the law of h_t under the sampling measure $\mu_t^{\otimes 2}$) based on the properties of the \mathcal{U} -processes.

All these points are resolved in the following proposition.

Proposition 7.1 (From \mathcal{U} to \mathcal{G}). *The following hold for the various model classes considered, except for insertion/deletion (treated in Section 8).*

- (a) *If \mathcal{U}^n converges to \mathcal{U} in path space, then \mathcal{G}^n converges to \mathcal{G} for the \mathbb{G}^\square -valued process, as well as for the $\mathbb{G}^\langle \rangle$ -valued functional.*
- (b) *If \mathcal{U} satisfies the $(\mathcal{L}^{\mathcal{U}}, \widehat{\Pi}^{*,\downarrow}, \Gamma_{\mathcal{U}})$ -martingale problem and the latter has a unique solution, then the $(\mathcal{L}^{\mathcal{G}}, \widehat{\Pi}^*, \Gamma_{\mathcal{G}})$ -martingale problem associated with \mathcal{U} has a unique solution \mathcal{G} that is given by the construction in Section 1.3.3. Furthermore, the functional $(\nu_t)_{t \geq 0}$ solves the martingale problem of (4.24) (see Remark 4.1, where the growth operator is suppressed and the domain of the operator is Π^*).*
- (c) *The equilibrium of \mathcal{G} is obtained by applying the construction of Section 1.3.3 to the equilibrium state of \mathcal{U} .*
- (d) *The strong dual representation of \mathcal{U} implies the strong dual representation of \mathcal{G} , in particular, the H -duality relation lifted from \mathbb{U}_1 to \mathbb{G}^\square .*
- (e) *The second-order property of the operators of \mathcal{G} can be transferred from \mathcal{U} . \square*

The details on how to prove these results are given next.

7.1.2 Proofs: functionals of connection-matrix distributions and joint laws

The structure of grapheme spaces. We are in a similar situation as in the theory of measure-valued processes, where, for example, the Fleming-Viot process in any positive time enters the set of atomic measures (which is a trap), after which it has a simpler structure. Here we consider the version of the process on $\widetilde{\mathbb{G}}_{\text{ultra}}^{\square}$ (recall the definition (1.28)). The task to move from \mathcal{U} to \mathcal{G} is facilitated by the structural property of such a grapheme \mathcal{G} : h is in fact a *coarser* ultrametric than r , the one of the space (U, r, μ) in which we embed the graph. Namely, we can view h as a function defining an ultrametric taking three values, 0, 1 and a further value > 1 , and define a *closed subspace* \mathbb{U}_1^* of \mathbb{U}_1 (and similarly for \mathbb{U}). In *our* models, the ultrametric structure on the state given by $\mathcal{U}_t \in \mathbb{U}_1$ is a *refinement* of the state from \mathbb{U}_1^* given by h , namely, $(U_t, h_t, \mu_t^\downarrow)$, the arrow indicating restrictions of μ_t to the coarser σ -algebra of the ultrametric h .

The more general properties are the following :

- For $t > 0$ our process has only completely connected components, i.e., lives in $\mathbb{G}_{\text{comp}}^{\square}, \mathbb{G}_{\text{comp}}^{\diamond}$.
- The solution of the martingale problem for the h -functional can itself be seen as a \mathbb{U}_1^* -valued process for some closed subspace \mathbb{U}_1^* of \mathbb{U}_1 , defined by the property that the metric takes only three values, 0, 1 and > 1 , with operators similar to the ones obtained in population genetics for the genealogy process. Being a coarse-graining, the solution is compatible with the underlying r .

The above observation defines a *grapheme space* $\mathbb{G}^* \subseteq \mathbb{G}^{\square}$: (i) where our processes turn out to live and where the underlying embedding in a metric space $\mathcal{U} = [U, r, \mu]$ is required to be from \mathbb{U}_1 (or \mathbb{U}); (ii) h is required to be coarser than r . In that case we can give an embedding explicitly, which allows us to view $\mathbb{G}_{\text{comp}}^{\diamond}$ or $\widetilde{\mathbb{G}}_{\text{comp}}^{\diamond}$ as contained in a closed subspace of the space \mathbb{U}_1^* of ultrametric probability measure spaces, such that the ultrametric takes on only three different values via an embedding. Here are the details of the construction.

When all connected components are completely connected, we can project, respectively, embed the subset of graphemes $\mathbb{G}_{\text{comp}}^{\square} \subseteq \mathbb{G}^{\square}, \mathbb{G}_{\text{comp}}^{\diamond} \subseteq \mathbb{G}^{\diamond}$ into \mathbb{U}_1 via a map ψ as

$$(7.2) \quad \mathbb{G}_{\text{comp}}^{\square} \xrightarrow{\psi} \mathbb{U}_1, \quad \text{the image } \mathbb{U}_1^{\text{graph}} \text{ of } \mathbb{G}_{\text{comp}}^{\square} \text{ is contained } \mathbb{U}_1^*,$$

where \mathbb{U}_1 is the space of equivalence classes of ultrametric measure spaces. Here, on \mathbb{G}^* , ψ is given by

$$(7.3) \quad [(\mathcal{I}, r), h, \mu] \xrightarrow{\psi} [\mathcal{I}, r^*, \mu^*] \in \mathbb{U}_1^*,$$

where μ^* is the restriction of μ to the σ -algebra associated with r^* , and r^* is an ultrametric given by

$$(7.4) \quad r^*(x, y) = \begin{cases} 0, & \text{if } x = y, \\ 1, & \text{if } h(x, y) = 1, \\ \infty, & \text{if } h(x, y) = 0 \text{ and } x \neq y. \end{cases}$$

Via the transformation $r \rightarrow 1 - e^{-r}$ we obtain an element of the classical \mathbb{U}_1 (with *finite* values for the metric), with distances taking values $\{0, 1, \infty\}$, respectively, $\{0, e^{-1}, 1\}$. On $\mathbb{G}_{\text{ultra}}^{\square}$, this mapping is a coarsening of the state, while on $\mathbb{G}_{\text{ultra}}^{\langle \rangle}$ it is an embedding and

$$(7.5) \quad \mathbb{U}_1^* \text{ is a closed subspace of } \mathbb{U}_1.$$

In the general case $\mathbb{G}^{\square}, \mathbb{G}^{\langle \rangle}$ the function h can be viewed as metric (not necessarily an ultrametric) when the third value is less than 2.

Proof of Proposition 7.1. We proceed in five steps (i)-(v).

(i) We first consider initial states in $\mathbb{G}_{\text{comp}}^{\square}$, or the single-root case, generating a closed set topologically and dynamically.

Case \mathbb{G}^{\square} . The above observation (together with the fact that balls in ultrametric spaces are open and closed) yields that the weak convergence of $\mathcal{U}_t^{(n)}$ to \mathcal{U}_t in \mathbb{U}_1 as $n \rightarrow \infty$ implies the convergence of $\mathcal{G}_t^{(n)}$ to \mathcal{G}_t when the component $h^{(n)}$, respectively, h is taken into account. This is due to the definition of $\mathcal{G}^{(n)}, \mathcal{G}$ as specific functionals of $\mathcal{U}^{(n)}, \mathcal{U}$. In particular, if we look at the finite- n versions of the functional ν giving the process $(\nu_t)_{t \geq 0}$, then we see that this process converges based on what is known about the underlying processes $\mathcal{U}^{(n)}$ and their limit \mathcal{U} , namely, their $2t$ -ball decomposition is also the decomposition in completely connected components, and the former converges as object in \mathbb{U}_1, \mathbb{U} .

Case $\mathbb{G}^{\langle \rangle}$. Consider evolution rules where the population dynamics has genealogies that evolve as \mathbb{U}_1 -valued processes, arising as limits of individual-based models, which are *strong Markov processes* $(\mathcal{U}_t)_{t \geq 0}$ and are characterised by a well-posed martingale problem. We have defined a function h , based on \mathcal{U} , so that we obtain a functional (7.3) giving the connection-matrix distribution Ψ that maps elements of \mathbb{U}_1 to $\mathbb{G}^{\langle \rangle}$, and gives rise to a *well-defined grapheme evolution* $(\mathcal{G}_t^{\langle \rangle})_{t \geq 0}$ on the state space $\mathbb{G}^{\langle \rangle}$, of which we need that it arises as the limit of finite-grapheme evolutions. In particular, the subgraph counts converge to the ones obtained as moments under the connection-matrix distribution. The latter comes from the fact that the frequency of k -subgraphs can be expressed in terms of $(\mu_n(t)(H^i))_{i \in \mathbb{N}} = (\mu_n^*(t)(H^i))_{i \in \mathbb{N}}$ with H^i , $i \in \mathbb{N}$, the completely connected components, and from the fact that the expressions converge when the processes $(\mathcal{U}_t^{(n)})_{t \geq 0}$ converge to $(\mathcal{U}_t)_{t \geq 0}$, since the completely connected components are open and closed balls in the metric given on \mathcal{U}_t .

We have to identify the possible *initial states* for the $n \rightarrow \infty$ limit, i.e., we need to know how càdlàg or continuous path moves inside $\mathbb{G}_{\text{comp}}^{\square}$, or move from the complement to the inside as the entrance law from the single-root case as continuous path, or instantaneously jump inside at time 0 and afterwards evolve like the latter entrance law. Finally, we can have mixtures of these three extreme types of behaviour.

The above means that we have to show that if H is not $\equiv 0$ or $\equiv 1$ within sets of a decomposition of the basic set, or is a mixture of both, then there is the instantaneous jump at time $t = 0$ into the $H \equiv 0$ state, taking place on the complement of the union of the completely connected components. In other words, looking first at the case $c > 0$, only completely connected components can avoid immediate destruction, by the growing

of completely connected components, founded by single vertices whose descendants survive until time $t > 0$ that are all connected due to the resampling mechanism, but to *no* vertices that are *not* in their clique. Thus, descendants of immigrants are connected precisely with all their subfamilies, which reduces the need to show this for all initial survivors for $t > 0$, based on the $c = 0$ mechanism. A similar but more subtle mechanism works for the case $c = 0$. Here we have to focus on the edges between different *weakly* completely connected components, where connections to other components or points are allowed. For the part of $h \equiv 0$, we have the same situation as with immigrants. We can focus on the completely connected components. We need that the change by resampling leads to a loss of connection to the outside of the weakly completely connected component from the lost vertices, and duplication of the connections to the outside that go out from the duplicated vertex. Therefore the unbiased subsequent changes are of order n , whereas the number of changes of vertices in a weakly completely connected component is of order 1. Therefore, as $n \rightarrow \infty$ we hit 0 with the number of edges between weakly completely connected components before the vertices begin to fluctuate. This jump follows also from the convergence of $\mathcal{U}^{(n)}$ to \mathcal{U} , with finitely many surviving initial individuals or immigrants entering before a time s for positive time in \mathcal{U} , absorbing the full measure and thus forcing out the complement of the closure of $\widetilde{\mathbb{G}}_{\text{comp}}^{\square}$ by right continuity.

Thus, altogether, we have proved part (a) of Proposition 7.1 on \mathbb{G}^{\square} and on \mathbb{G}^{\diamond} .

(ii) Next, we show that the structure introduced above implies that \mathcal{G} on \mathbb{G}^{\square} solves a well-posed martingale problem for $\mathcal{G}_0 \in \mathbb{G}_{\text{comp}}^{\square}$ or the single-root case. The action of the operator on $\widehat{\Pi}^{*,\downarrow}$ arises from the construction of \mathcal{G} as a functional of \mathcal{U} , and \mathcal{U} itself evolves according to an operator not containing information on h , and follows the action of the operator of the process \mathcal{U} acting on $\Pi^{*,\downarrow}$, which gives the first piece of our martingale problem for \mathcal{G} . We have to add functions from the algebra Π^* that determine h , as follows: h is a function on $[U_t, r_t, \mu_t]$, the change of the latter gives the action of the operator on Π^* , with the same action of merging variables because h can be viewed as giving an ultrametric, and the growth operator is identically zero because h registers only when a distance is in $(0, \infty)$, so there is no change for $t > 0$. We next take the union of the algebras $\widehat{\Pi}^{*,\downarrow}$ and Π^* , and consider the generated algebra $\widehat{\Pi}^*$ (recall Section 3) where we give the action on products of a function from each of the sets, which still is the merging of variables. This gives us the martingale problem that is solved by \mathcal{G} , for $t > 0$.

We have to also discuss general h_0 and \mathcal{U}_0 not in the *closure* of $\widetilde{\mathbb{G}}_{\text{comp}}$, and do so based on the martingale problem. We see from the approximation that in our scaling the connections between *weakly* completely connected components disappear at a larger rate than the latter grow to full frequency of completely connected components. Since the martingale problem implies the duality via the generator criterion, and since we have right-continuous paths, we must have the h -component not in the closure of $\mathbb{G}_{\text{comp}}^{\square}$ (which is dynamically closed) jump to dust, i.e., to $h \equiv 0$, at infinite rate at time 0, as prescribed.

Next, we turn to \mathbb{G}^{\diamond} . The claim is that $(\mathcal{G}_t^{\diamond})_{t \geq 0}$ is itself a Markov process on \mathbb{G}^{\diamond} , specified by a well-posed martingale problem that, however, is not defined on a closed subspace of $\widehat{\mathcal{W}}$, except when restricted to the closure of $\mathbb{G}_{\text{comp}}^{\diamond}$. For every stochastic process $(\mathcal{U}_t^*)_{t \geq 0}$ taking values in \mathbb{U}_1^* , via the embedding ψ of (7.2)–(7.3) restricted *on* $\mathbb{G}_{\text{comp}}^{\diamond}$, we

obtain a stochastic process of graphemes and hence of connection-matrix distributions

$$(7.6) \quad (\nu_t)_{t \geq 0} = (\psi^{-1}(\mathcal{U}_t^*))_{t \geq 0}.$$

Consider the $(\mathcal{L}^*, \Pi^*, \mathcal{G}_0)$ -martingale problem, where Π^* is the set of *connection polynomials* on \mathbb{G}^\square , and (\mathcal{L}^*, Π^*) arises as the *extension of* (\mathcal{L}, Π) *from* \mathbb{U}_1^* *to* \mathbb{G}^\square , with Π the polynomials in \mathbb{U}_1^* . Namely, the coarsening of \mathcal{U} induced by the map Ψ on \mathcal{G}^\diamond gives a \mathbb{U}_1^* -valued process whose operator we can determine from the one of the underlying \mathcal{U} . We see from (7.4) that in our case, for $t > 0$,

$$(7.7) \quad \mathcal{L}^{*,\text{FV}}(\Phi)[\mathcal{U}] = \mathcal{L}^{\text{FV}}([\mathcal{U}]) - \mathcal{L}^{\text{growth}}([\mathcal{U}]), \quad \mathcal{U} \in \mathbb{U}_1,$$

with \mathcal{L}^{FV} the operator for the process $(\mathcal{U}_t)_{t \geq 0}$ on \mathbb{U}_1 . This means that for this functional not the metric but *only the sampling measure* evolves further for $t > 0$. To incorporate $t = 0$, note that the metric jumps instantaneously to 1 for pairs of descendants of a founder as time starts running, and to a larger number otherwise. However, this information we have *only in the associated* \mathcal{U} . Therefore we have to view the process as an entrance law from time 0 in the dust state, where $h = 0$ on \mathbb{G}^\diamond and not in the union of completely connected components. This process is Markovian on \mathbb{G}^\diamond , which completes the first task. This latter follows because the operator does not rely on information about \mathcal{U} , and uniqueness follows from duality.

This settles the second part of part (b) of Proposition 7.1.

(iii) We have a moment duality, linking the operator $\mathcal{L}^{*,\text{FV}}$ to the coalescence operator $\mathcal{L}^{*,\text{coal}}$ on Π^* , when we work with the polynomials on \mathbb{U}_1^* , since the corresponding Fleming-Viot processes \mathcal{U} are in duality including the growth operator. The Fleming-Viot operator and the growth operator separately satisfy the generator criterion and hence, in the backward *and* the forward direction, the growth operator is dropped when passing from r_t to r_t^* , which carries over to the functional that solves the H -duality for positive time. More precisely, for the duality relation of the underlying $\mathbb{U}_1(\bar{\mathbb{U}})$ -valued process $(\mathcal{U}_t)_{t \geq 0}$ of $(\mathcal{G}_t)_{t \geq 0}$, we know that the dual on \mathbb{U}_1 is driven by a *coalescent* on the partitions of \mathbb{N} . Indeed, the right-hand side has the *coalescent with emigration* at rate c into a cemetery (the entrance law from infinity, where the partition elements at time t give the connected components of the grapheme \mathcal{C}_t^* for the coalescent \mathcal{C}_t). We have to combine this with a duality covering the joint law of h and the \mathcal{U} -component, by having in the duality function (6.2) both φ_r and φ_h different from just constants. But this is immediate, since in the action of the dual operator of the second-order part the merging of variables is the same in both parts.

This settles the H -duality.

We want to reformulate the duality in the *framework of* \mathbb{G}^\square to a stronger statement, i.e., we want to translate the duality relation on \mathbb{U}_1^* into a *strong duality relation on* \mathbb{G}^\square . The duality relation for \mathcal{U}^{FV} on \mathbb{U}_1^* says that if we transform the distances of the state \mathcal{U} as in (7.4), then the new state, denoted by \mathcal{U}^\downarrow , satisfies

$$(7.8) \quad \text{LAW} \left[\mathcal{U}_t^{\text{FV},\downarrow} \right] = \text{LAW} [\mathcal{C}_t],$$

where \mathcal{C}_t is the ultrametric probability measure space spanned by the *entrance law of the coalescent*. By the definition of h as functional of \mathcal{U} , respectively, \mathcal{C} in the forward,

respectively, the backward direction, this gives (recall that here h is also an ultrametric)

$$(7.9) \quad \text{LAW} \left[\mathcal{G}_t^{d,c,\theta} \right] = \text{LAW} \left[\mathcal{G}_0 \vdash \mathcal{C}_t^{*,d,c,\theta} \right], \quad t > 0.$$

This dual representation is associated with a duality relation that has duality function H , and is connected to a *generator relation*, which is known and proved on \mathbb{U}_1 in the literature, and above for the h -part, and which guarantees uniqueness of the martingale problem on \mathbb{G}^\square .

From the duality relation w.r.t. H we obtain the strong duality by using the *tightness* properties for the collection of the finite dual processes (see [GPW13]) and the fact that $H(\cdot, \cdot)$ determines states and the law of states. Transforming both sides of the duality relation, we get (7.9), and the above embedding in \mathbb{U}_1^* carries over to $\mathbb{G}_{\text{comp}}^{\langle \rangle}$, and trivially to $\mathbb{G}^{\langle \rangle}$, since we add the immediate jump to the former by modifying (7.9), replacing \mathcal{G}_0 by \mathcal{G}_0 projected on the closure of $\mathbb{G}_{\text{comp}}^\square$ in the right-hand side.

This settles part (d) of Proposition 7.1.

(iv) From (7.9) we conclude that the equilibrium process of $\mathcal{G}^{\langle \rangle}$ arising as functional of \mathcal{G}^\square is described by the equilibrium of a known *measure-valued process*, namely, the Fleming-Viot process with emigration/immigration. Indeed, project the process \mathcal{G}^\square on its measure component $(\mu_t)_{t \geq 0}$, and afterwards to the subalgebra given by the r^* -balls for $r^* = 1$, to get the weight vector of the $2t$ -balls. Take the same operations in the right-hand side of (7.9). This leads to the same right-hand side in (7.9) as in the duality relation for the well studied *measure-valued* Fleming-Viot process with emigration/immigration, allowing us to use [DGV95], where the identification of the equilibrium as well as convergence to equilibrium is given.

This settles part (c) of Proposition 7.1.

(v) The fact that the operator for \mathcal{G} is second order follows from the fact that the first-order terms can be verified as such with [DGP12], whereas the second-order term acting on the h -component and the r -component is of the same form and hence [DGP12] *again* can be used to get Proposition 7.1(e).

This completes the proof of Proposition 7.1.

Let us look in more detail at the martingale problem for the underlying process \mathcal{U} and the facts that are needed as assumptions for Proposition 7.1. The *existence* of a solution to the martingale problem on \mathbb{U}_1 arises via the finite-population approximation, by considering a suitable dynamics with n individuals and letting $n \rightarrow \infty$. Since the dynamics we have given for the vertices and the edges can be viewed as the individual-based Moran process or the Galton-Watson process, both enriched with edges, the convergence properties for graphs fit with the convergence properties for populations. We can therefore use the argument that the genealogies $\mathcal{U}^{d,c,\theta,(n)}, \mathcal{U}^{b,c,\theta,(n)}$ of the finite n -populations converge to a limit, namely, $\mathcal{U}^{d,c,\theta}, \mathcal{U}^{b,c,\theta}$, respectively. This is carried out in [GPW13] for the Fleming-Viot process, in [DG19] for the Galton-Watson process, and in [DGP12] for the model with mutation and selection, leading to a limiting process on \mathbb{U} . For the model with immigration/emigration we have to use the results in [DG19, GSW16, DGP23a, GRG21], but the necessary tightness nevertheless comes out of the same ideas in the proof (see,

in particular, [GSW16, Section 2] and [DG19], where this is exhibited for a particular convergence result, but the technique carries over to our setup).

The *uniqueness* for the martingale problem for \mathcal{U} follows from the duality, as well as the $t \rightarrow \infty$ convergence of that limiting dynamics to *equilibrium*, as verified in the same references for the underlying $\mathcal{U}^{d,c,\theta}, \mathcal{U}^{b,c,\theta}$, where the latter is proved via convergence of the dual, which gives the claim right away (the former is a general fact for standard martingale problems).

In the proof of Theorem 1.14, which concerns the extension to mechanisms including *selection*, \mathcal{L} requires marks. The addition of *mutation* requires not just marks, but marks that are *not* inherited anymore. This is addressed separately in Section 7.3.

Having *only additional selection* works fine, namely, the approximation by finitely many vertices follows for the underlying \mathcal{U} from [GPW13, DG20], respectively, [DG19], which give tightness in path space, in particular, the dust-free condition for the grapheme and convergence to a limit grapheme. Selection is handled via [DGP12]. We give the list of relevant references in Section 7.4.

Results from the literature together with Proposition 7.1 settle the claims in Theorems 1.9, 1.14, 1.20, 1.22, and in some cases of Theorem 1.26, namely, without mutation. \square

7.2 Proof of the two corollaries

In this section we complete the proof of Corollaries 1.23 and 1.24. We have to argue that the weight of μ_t of the completely connected components is given by the measure-valued diffusions with emigration and immigration in [DGV95], respectively, [DG96] with values in $\mathcal{M}_1([0, 1])$, respectively, $\mathcal{M}_{\text{fin}}([0, 1])$. Start at time 0 in the Lebesgue measure. Then the states are atomic measures for every $t > 0$, and for $t \rightarrow \infty$ converge to an equilibrium that can be specified explicitly, as we did in (1.42) and (1.44). How to connect this to our process \mathcal{G} ?

We have to code the weight vector given by the mass of the $2t$ -balls by an atomic measure in $[0, 1]$, respectively, $[0, \infty)$, by associating with the ball a unique number in $[0, 1]$. Take the size-ordered mass vector and assign i.i.d. uniformly distributed numbers from $[0, 1]$ to the corresponding balls. Then take the atomic measure of the observation with the weight under μ_t of the corresponding ball.

Next, we look at the dynamics of this object under the process $(\mathcal{U}_t)_{t \geq 0}$ underlying the process $(\mathcal{G}_t)_{t \geq 0}$. For that we consider the action of the generator on a subclass of test functions of the martingale problems for the process \mathcal{G} , which are given by polynomials $\Phi^{n,\varphi}$, where the function φ identifies the mass content of the different $2t$ -balls. We can choose here the indicators of these balls, which are continuous functions due to ultrametricity. Then we obtain, via formulas (4.20), (4.22), respectively, Section 4.2.2 for the Dawson-Watanabe models, the same expression as in the given references for the measure-valued process with immigration and emigration treated in [DGV95, Chapters 2(a) and 2(b)], documenting the necessary results and their origin in the literature. Similarly, for the $\bar{\mathbb{G}}$ -valued case and the Dawson-Watanabe model use [DG96]: the results from there prove

our claims.

7.3 Adding mutation

If in the extension of our basic processes $\mathcal{G}^{c,d,\theta}$, respectively, $\mathcal{G}^{c,b,\theta}$ (with values in $\mathbb{G}^{\mathbb{I},V}$ beyond the process $\mathcal{G}^{c,d,\theta,s}$ with selection) we add *mutation* to the selection, then in order to get \mathcal{G} from the underlying genealogy process \mathcal{U} we have to change the definition of h as a function of \mathcal{U} , which must now be based on the *types* rather than on the *founder* at time 0 only, and which we might or might not describe via marks. Recall that now two vertices are connected when both carry the same type. In particular, the ultrametric induced by h is no longer related as easily as before to the ultrametric r_t . Typically, *different* founders with *part* of their descendants give rise to the completely connected components. Nonetheless, it is still true that completely connected components are finite or countable unions of balls of $U \times V$ that are open and closed, because the vertices descending from an individual without mutation are connected, since the time of the last mutation. Because they have the same type, this corresponds to a ball of radius twice the time back to the common ancestor and lost mutant. The part suffering mutation in that time span founds a new family, which we have to take out of the ball. Iterate this scheme in the order of the relevant mutations in the tree trunks.

We also need to give attention to the topology of the mark set, to suit our purposes. Before we do so we observe the following. The union of balls taken out leaves a countable union of balls in $U \times V$ that are completely connected. Hence the vertices descending from an individual without mutation without further mutation are a completely connected subset, because they have the same type. This correspond to a ball in U of radius twice the time back to the common ancestor, but all marked with a specific type. Hence we are still able to conclude the approximation via finite models by using the tightness properties that we have for \mathbb{U}_1, \mathbb{U} from the literature. To control the distribution of masses, in particular, their uniform summability, we control the number of ancestors of at least $1 - \varepsilon$ of the mass over bounded time intervals, using backward the dual and forward the diffusion of the frequency of the part of the population descending from a finite set of ancestors. Here we use that the limit of the process projected on V is a well-known measure-valued diffusion, which for $t > 0$ has atomic states and is the limit of its finite counterparts. The \mathbb{U}_1 -valued process projecting on the genealogy has a subfamily decomposition in $2(t - s)$ -balls (disjoint and open and closed) for $s \in [0, t)$, of which at most countably many are charged by the sampling measure (compare with [DG14]). Also this decomposition arises as the limit of the decomposition of the finite counterparts.

We next give more detail on the decomposition above and the necessary topology of K . The claim is that $U \times V$ is decomposed for finite n , and in the limit consists of at most countably many balls in U , each marked with a single type. These are completely connected components, since the descendants of a mutant surviving at the reference time are elements of a ball with radius $2(t - t_{\text{mut}})$, with t_{mut} the time of mutation. The mutants occurring in this ball and surviving have a different type, and give rise to a sub-ball taken out of all the descendants at time t . If we *strengthen the topology* on V to the *discrete* topology, then we obtain a decomposition into open and closed balls, thereby creating the completely connected components.

The remaining parts of the argument are the same as in the models we treated before.

Regarding the equilibrium states, it is known from [DGP12] that the underlying process $(\mathcal{U}_t^{c,d,\theta,m,s})_{t \geq 0}$ has a unique equilibrium. For the process arising when we project on the mark component, this observation becomes a statement (in the case we are treating) about a measure-valued process on the type space V . Therefore, in the case $m = 0$, it is clear that we can transfer the argument to the grapheme process, which is a function of \mathcal{U} . With mutation the argument becomes a bit more subtle as we saw above, since now the function giving \mathcal{G} from \mathcal{U} is at first sight of a somewhat different character, and we have to show that the process run for time $t \rightarrow \infty$ converges to a state obtained by applying the new function applied to \mathcal{U}_∞ as well. But, this follows from the fact that the level sets of h as finite or countable union of balls are continuous.

7.4 Completion of the proof of the main theorems

In this section we complete the proof of Theorems 1.9, 1.14 and 1.18, the duality Theorem 1.20, and the extension Theorem 1.26. We do so by providing more details on the references given in Sections 7.1.1, 7.1.2 in order to justify the *claims* made for the *genealogy process* \mathcal{U} . There we saw how to make the step from $(\mathcal{U}_t)_{t \geq 0}$ to $(\mathcal{G}_t)_{t \geq 0}$. Relevant references and theorems to the properties of genealogy-valued processes \mathcal{U} necessary for a particular claim on \mathcal{G} are listed below for each of our theorems.

A general introduction to the ideas and orientation in the field can be found in [DG20], which gives some orientation to be able to work with the references we suggest below. A further general remark is that the proof in [GSW16, Section 2] contains many steps and arguments that are *easily adapted to different situations*.

Here is the list of references:

- Theorem 1.9: [GSW16] Section 1.3.1, in particular, Theorem 1.12, Theorem 1.17 and the arguments in Section 2.
- Theorem 1.14: [GPW13] Proposition 2.22, [GSW16], Section 2.
- Theorem 1.18: [GSW16] Theorem 1.17, Section 2.
- Theorem 1.20: [GSW16] Theorem 1.17, Section 2.
- Theorem 1.26: [DGP12] Theorem 1, Theorem 3.

8 Proofs for grapheme diffusions with non-completely connected components

In this section we add *insertion and deletion* of edges in order to obtain grapheme diffusions with equilibria whose (*path* connected) components are not necessarily *completely* connected. The key steps are as follows.

We show that for the process denoted by $\mathcal{G}^{(a^+, a^-)}$ (being a solution of the martingale problem) there exists a unique underlying process \mathcal{U} that generates a unique grapheme

process *without* insertion and deletion (as we know already), say $\mathcal{G}^{(0,0)}$. In the model considered in this paper, the edges no longer evolve as a function of the (U, r, μ) , because additional randomness enters.

We analyse the situation by *conditioning* $\mathcal{G}^{(a^+, a^-)}$ on the entire path of the underlying process $([U_t, r_t, \mu_t])_{t \geq 0}$, which evolves *autonomously* as a strong Markov process, and study the evolution $(h_t)_{t \geq 0}$, or better the connection-matrix distribution $(\nu_t)_{t \geq 0}$, which itself does *not* influence the evolution of $(\mathcal{U}_t)_{t \geq 0}$. Note that conditioning on \mathcal{U} is *equivalent* to conditioning on $\mathcal{G}^{(0,0)}$, since the latter is a *function* of \mathcal{U} and also contains \mathcal{U} .

Thus, we have a martingale problem generating the autonomous evolution of the path $(\mathcal{U}_t)_{t \geq 0}$, referred to as the solution of the martingale problem of the condition, i.e., this process exists uniquely with continuous paths and is strong Markov. (The same holds for $\mathcal{G}^{(0,0)}$, as we already know.) Given a complete path of \mathcal{U} , we obtain a conditional martingale problem from (4.30)–(4.31) for $(h_t)_{t \geq 0}$ or $(\nu_t)_{t \geq 0}$, by inserting into the expression the data from the given path \mathcal{U} . This gives us a specific time-inhomogeneous spin-flip system operator in random environment, with the latter given by the condition.

The strategy to treat the martingale problem in Section 8.1, 8.2 is as follows:

- (i) *Given a path* of $(\mathcal{U}_t)_{t \geq 0}$, we let ν_s be the time- s marginal of the law of the functional ν of the $\mathbb{N} \times \mathbb{N}$ connection-matrix distribution given by the evolution $(\nu_s((\mathcal{U}_t)_{t \geq 0}))_{s \geq 0}$, which we characterise by the *conditional* martingale problem derived above (for which we must show that it is well-posed and depends *measurably on* \mathcal{U}).
- (ii) We show that the full martingale problem *implies* that the condition \mathcal{U} evolves according to its own well-posed martingale problem.
- (iii) We show that $(\nu_s^\infty)_{s \geq 0}$ must solve uniquely the *conditional martingale problem* for almost surely all realisations of the path \mathcal{U} . Integration over the condition (recall (i) for measurability) gives the unique solution $\mathcal{G}^{(a^+, a^-)}$ of the martingale problem on \mathbb{G}^\square .
- (iv) With the help of the conditional duality, we can treat the equilibrium as well: we show that the equilibrium is the $t \rightarrow \infty$ limit of $\mathcal{G}_t^{(a^+, a^-)}$ and arises by integrating the equilibrium of $\mathcal{G}_\infty^{(0,0)} = [(U_\infty, r_\infty, \mu_\infty, h_\infty)]$ for the state $[(U_\infty, r_\infty, \mu_\infty, h_\infty)]$.

8.1 Completion of the proof: martingale problem

In this section we complete the proof of Theorem 1.29 concerning the martingale problem part and properties of the solution, i.e., parts (a), (b) and (d). The fact that we deal with a conditional martingale problem requires some care, even though the basic idea is simple. We proceed in three steps to realise the program described above.

Step 1. We have to specify the evolution of the *connection-matrix distribution*, which we do by giving the evolution of the finite-dimensional $(\nu_t^m)_{t \geq 0}$, $m \in \mathbb{N}$, and hence of $(\nu_t)_{t \geq 0}$. This arises when we restrict the action of the operator for a given path of $\mathcal{G}^{(0,0)}$ to functions in Π^* . Note that the *connection-matrix distribution* of the *virtual edges* $(\nu_t^{(0,0)})$ is given. We obtain a process that has continuous path, since $\mathcal{G}^{(0,0)}, \mathcal{U}$ have this property. Consider the operator of the martingale problem as introduced in Section 4.3. Via the corresponding conditional martingale problem (recall (4.30)) this operator determines *uniquely*, for every path of $(\mathcal{U})_{t \geq 0}$, a process $(\nu_t)_{t \geq 0}$ with values in the connection-matrix distributions, since

it is the generator of a time-inhomogeneous spin-flip system with independently acting components on $\{0, 1\}^{\mathbb{N}}$. The latter is, of course, well-posed as the evolution equation of a simple Markov pure-jump process.

Step 2. From the martingale problem of the process $\mathcal{G}^{(a^+, a^-)}$, we identify the martingale problem of the underlying process \mathcal{U} . This in turn determines a unique process $\mathcal{G}^{(0,0)}$, where the insertion/deletion of edges is suppressed. For that purpose, we consider the action of the operator for $\mathcal{G}^{(a^+, a^-)}$ on the test function. This operator depends only on $[U_t, r_t, \mu_t]$ and is, by inspection, immediately identified as the operator acting on functions of the underlying process \mathcal{U} in $\widehat{\Pi}^{*, \downarrow}$. Note that \mathcal{U} agrees completely with the one we obtain as underlying process for $\mathcal{G}^{(0,0)}$. Furthermore, it determines uniquely a solution in \mathbb{U}_1 , and therefore $([U_t, r_t, \mu_t])_{t \geq 0}$ must indeed be a path of the process \mathcal{U} underlying $\mathcal{G}^{(0,0)}$ and $\mathcal{G}^{(a^+, a^-)}$. The point here is that $\widehat{\Pi}^*$ is the algebra generated by $\widehat{\Pi}^{*, \downarrow}$ and $\widehat{\Pi}^*$, each of which belongs to one of the components of the grapheme, $[U_t, r_t, \mu_t]$, respectively, $[h]$ and, due to the form of the operator for the two components, the operator is lifted immediately from the one of the \mathbb{G}^{\square} -valued martingale problem.

Step 3. Here we come to the modified dual representation. We have a dual representation of the underlying process $\mathcal{G}^{(0,v)}$ or \mathcal{U} and of the conditional martingale problem. This means that the process $\mathcal{G}^{(0,0)}$ or \mathcal{U} , and the conditional process giving h^* based on $h^{(0,0)}$, are uniquely determined and are continuous in the initial condition. Therefore there is a unique process, arising from the recipe above, that solves the original martingale problem. To see that this is the *unique* solution of the latter, it suffices to have a duality for the original martingale problem. But this follows from the form of the operator, and from the duality relations for $\mathcal{G}^{(0,0)}$ and the additional generator term associated with the spin-flip system, dual to the *reversed* spin-flip system.

8.2 Completion of the proof: equilibrium

What remains to be shown is that the equilibrium exists and is as described in part (c) of the theorem. However, this follows from the dual representation, which guarantees existence of the limit as $t \rightarrow \infty$. The dual of $\mathcal{G}_t^{(0,0)}$ converges as $t \rightarrow \infty$ to an equilibrium that we identified via the dual, and for which we explicitly obtained the forward equilibrium state of the functional ν . Subsequently, we observe that the pruning of the edges yields the frequencies of edges as the current law of the spin-flip system on $\{0, 1\}^{\mathbb{N}}$, which approaches the limit state as $t \rightarrow \infty$ given by $(\frac{a^+}{a^+ + a^-}, \frac{a^-}{a^+ + a^-})^{\otimes \mathbb{N}}$.

The equilibrium of the conditional process is independent of the condition. Therefore the convergence of $\mathcal{G}^{(0,0)}$ to an equilibrium, which is independent of the edge process, indeed means convergence to the claimed equilibrium, which is the product of the law of an \mathbb{N} -sample $h^{(0,0)}$ of $\mathcal{G}_{\infty}^{(0,0)}$ and of the spin-flip system h_{∞}^* on $\{0, 1\}^{\mathbb{N}}$. The equilibrium v_{∞} is obtained as the law of (denote with \times the point-wise multiplication of two function on the same space \mathbb{N})

$$(8.1) \quad h^{(0,0)} \times h^*.$$

References

- [AdHR21] S. Athreya, F. den Hollander, and A. Röllin. Graphon-valued stochastic processes from population genetics. *Ann. Appl. Probab.*, 31(4):1724–1745, 2021.
- [BCL⁺08] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi. Convergent graph sequences I: Subgraph frequencies, metric properties, and testing. *Adv. Math.*, 219:1801–1851, 2008.
- [BCL⁺12] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi. Convergent sequences of dense graphs II: Multiway cuts and statistical physics. *Ann. Math.*, 176:151–219, 2012.
- [BdHM22] P. Braunsteins, F. den Hollander, and M. Mandjes. Graphon-valued processes with vertex-level fluctuations. Available at arXiv:2209.01544[math.PR] (to appear in *Stoch. Proc. Appl.*), 2022.
- [BdHM23] P. Braunsteins, F. den Hollander, and M. Mandjes. A sample-path large deviation principle for dynamic Erdős-Rényi random graphs. *Ann. Appl. Probab.*, 33:3278–3320, 2023.
- [Cha15] S. Chatterjee. *Large Deviations for Random Graphs*. École d’Été de Probabilités de Saint-Flour XLV. Springer, 2015.
- [CK18] J. Cerný and A. Klimovsky. Markovian dynamics of exchangeable arrays. In *Genealogies of Interacting Particle Systems, Lecture Notes Series 38, Institute for Mathematical Sciences, National University of Singapore*. World Scientific, Singapore, 2018.
- [Cra16] H. Crane. Dynamic random networks and their graph limits. *Ann. Appl. Probab.*, 26:691–721, 2016.
- [Cra17] H. Crane. Exchangeable graph-valued Feller processes. *Probab. Theory Relat. Fields*, 168:849–899, 2017.
- [CV11] S. Chatterjee and S.R.S. Varadhan. The large deviation principle for the erdős-rényi random graph. *European J. Comb.*, 32:1000–1017, 2011.
- [Daw93] D.A. Dawson. *Measure-valued Markov processes*. Springer, 1993.
- [DG96] D.A. Dawson and A. Greven. Multiple space-time scale analysis for interacting branching models. *Electron. J. Probab.*, 1(14):1–84, 1996.
- [DG14] D. A. Dawson and A. Greven. *Spatial Fleming-Viot models with selection and mutation*, volume 2092 of *Lecture Notes in Math*. Springer Verlag, 2014.
- [DG19] A. Depperschmidt and A. Greven. Genealogy-valued Feller diffusion. Available at arXiv:1904.02044[math.PR], 2019.

- [DG20] A. Depperschmidt and A. Greven. Stochastic evolution of genealogies of spatial populations: State description, characterization of dynamics and properties, pages 39–79. In: *Genealogies of Interacting Particle systems*, World Scientific, Singapore, 2020.
- [DGKR15] P. Diao, D. Guillet, A. Khare, and B. Rajaratnam. Differential calculus on graphon space. *J. Combin. Theory Ser. A*, 133:183–227, 2015.
- [DGP11] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Marked metric measure spaces. *Electron. Commun. Probab.*, 16:174–188, 2011.
- [DGP12] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Tree-valued Fleming-Viot dynamics with mutation and selection. *Ann. Appl. Probab.*, 22(6):2560–2615, 2012.
- [DGP23a] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Duality and the well-posedness of a martingale problem. Available at arXiv:1904.01564[math.PR], 2023.
- [DGP23b] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Evolving genealogies of interacting Fleming-Viot processes including recombination. preprint, 2023.
- [DGV95] D.A. Dawson, A. Greven, and J. Vaillancourt. Equilibria and quasi-equilibria for infinite collections of interacting Fleming-Viot processes. *Trans. Amer. Math. Soc.*, 347:2277–2360, 1995.
- [EK86] S.N. Ethier and T. Kurtz. *Markov Processes. Characterization and Convergence*. John Wiley, New York, 1986. Paperback reprint in 2005.
- [GdH23] A. Greven and F. den Hollander. Spatial populations with seed-bank: finite-systems scheme. *Electron. J. Probab.*, 28:1–66, paper no. 84, 2023.
- [GdHKW23] A. Greven, F. den Hollander, A. Klimovsky, and A. Winter. The grapheme-valued Wright-Fisher diffusion with mutation. Available at arXiv:2308.11598[math.PR] (to appear in *Stoch. Proc. Appl.*), 2023.
- [GdHO21] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: renormalisation on the hierarchical group. to appear in *Memoirs Amer. Math. Soc.*, 2021.
- [GdHO22] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: well-posedness, duality and equilibrium. *Electron. J. Probab.*, 27:1–88, paper no. 18, 2022.
- [GKW23] A. Greven, A. Klimovsky, and A. Winter. Ancestral web and measure-valued path representations. 2023.
- [GPW09] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (Λ -coalescent measure tree). *Probab. Theory Relat. Fields*, 145(1):285–322, 2009.

- [GPW13] A. Greven, P. Pfaffelhuber, and A. Winter. Tree-valued resampling dynamics: martingale problems and applications. *Probab. Theory Relat. Fields*, 155(3–4):789–838, 2013.
- [GRG21] A. Greven, T. Rippl, and P. Glöde. Branching processes – a general concept. *ALEA*, 18:635–706, 2021.
- [GSW16] A. Greven, R. Sun, and A. Winter. Continuum space limit of the genealogies of interacting Fleming-Viot processes on \mathbb{Z} . *Electron. J. Probab.*, 21:Paper No. 58, 64, 2016.
- [HN12] H. Hatami and S. Norine. The entropy of random-free graphons and properties. Available at arXiv:1205.3529[math.PR], 2012.
- [Jan13] S. Janson. Graphons, cut norm and distance, couplings and rearrangements. *NYJM Monographs*, 2013.
- [LS07] L. Lovasz and B. Szegedy. Szemerédi’s lemma for the analyst. *Geometric and Functional Analysis*, (1)(17):252–270, 2007.
- [LW21] W. Löhr and A. Winter. Spaces of algebraic measure trees and triangulations of the circle. *Bull. Soc. Math. France*, 149(1):55–117, 2021.
- [Pet09] L.A. Petrov. A two-parameter family of infinite-dimensional diffusions on the kingman simplex. *Funct. Anal. Appl.*, 43:279–296, 2009.
- [VR15] V. Veitch and D.M. Roy. The class of random graphs arising from exchangeable random measures. Available at arXiv:1512.03099[math.PR], 2015.