**A sample-path large deviation principle
for dynamic Erdös Renyi random graphs**

P. Braunsteins, F. den Hollander, M. Mandjes

A SAMPLE-PATH LARGE DEVIATION PRINCIPLE
FOR DYNAMIC ERDŐS-RÉNYI RANDOM GRAPHS

PETER BRAUNSTEINS, FRANK DEN HOLLANDER, AND MICHEL MANDJES

ABSTRACT. We consider a dynamic Erdős-Rényi random graph (ERRG) on n vertices in which each edge switches on at rate λ and switches off at rate μ, independently of other edges. The focus is on the analysis of the evolution of the associated empirical graphon in the limit as $n \to \infty$. Our main result is a large deviation principle (LDP) for the sample path of the empirical graphon observed until a fixed time horizon. The rate is $\binom{n}{2}$, the rate function is a specific action integral on the space of graphon trajectories. We apply the LDP to identify (i) the most likely path that starting from a constant graphon creates a graphon with an atypically large density of d-regular subgraphs, and (ii) the mostly likely path between two given graphons. It turns out that bifurcations may occur in the solutions of associated variational problems.

1. Introduction and main results

Section 1.1 provides motivation and background, Section 1.2 introduces graphs and graphons, Section 1.3 recalls the LDP for the inhomogeneous ERRG, Section 1.4 defines a switching dynamics for the ERRG, Section 1.5 states the sample-path LDP for the latter, while Section 1.6 offers a brief discussion and announces two applications.

1.1. **Motivation and background.** Graphons arise as limits of dense graphs, i.e., graphs in which the number of edges is of the order of the square of the number of vertices. The theory of graphons – developed in [17], [18], [2], [3] – aims to capture the limiting behaviour of large dense graphs in terms of their subgraph densities (see [16] for an overview). Both typical and atypical behaviour of random graphs and their associated graphons have been analysed, including LDPs for homogeneous and inhomogeneous Erdős-Rényi random graphs [7], [10].

Most of the theory focusses on *static* random graphons, although recently some attempts have been made to include *dynamic* random graphons [22], [8], [9], [5], [1]. The goal of the present paper is to generalise the LDP in [7] to a sample-path LDP for a dynamic random graph in which the edges switch on and off in a random fashion. The equilibrium of the dynamics coincides with the setup of [7], so that our sample-path LDP is a true dynamic version of the static LDP derived in [7]. The corresponding large deviation rate function turns out to be an action integral. We consider two applications that look at optimal paths for graphons that realise a prescribed large deviation. We find that bifurcations may occur in the solutions of the associated variational problems.

1.2. **Graphs and graphons.** There is a natural way to embed a simple graph on n vertices in a space of functions called *graphons*. Let $\mathscr{W}$ be the space of functions $h\colon [0,1]^2 \to [0,1]$ such that $h(x,y) = h(y,x)$ for all $(x,y) \in [0,1]^2$, formed after taking the quotient with respect to the equivalence relation of almost everywhere equality. A finite simple graph G on n vertices can be represented as a graphon $h^G \in \mathscr{W}$ by setting

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

This object is referred to as an *empirical graphon* and has a block structure (see Figure 1). The space of graphons $\mathscr{W}$ is endowed with the *cut distance*

$$d_\square(h_1, h_2) := \sup_{S,T \subseteq [0,1]} \left| \int_{S \times T} \mathrm{d}x\,\mathrm{d}y\,[h_1(x,y) - h_2(x,y)] \right|, \quad h_1, h_2 \in \mathscr{W}. \tag{1.2}$$

The space $(\mathscr{W}, d_\square)$ is not compact.

On $\mathscr{W}$ there is a natural equivalence relation, referred to as '$\sim$'. More precisely, with $\mathscr{M}$ denoting the set of measure-preserving bijections $\sigma\colon [0,1] \to [0,1]$, we write $h_1(x,y) \sim h_2(x,y)$ when there exists a $\sigma \in \mathscr{M}$ such that $h_1(x,y) = h_2(\sigma(x), \sigma(y))$ for all $(x,y) \in [0,1]^2$. This equivalence relation induces the quotient space $(\widetilde{\mathscr{W}}, \delta_\square)$, where $\delta_\square$ is the *cut metric* defined by

$$\delta_\square(\tilde{h}_1, \tilde{h}_2) := \inf_{\sigma_1, \sigma_2 \in \mathscr{M}} d_\square(h_1^{\sigma_1}, h_2^{\sigma_2}), \quad \tilde{h}_1, \tilde{h}_2 \in \widetilde{\mathscr{W}}. \tag{1.3}$$
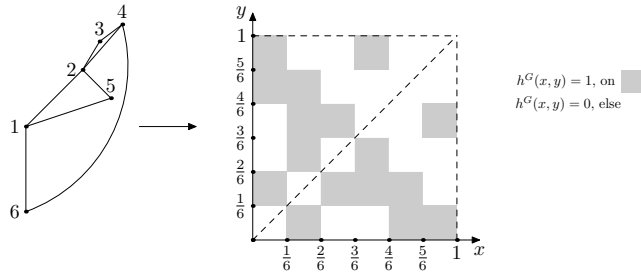
Figure 1. An example of an empirical graphon.

The space $(\tilde{\mathscr{W}}, \delta_\square)$ *is* compact [17, Lemma 8].

Suppose that H is a simple graph on k vertices. The *homomorphism density* of H in $G \supseteq H$ is defined as

$$t(H, G) = t(H, h^G) := \int_{[0,1]^k} \mathrm{d}x_1 \ldots \mathrm{d}x_k \prod_{\{i,j\} \in E(H)} h^G(x_i, x_j), \tag{1.4}$$

where $E(H)$ is the set of edges of H and $k = |E(H)|$. The homomorphism densities are continuous with respect to the cut metric [6, Proposition 3.2].

1.3. **LDP for the inhomogeneous ERRG.** Let $r \in \mathscr{W}$ be a *reference graphon* satisfying

$$\exists\, \eta > 0\colon \qquad \eta \leqslant r(x,y) \leqslant 1 - \eta \quad \forall\, x, y \in [0,1]^2. \tag{1.5}$$

Fix $n \in \mathbb{N}$ and consider the random graph G_n with vertex set $[n] = \{1, \ldots, n\}$ where the pair of vertices $i, j \in [n]$, $i \neq j$, is connected by an edge with probability $r(\frac{i}{n}, \frac{j}{n})$, independently of other pairs of vertices. Write $\mathbb{P}_n$ to denote the law of G_n. Use the same symbol for the law on $\mathscr{W}$ induced by the map that associates with the graph G_n its graphon h^{G_n}. Write $\tilde{\mathbb{P}}_n$ to denote the law of $\tilde{h}^{G_n}$, the equivalence class associated with h^{G_n}.

The following LDP has been proven in [10] and is an extension of the celebrated LDP for the homogeneous ERRG derived in [7].

Theorem 1.1. [LDP for inhomogeneous ERRG] *Subject to* (1.5), *the sequence of probability measures* $(\tilde{\mathbb{P}}_n)_{n \in \mathbb{N}}$ *satisfies the LDP on* $(\tilde{\mathscr{W}}, \delta_\square)$ *with rate* $\binom{n}{2}$, *i.e.,*

$$\begin{aligned}
\limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \tilde{\mathbb{P}}_n(\mathcal{C}) &\leqslant -\inf_{\tilde{h} \in \mathcal{C}} J_r^*(\tilde{h}) \quad \forall\, \mathcal{C} \subseteq \tilde{\mathscr{W}} \text{ closed,} \\
\liminf_{n \to \infty} \frac{1}{\binom{n}{2}} \log \tilde{\mathbb{P}}_n(\mathcal{O}) &\geqslant -\inf_{\tilde{h} \in \mathcal{O}} J_r^*(\tilde{h}) \quad \forall\, \mathcal{O} \subseteq \tilde{\mathscr{W}} \text{ open.}
\end{aligned} \tag{1.6}$$

Here the rate function $J_r^* \colon \tilde{\mathscr{W}} \to \mathbb{R}$ *is the lower semi-continuous envelope of the function* J_r *given by*

$$J_r(\tilde{h}) = \inf_{\sigma \in \mathscr{M}} I_r(h^\sigma), \tag{1.7}$$

where h is any representative of $\tilde{h}$ and

$$I_r(h) := \int_{[0,1]^2} \mathrm{d}x\, \mathrm{d}y\, \mathcal{R}\big(h(x,y) \mid r(x,y)\big), \quad h \in \mathscr{W}, \tag{1.8}$$

with

$$\mathcal{R}(a \mid b) := a \log \tfrac{a}{b} + (1 - a) \log \tfrac{1-a}{1-b} \tag{1.9}$$

the relative entropy of two Bernoulli distributions with success probabilities $a \in [0, 1]$, $b \in (0, 1)$ (with the convention $0 \log 0 = 0$).

It is clear that J_r^* is a good rate function, i.e., $J_r^* \not\equiv \infty$ and J_r^* has compact level sets. It was shown in [21] that (1.5) can be weakened: Theorem 1.1 holds when $0 < r < 1$ almost everywhere under the integrability conditions $\log r, \log(1 - r) \in L^1([0, 1]^2)$. Moreover, it was shown in [21] that J_r is lower semi-continuous on $\widetilde{\mathscr{W}}$, and so $J_r^* = J_r$. In [4] the case where r is a block graphon is considered, which is allowed to take the value 0 or 1 on some blocks.

1.4. Dynamics for the inhomogeneous ERRG.

We now allow the edges to alternate between being active and inactive, thereby creating a dynamic version of the setup studied in [7]. Let $\mathbb{G}_n$ be the set of simple graphs with n vertices. Fix a time horizon $T \in (0, \infty)$. Consider a continuous-time Markov process $\{G_n(t)\}_{t \in [0,T]}$ with state space $\mathbb{G}_n$, starting from a given graph $G_n(0)$. The edges in $G_n(t)$ update *independently* after exponentially distributed times, according to the following rules:

○ an inactive edge becomes active at rate $\lambda \in (0, \infty)$;
○ an active edge becomes inactive at rate $\mu \in (0, \infty)$.

Throughout the paper, the transition rates $\lambda, \mu \in (0, \infty)$ are held fixed. Let $p_{01,t}$ ($p_{11,t}$) denote the probability that an initially inactive (active) edge is *active* at time t. Then

$$p_{01,t} = \frac{\lambda - \lambda \, \mathrm{e}^{-t(\lambda+\mu)}}{\lambda + \mu}, \qquad p_{11,t} = \frac{\lambda + \mu \, \mathrm{e}^{-t(\lambda+\mu)}}{\lambda + \mu}. \tag{1.10}$$

We can represent $\{G_n(t)\}_{t \in [0,T]}$ as a graphon-valued process. Abbreviate

$$f_{n,t} := h^{G_n(t)}, \qquad f_n := (f_{n,t})_{t \in [0,T]}. \tag{1.11}$$

Let $\mathscr{W} \times [0, T]$ be the set of $\mathscr{W}$-valued paths on the time interval $[0, T]$. On the space $(\mathscr{W}, d_\square)$, we can define the Skorohod topology on $\mathscr{W}$-valued paths in the usual way, namely,

$$D = D([0, T], \mathscr{W}) = \text{set of càdlàg paths in } \mathscr{W}, \tag{1.12}$$

and equip D with a metric that induces the Skorohod topology. Define

$$\mu_n(B) := \mathbb{P}_n(f_n \in B), \qquad \tilde{\mu}_n(B) := \mathbb{P}_n(\tilde{f}_n \in B), \tag{1.13}$$

for B in the Borel sigma-algebra induced by the metric.

Note that the initial graphon $f_{n,0}$ effectively plays the role of the reference graphon r in the static setting of an inhomogeneous ERRG treated in [10].

1.5. Main theorem: sample-path LDP.

In order to state our main theorem (the sample-path LDP in Theorem 1.4 below), we first state a few simpler LDPs.

1.5.1. *LDP for local edge density.* Fix $t \in [0, T]$, $(x, y) \in [0, 1]^2 \backslash D^*$, with D^* the diagonal, and $\Delta > 0$ small enough so that $[x, x + \Delta) \times [y, y + \Delta) \cap D^* = \varnothing$. Let

$$\bar{u}_{t,n} = \bar{u}_{t,n}(x, y) = \frac{1}{\Delta^2} \int_{[x,x+\Delta) \times [y,y+\Delta)} d\bar{x} \, d\bar{y} \, f_{n,t}(\bar{x}, \bar{y})$$

$$= \frac{1}{(n\Delta)^2} \sum_{(i,j) \in n[x,x+\Delta) \times n[y,y+\Delta)} 1_{\{i \text{ and } j \text{ are connected in } G_n(t)\}} \qquad (1.14)$$

denote the proportion of active edges in $[x, x + \Delta) \times [y, y + \Delta)$ at time t (for simplicity we pretend that $nx, ny, n\Delta$ are integer). Fixing an initial proportion of active edges $\bar{u}_{0,n} = \bar{u}$ that is independent of n, we see that the moment generating function of $\bar{u}_{t,n}$, defined by $M_{\bar{u}_{t,n}}(s) := \mathbb{E}_n[e^{s\bar{u}_{t,n}}]$, $s \in \mathbb{R}$, equals

$$M_{\bar{u}_{t,n}}(s) = [(1 - p_{11,t}) + e^{sN^{-1}} p_{11,t}]^{N\bar{u}} [(1 - p_{01,t}) + e^{sN^{-1}} p_{01,t}]^{N(1-\bar{u})} \qquad (1.15)$$

with $N := (n\Delta)^2$, the total number of edges in $[x, x + \Delta) \times [y, y + \Delta)$. Here, N^{-1} is the contribution to $\bar{u}_{t,n}$ from a single active edge. Hence

$$\lim_{n \to \infty} N^{-1} \log M_{\bar{u}_{t,n}}(vN) = J_{t,v}(\bar{u}), \qquad v \in \mathbb{R}, \qquad (1.16)$$

with

$$J_{t,v}(\bar{u}) := \bar{u} \log[(1 - p_{11,t}) + e^v p_{11,t}] + (1 - \bar{u}) \log[(1 - p_{01,t}) + e^v p_{01,t}]. \qquad (1.17)$$

Then, by the Gärtner-Ellis theorem [14, Chapter V], the sequence $(\bar{u}_{t,n})_{n \in \mathbb{N}}$ satisfies the LDP on $\mathbb{R}$ with rate N and with good rate function

$$I_{1,t}(\bar{u}, w) := \sup_{v \in \mathbb{R}} [vw - J_{t,v}(\bar{u})], \qquad w \in \mathbb{R}, \qquad (1.18)$$

which is the Legendre transform of (1.17). We use the indices $1, t$ to indicate that (1.18) is the rate function for 1 time lapse of length t.

1.5.2. *Two-point LDP.* If we extend the domain of $I_{1,t}(\bar{u}, w)$ in (1.18) to $\mathscr{W}^2$ by putting

$$I_{1,t}(u, h) := \int_{[0,1]^2} dx \, dy \, I_{1,t}(u(x, y), h(x, y)), \qquad (1.19)$$

then we obtain a candidate rate function for a *two-point* LDP. However, $I_{1,t}$ is not necessarily well defined on $\tilde{\mathscr{W}}^2$ because for $u_1 \sim u_2$ and $h_1 \sim h_2$ it may be that $I_{1,t}(u_1, h_1) \neq I_{1,t}(u_2, h_2)$. To define a valid candidate rate function, we put

$$\tilde{I}_{1,t}(\tilde{u}, \tilde{h}) := \inf_{\sigma_1, \sigma_2 \in \mathscr{M}} I_{1,t}(u^{\sigma_1}, h^{\sigma_2}) = \inf_{\sigma_2 \in \mathscr{M}} I_{1,t}(u, h^{\sigma_2}) = \inf_{\sigma_1 \in \mathscr{M}} I_{1,t}(u^{\sigma_1}, h), \qquad (1.20)$$

noting that $I_{1,t}(u^{\sigma_1}, h^{\sigma_2}) = I_{1,t}(u, h^{\sigma_2 \circ \sigma_1^{-1}}) = I_{1,t}(u^{\sigma_1 \circ \sigma_2^{-1}}, h)$ and $\sigma_2 \circ \sigma_1^{-1}, \sigma_1 \circ \sigma_2^{-1} \in \mathscr{M}$.

Define

$$\mu_{n,t}(B) := \mathbb{P}_n(f_{n,t} \in B), \qquad \tilde{\mu}_{n,t}(B) := \mathbb{P}_n(\tilde{f}_{n,t} \in B) \qquad (1.21)$$

for B in the Borel sigma-algebra.

Theorem 1.2. [Two-point LDP] *Suppose that* $\lim_{n \to \infty} \delta_\square(\tilde{f}_{n,0}, \tilde{u}) = 0$ *for some* $\tilde{u} \in \tilde{\mathscr{W}}$. *Then the sequence of probability measures* $(\tilde{\mu}_{n,T})_{n \in \mathbb{N}}$ *satisfies the LDP on* $\tilde{\mathscr{W}}$ *with rate* $\binom{n}{2}$ *and with good rate function* $\tilde{I}_{1,T}(\tilde{u}, \tilde{h})$.

1.5.3. *Multi-point LDP.* The *multi-point* candidate rate function follows from the two-point candidate rate function by iteration. Let $\mathscr{J}$ denote the collection of all ordered finite subsets of $[0, T]$, i.e., $j \in \mathscr{J}$ if $j = (t_0, t_1, \ldots, t_k)$ with $0 = t_0 < t_1 < \cdots < t_k = T$ for some $k = |j| \in \mathbb{N}$. For $\tilde{g} \in \mathscr{W} \times [0, T]$ and $j \in \mathscr{J}$, let

$$p_j(\tilde{g}) = (\tilde{g}_{t_0}, \tilde{g}_{t_1}, \ldots, \tilde{g}_{t_{|j|}}) \in \tilde{\mathscr{W}}^{|j|+1}. \tag{1.22}$$

Theorem 1.3. [Multi-point LDP] *Suppose that* $\lim_{n \to \infty} \delta_\square(\tilde{f}_{n,0}, \tilde{u}) = 0$ *for some* $\tilde{u} \in \tilde{\mathscr{W}}$. *Then, for every* $j \in \mathscr{J}$, *the sequence of probability measures* $(\tilde{\mu}_n \circ p_j^{-1})_{n \in \mathbb{N}}$ *satisfies the LDP on* $\tilde{\mathscr{W}}^{|j|+1}$ *with rate* $\binom{n}{2}$ *and with good rate function*

$$\tilde{I}_j\left((\tilde{h}_i)_{i=0}^{|j|}\right) := \sum_{i=1}^{|j|} \tilde{I}_{1, t_i - t_{i-1}}(\tilde{h}_{i-1}, \tilde{h}_i) \tag{1.23}$$

with $\tilde{h}_0 = \tilde{u}$.

1.5.4. *Sample-path LDP.* Let $\mathcal{AC}$ denote the set of functions $h \in \mathscr{W} \times [0, T]$ such that $t \mapsto h_t(x, y)$ is absolutely continuous for almost all $(x, y) \in [0, 1]^2$. For $h \in \mathcal{AC}$, put

$$h'_t(x, y) = \left.\frac{\partial h_s(x, y)}{\partial s}\right|_{s=t}. \tag{1.24}$$

To write down a candidate rate function for the sample-path LDP, fix $\Delta t > 0$ such that $T/\Delta t \in \mathbb{N}$. We will show that

$$I_j\left((h_{i\Delta})_{i=0}^{T/\Delta t}\right) = \sum_{i=1}^{T/\Delta t} I_{1, \Delta t}(h_{(i-1)\Delta t}, h_{i\Delta t}) \to I(h), \qquad \Delta t \downarrow 0, \tag{1.25}$$

with

$$I(h) := \begin{cases} \frac{1}{2} \int_0^T \mathrm{d}t \int_{[0,1]^2} \mathrm{d}x\, \mathrm{d}y\, \mathcal{L}(h_t(x, y), h'_t(x, y)), & h \in \mathcal{AC}, \\ \infty, & h \notin \mathcal{AC}, \end{cases} \tag{1.26}$$

where

$$\mathcal{L}(a, b) = \sup_{v \in \mathbb{R}} \left[vb - \lambda(\mathrm{e}^v - 1)(1 - a) - \mu(\mathrm{e}^{-v} - 1)a\right], \qquad a \in [0, 1], b \in \mathbb{R}. \tag{1.27}$$

As before, I in (1.26) is not necessarily well defined on $\mathscr{W} \times [0, T]$, and therefore is not a valid candidate function. For this reason we extend the equivalence relation $\sim$ on $\mathscr{W}$ to the equivalence relation $\sim$ on $\mathscr{W} \times [0, T]$ obtained by defining, for every $h_1, h_2 \in \mathscr{W} \times [0, T]$,

$$h_1 \sim h_2 \quad \text{if and only if} \quad (h_1)_t \sim (h_2)_t \quad \forall\, t \in [0, T], \tag{1.28}$$

and writing $\tilde{h}$ to denote the equivalence class of $h \in \mathscr{W} \times [0, T]$.

Theorem 1.4. [Sample-path LDP] *Suppose that* $\lim_{n \to \infty} \delta_\square(\tilde{f}_{n,0}, \tilde{u}) = 0$ *for some* $\tilde{u} \in \tilde{\mathscr{W}}$. *Then the sequence* $(\tilde{\mu}_n)_{n \in \mathbb{N}}$ *satisfies the LDP on* $\tilde{\mathscr{W}} \times [0, T]$ *with rate* $\binom{n}{2}$ *and with good rate function*

$$\tilde{I}(\tilde{h}) := \inf_{\substack{h \in \mathscr{W} \times [0,T]: \\ h \sim \tilde{h}}} I(h), \qquad \tilde{h} \in \tilde{\mathscr{W}} \times [0, T], \tag{1.29}$$

with $\tilde{h}_0 = \tilde{u}$.

1.6. **Discussion.** Theorems 1.3 and 1.4 are LDPs for the dynamic inhomogeneous Erdős-Rényi random graph with independent edge switches. The fact that the rate is $\binom{n}{2}$, the total number of edges, is natural because a positive fraction of the states of the edges must switch in order to produce a change in the graphon. The fact that the rate function in the sample-path LDP is an *action integral* is also natural, because what matters is both the value of the graphon and the gradient of the graphon integrated along the sample path (due to the exponentiality of the underlying switching mechanism).

Even though the shape of the rate function in (1.29) can be guessed through standard large deviations arguments, the proof of the LDP requires various non-standard steps. Specifically, we are facing the following challenges:

- In Theorem 1.1 the edge probabilities are determined by a single (typically smooth) reference graphon, whereas in Theorem 1.3 the edge probabilities at time T are determined by a sequence of (inherently rough) empirical graphons. This adds a layer of complexity to the proof, and requires a series of approximations that are technically demanding.
- Theorem 1.1 is an LDP on the quotient space $\widetilde{\mathscr{W}}$, whereas Theorem 1.4 is an LDP on the quotient space of *paths* $\widetilde{\mathscr{W}} \times [0, T]$. This leads to various complications in the proofs, as is also evident from the variational problems that arise when we apply the LDP. While the LDP for the static inhomogeneous ERRG is covered by [6], [7] and the sample-path LDP for collections of switching processes is studied in e.g. [23], the dynamic inhomogeneous ERRG considered in the present paper faces the hurdles encountered in both these works.

Several extensions may be thought of. In order to achieve a space-inhomogeneous dynamics, we may replace λ, μ by graphons $\lambda(x, y), \mu(x, y)$, for $(x, y) \in [0, 1]^2$, that are bounded away from 0 and 1, and let the edge between i and j switch on at rate $\lambda(\frac{i}{n}, \frac{j}{n})$ and switch off at rate $\mu(\frac{i}{n}, \frac{j}{n})$. In addition, these graphons may vary over time, in order to capture a time-inhomogeneous dynamics. Both extensions are straightforward and are therefore not addressed in the present paper. A challenging extension would be to consider dynamics where the switches of the edges are dependent (cf. the setup analysed in [1]).

The two applications to be described in Section 2 show that the dynamics is a source of *new phenomena*. The fact that dynamics brings extra richness is no surprise: the area of *interacting particle systems* is a playground with a long history [15].

1.7. **Outline.** Section 2 describes two applications of Theorems 1.3 and 1.4, formulated in Theorems 2.1 and 2.4 below. Section 3 contains the proof of Theorem 1.2, Section 4 the proof of Theorems 1.3 and 1.4, and Section 5 the proof of Theorems 2.1 and 2.4. The applications show that the dynamics introduces interesting bifurcation phenomena.

2. Applications

Section 2.1 identifies the most likely path the process takes if it starts from a constant graphon and ends as a graphon with an atypically large density of d-regular graphs. Section 2.2 identifies the mostly likely path between two given graphons. In these applications, the LDPs presented in Section 1 come to life.

2.1. **Application 1.** Suppose that the initial graphon u is constant, i.e., $u \equiv c$ for some $c \in [0,1]$. Condition on the event that at time $T > 0$ the density of d-regular graphs in $G_n(T)$ is at least r^d, where r corresponds to an atypically large edge density compared to u, i.e.,

$$r > c\, p_{11,T} + (1 - c)\, p_{01,T}. \tag{2.1}$$

A natural question is the following. Is the graph $G_n(T)$ conditional on this event close in the cut distance to a typical outcome of an ERRG with edge probability r? Phrased differently, are the additional d-regular graphs formed by extra edges (i) sprinkled uniformly, or (ii) arranged in some special structure?

2.1.1. *Phase transition.* The next theorem, which can be thought of as the dynamic equivalent of [19, Thm. 1.1], answers the above questions when the initial graphon is constant.

Theorem 2.1. [Phase transition] *Fix a constant initial graphon u. Let H be a d-regular graph for some $d \in \mathbb{N}\backslash\{1\}$, and $e(H)$ the number of edges in the graph H. Suppose that $\delta_\square(\tilde{f}_{n,0}, \tilde{u}) \to 0$ and that r satisfies (2.1).*

(i) *If the point $(r^d, I_{1,T}(u,r))$ lies on the convex minorant of $x \mapsto I_{1,T}(u, x^{1/d})$, then*

$$\lim_{n\to\infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}\left(t(H, f_{n,T}) \geqslant r^{e(H)} \right) = -I_{1,T}(u,r), \tag{2.2}$$

and for every $\varepsilon > 0$ there exists a $C > 0$ such that

$$\mathbb{P}\left(\delta_\square(f_{n,T}, r) < \varepsilon \;\middle|\; t(H, f_{n,T}) \geqslant r^{e(H)} \right) \geqslant 1 - e^{-Cn^2}, \quad n \in \mathbb{N}. \tag{2.3}$$

(ii) *If the point $(r^d, I_{1,T}(u,r))$ does not lie on the convex minorant of $x \mapsto I_{1,T}(u, x^{1/d})$, then*

$$\lim_{n\to\infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}\left(t(H, f_{n,T}) \geqslant r^{e(H)} \right) > -I_{1,T}(u,r), \tag{2.4}$$

and there exist $\varepsilon, C > 0$ such that

$$\mathbb{P}\left(\inf_{s\in[0,1]} \delta_\square(f_{n,T}, s) > \varepsilon \;\middle|\; t(H, f_{n,T}) \geqslant r^{e(H)} \right) \geqslant 1 - e^{-Cn^2}, \quad n \in \mathbb{N}. \tag{2.5}$$

In (2.3) and (2.5) the $\delta_\square$-distance is towards the constant graphons r and s, respectively. We say that $G_n(T)$ is in the

- *symmetric phase* (S) when the condition of Theorem 2.1(i) holds,
- *symmetry breaking phase* (SB) when the condition of Theorem 2.1(ii) hols.

We next explore some consequences of Theorem 2.1. To avoid redundancy we set $\mu = 1$ and put

$$p^* = \lambda/(1 + \lambda), \tag{2.6}$$

so that $\lambda = p^*/(1 - p^*)$. Note that p^* is the stationary probability that an edge is active. The following two propositions provide a partial phase classification.

Proposition 2.2. [Short-time SB] *If $u < r$, then for T sufficiently small $G_n(T)$ is SB.*

Proposition 2.3. [Monotonicity]

(i) *If $u = 0$ and $G_n(T)$ is S, then $G_n(T')$ is S for all $T' > T$.*
(ii) *If $u = 1$ and $G_n(T)$ is SB, then $G_n(T')$ is SB for all $T' > T$.*

2.1.2. *Numerics.* A natural choice of the constant initial graphon is $u = p^*$, i.e., the dynamics starts at a typical outcome of its stationary state. Figure 2 illustrates the consequences of varying T for the case where $d = 2$ (i.e., triangles). For large T and $p^* = \frac{1}{7}, \frac{1}{8}$, $G_n(T)$ is S for all $r \in [0,1]$, while for large T and $p^* = \frac{1}{9}, \frac{1}{10}$ there exists r such that $G_n(T)$ is SB. To understand why, observe that, for large T, $G_n(T)$ behaves like an Erdős-Rényi random graph with edge probability p^*. According to [19, Theorem 1.1], if $\mathrm{ERRG}_n(p^*)$ is an Erdős-Rényi random graph with edge probability p^*, then it is S for all $r \in [0,1]$ if and only if

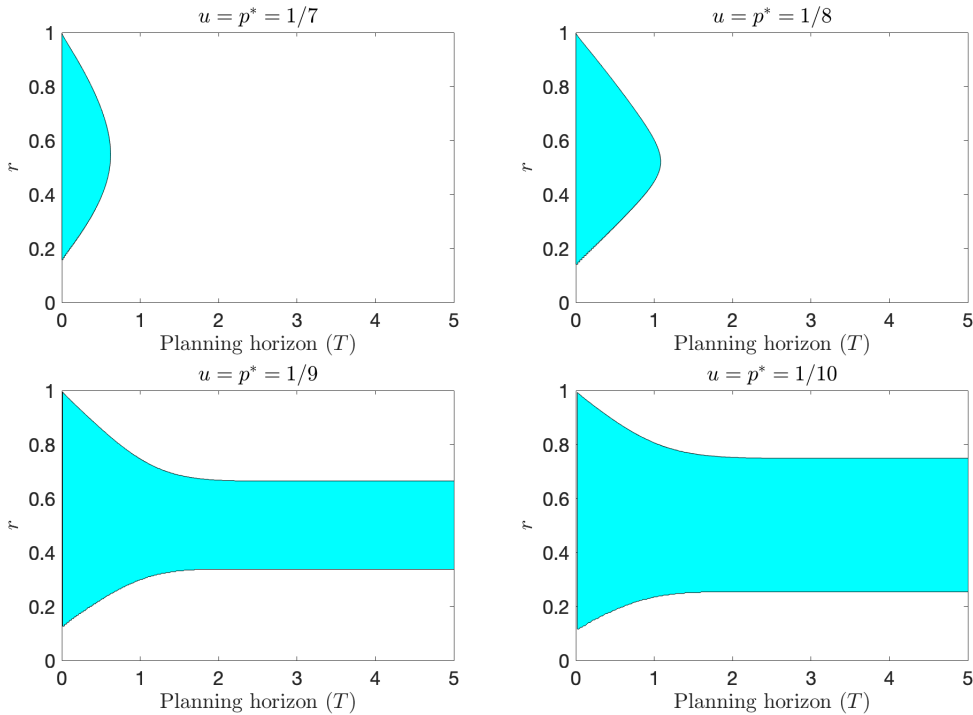$$p^* \geqslant (e^2 + 1)^{-1}. \tag{2.7}$$

Figure 2. Phase diagrams in (T, r) for $d = 2$ and $u \equiv p^*$ with $p^* = \frac{1}{7}, \frac{1}{8}, \frac{1}{9}, \frac{1}{10}$. The shaded region corresponds to SB, the unshaded region to S. Observe that SB prevails for small T and $r > u$.

A visual inspection of Figure 2 indicates that, as the planning horizon T increases, $G_n(T)$ can transition from SB to S. An informal explanation is the following. For small T it is more costly to add extra edges than for large T. Hence, for small T we expect to see graphs where the extra triangles are formed through the addition of a small number of extra edges arranged in a special structure (corresponding to SB), rather than through the addition of a large number of extra edges sprinkled uniformly (corresponding to S).

Because of the lack of structural results, we numerically consider additional values of p^*, namely, near the critical value $(e^2 + 1)^{-1}$. In Figure 3 we pick $u = 0$ (left column), $u = p^*$ (center column), and $u = 1$ (right column), and $p^* = (e^2 + 1)^{-1}$ (top row), $p^* = (e^2 + 1)^{-1} - 10^{-4}$ (middle row), and $p^* = (e^2 + 1)^{-1} - 2 \times 10^{-4}$ (bottom row). Note
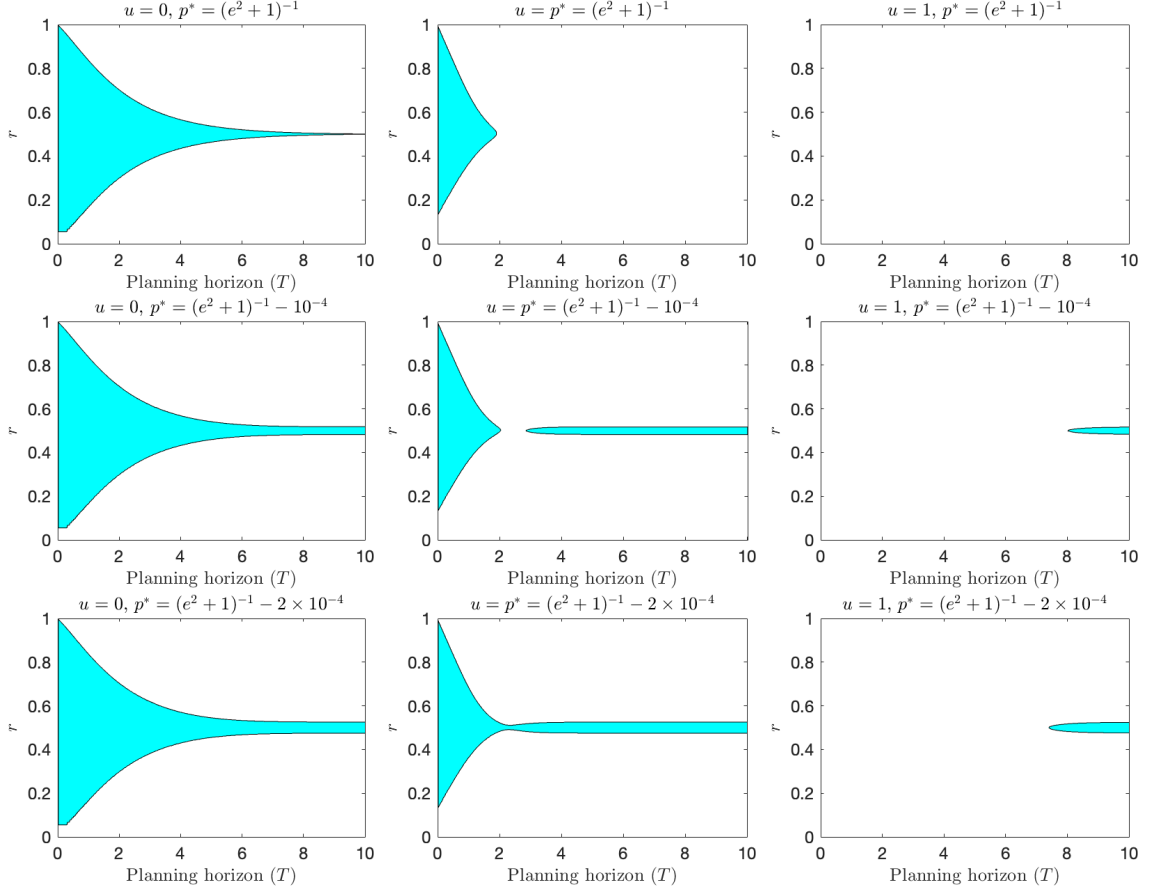
Figure 3. Phase diagrams in (T, r) for $u = 0, p^*, 1$ with p^* close to $(e^2 + 1)^{-1}$. The shaded region corresponds to SB, the unshaded region to S.

that, in line with Proposition 2.3, for $u = 0$ or $u = 1$ we observe at most one phase transition in the planning horizon T: from SB to S when $u = 0$ and from S to SB when $u = 1$. However, this is not so when $u = p^*$: when $u = p^* = (e^2 + 1)^{-1} - 10^{-4}$ and $u = p^* = (e^2 + 1)^{-1} - 2 \times 10^{-4}$, there are values of r such that, as T increases, $G_n(T)$ transitions from SB to S and back from S to SB. In other words, two phase transitions occur in the planning horizon T, i.e., a *re-entrant phase transition* is observed.

The re-entrant phase transition in T is quite distinct from the re-entrant phase transition in r (which was first observed in [7] and is evident from Figures 2 and 3). It is difficult to find a probabilistic explanation for the re-entrant phase transition in T. However, once we observe that, when $u = 0$, $G_n(T)$ can transition from SB to S and, when $u = 1$, $G_n(T)$ can transition from S to SB, then it is plausible that both are possible when we consider the intermediate value $u = p^*$. Moreover, in the light of Proposition 2.2, when $u = p^*$, $G_n(T)$ can only transition from S to SB *after* it has transitioned from SB to S.

2.2. **Application 2.** Suppose that the graphon valued process starts near a graphon $\tilde{u}$ at time 0, and is conditioned to end near another graphon $\tilde{r}$ at time T. A natural question

is the following. Is the most likely path necessarily unique, or is it possible that there are multiple most likely paths?

2.2.1. *Optimal paths.* The next theorem shows that we can answer this question by studying the set $\tilde{H}^* \subseteq \tilde{\mathscr{W}} \times [0, T]$ of paths that minimise $\tilde{I}$ subject to the condition that the path starts at $\tilde{u}$ and ends at $\tilde{r}$. For $\eta > 0$, put

$$\tilde{H}_\eta := \{\tilde{h} \in \tilde{\mathscr{W}} \times [0, T] \colon \delta_\square(\tilde{h}_T, \tilde{r}) \leqslant \eta\}, \tag{2.8}$$

and for $\tilde{h}, \tilde{h}' \in \tilde{\mathscr{W}} \times [0, T]$ define

$$\delta_\square^\infty(\tilde{h}, \tilde{h}') := \sup_{t \in [0,T]} \delta_\square(\tilde{h}_t, \tilde{h}'_t). \tag{2.9}$$

Theorem 2.4. [Optimal paths] *Let $\tilde{H} \subseteq \tilde{\mathscr{W}} \times [0, T]$ be the set of paths starting at $\tilde{u}$ and ending at $\tilde{r}$. Let $\tilde{H}^* \subseteq \tilde{H}$ be the set of minimisers of $\tilde{I}$ in $\tilde{H}$. Then $\tilde{H}^*$ is non-empty and compact. In addition, if $\lim_{n \to \infty} \delta_\square(\tilde{f}_{n,0}, \tilde{u}) = 0$, then*

$$\lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}\big(\delta_\square^\infty(\tilde{f}_n, \tilde{H}^*) \geqslant \varepsilon \mid \tilde{f}_n \in \tilde{H}_\eta\big) \leqslant -C, \tag{2.10}$$

where $C > 0$ is a constant that depends on $\tilde{u}$, $\tilde{r}$, T and ε.

2.2.2. *Variational problems.* To better understand the set $\tilde{H}^*$, we next solve two related variational problems, each with its own probabilistic interpretation.

Lemma 2.5. [Identification of minimiser] *Pick $u \in [0, r]$. Let*

$$f_{u \to r}^*(t) := \operatorname*{arg\,min}_{s \in [0,1]} [I_{1,t}(u, s) + I_{1,T-t}(s, r)], \qquad t \in [0, T]. \tag{2.11}$$

Then $f_{u \to r}^*$ is the unique minimiser of

$$I(f) = \int_0^T \mathrm{d}t\, \mathcal{L}(f(t), f'(t)), \tag{2.12}$$

subject to the condition $f(0) = u$ and $f(T) = r$, where $\mathcal{L}$ is defined in (1.27). In addition, $I(f_{u \to r}^*) = I_{1,T}(u, r)$.

Remark 2.6. Let $\{X_i(t)\}_{t \geqslant 0}$, $i \in \mathbb{N}$, be independent processes switching between active and inactive, with λ the rate of becoming active and μ the rate of becoming inactive, as before. Define

$$L_n(t) = \frac{1}{n} \sum_{i=1}^n X_i(t), \qquad \lim_{n \to \infty} L_n(0) = u. \tag{2.13}$$

Then, informally, we can interpret $f_{u \to r}^*$ as the most likely path that $(L_n(t))_{t \in [0,T]}$ takes from u to r when n is large. Such paths $f_{u \to r}^*$ can be efficiently computed (see Lemma 5.2 below). An illustration is given in Figure 4: in the right-hand-side the time horizon T is relatively large, so that the least costly way to reach r is by first falling back towards the equilibrium value $\frac{2}{3}$ and afterwards moving towards r in a relatively short time interval before T, whereas in the left-hand-side the the time horizon T is relatively small, so that the least costly way to reach r is by immediately moving towards it.
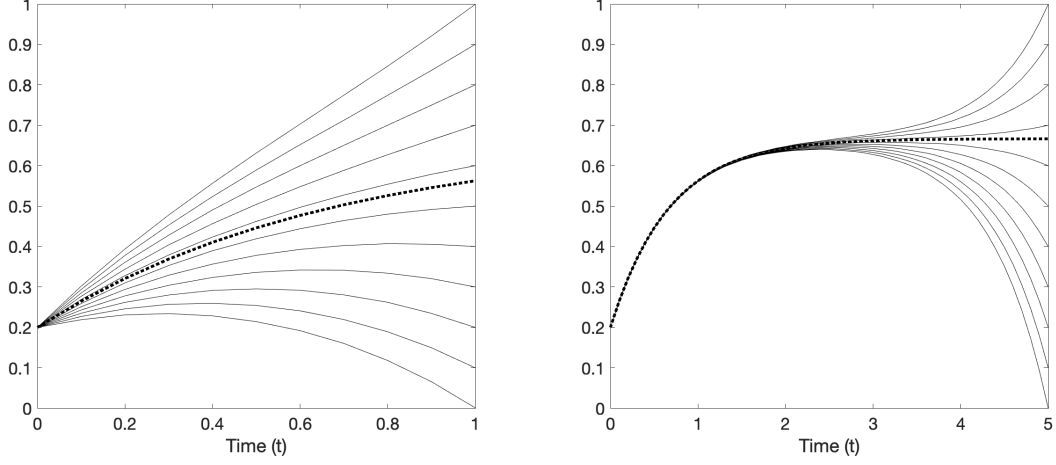
Figure 4. Solid curves: the paths $t \mapsto f_{u \to r}^*(t)$ for $\lambda = 1$, $\mu = \frac{1}{2}$, $u = \frac{1}{5}$ and $r = 0, \frac{1}{10}, \frac{2}{10}, \ldots, 1$ when $T = 1$ (left) and $T = 5$ (right). Dashed curve: most likely path without the terminal condition.

We need to define what we mean when we say that two paths $h, g \in \mathscr{W} \times [0, T]$ are equal. Define the equivalence relation '$\equiv$' by writing $h \equiv g$ if and only if

$$\text{Leb} \left\{ (x, y) \in [0, 1]^2 : \text{there exists a } t \in [0, T] \text{ such that } h_t(x, y) \neq g_t(x, y) \right\} = 0. \quad (2.14)$$

Below when we write $\mathscr{W} \times [0, T]$ we assume that this is the quotient space formed by the equivalence relation '$\equiv$'.

Lemma 2.7. [Identification of minimiser] *Set $u, r \in \mathscr{W}$. Let*

$$h_{u \to r}^*(x, y, t) = f_{u(x,y) \to r(x,y)}^*(t), \qquad t \in [0, T], \ (x, y) \in [0, 1]^2. \quad (2.15)$$

Then $h_{u \to T}^*$ is the unique minimiser of

$$I(h) = \int_0^T dt \int_{[0,1]^2} dx \, dy \, \mathcal{L}\big(h_t(x, y), h_t'(x, y)\big), \quad (2.16)$$

subject to the condition that $h_0 = u$ and $h_T = r$, where $\mathcal{L}$ is defined in (1.27). In addition, $I(h_{u \to r}^*) = I_{1,T}(u, r)$.

We next turn our attention to the original variational problem on $\tilde{\mathscr{W}} \times [0, T]$. If $\tilde{h} \in \tilde{H}^*$, then, armed with Lemma 2.7 and the specific form of $\tilde{I}$, we may expect that there exists a representative h of $\tilde{h}$ such that

$$I(h) = I_{1,T}(u, r^\sigma), \qquad h_T = r^\sigma, \quad (2.17)$$

for some $\sigma \in \mathscr{M}$. By Lemma 2.7, the only such paths are of the form $h_{u \to r}^*$. Theorem 2.8 below, which applies when u and r are *block graphons*, shows that we may restrict our attention to the equivalence classes of these paths, i.e., the set $\{\tilde{h}_{u \to r^\sigma}^*\}_{\sigma \in \mathscr{M}}$, and implies that we can replace the variational problem on $\tilde{\mathscr{W}} \times [0, T]$ by a significantly simpler one, in terms of permutations of the target graphon r.

For $I \in \mathbb{N}$, let $\mathscr{W}^{(I)}$ denote the space of block graphons with I^2 blocks, so that for any $g \in \mathscr{W}^{(I)}$ there exist *block endpoints* $0 = x_0 < x_1 < \cdots < x_I = 1$ such that

$$g(x, y) = g_{ij} \qquad \forall\, x, y \in [x_{i-1}, x_i) \times [x_{j-1}, x_j). \tag{2.18}$$

For $u \in \mathscr{W}^{(I)}$ and $r \in \mathscr{W}^{(J)}$ with block endpoints $0 = a_0 < a_1 < \cdots < a_I = 1$ and $0 = b_0 < b_1 < \cdots < b_J = 1$, respectively, let $\alpha \colon \mathscr{M} \mapsto [0,1]^{I \times J}$ be defined by

$$\alpha(\sigma)_{ij} := \operatorname{Leb}\{x \in [0,1] \colon x \in [a_{i-1}, a_i),\, \sigma(x) \in [b_{j-1}, b_j)\}, \qquad \sigma \in \mathscr{M}. \tag{2.19}$$

Note that α can map to any value in the compact set

$$V = \left\{ \boldsymbol{v} \in [0,1]^{I \times J} \colon \sum_{j \in J} v_{ij} = a_i - a_{i-1} \ \forall\, i \in I,\ \sum_{i \in I} v_{ij} = b_j - b_{j-1} \ \forall\, j \in J \right\}. \tag{2.20}$$

It is important to point out that, for any $\sigma_1, \sigma_2 \in \mathscr{M}$ with $\alpha(\sigma_1) = \alpha(\sigma_2)$,

$$\tilde{h}^*_{u \to r^{\sigma_1}} = \tilde{h}^*_{u \to r^{\sigma_2}}, \qquad I_{1,T}(u, r^{\sigma_1}) = I_{1,T}(u, r^{\sigma_2}). \tag{2.21}$$

For $v \in [0,1]^{I \times J}$, let $\sigma_{\boldsymbol{v}}$ be any element of $\mathscr{M}$ such that $\alpha(\sigma) = \boldsymbol{v}$.

Theorem 2.8. [Optimal paths] *Suppose that $u \in \mathscr{W}^{(I)}$ and $r \in \mathscr{W}^{(J)}$ for some $I, J \in \mathbb{N}$. Then*

$$\tilde{H}^* \subseteq \tilde{F}^* := \{\tilde{h}^*_{u \to r^\sigma}\}_{\sigma \in \mathscr{M}} = \{\tilde{h}^*_{u \to r^{\sigma_{\boldsymbol{v}}}}\}_{\boldsymbol{v} \in V}. \tag{2.22}$$

Moreover, if V^* is the set of $\boldsymbol{v} \in V$ that minimise $I_{1,T}(u, r^{\sigma_{\boldsymbol{v}}})$, then V^* is non-empty and

$$\tilde{H}^* = \{\tilde{h}^*_{u \to r^{\sigma_{\boldsymbol{v}}}}\}_{\boldsymbol{v} \in V^*}. \tag{2.23}$$

The requirement that u and r be block graphons is harmless, as block graphons can be used to approximate any graphon arbitrarily closely in L^2 [6, Proposition 2.6]. The following corollary is immediate because the constant graphon is invariant under permutation.

Corollary 2.9. [Uniqueness] *If either u or r is a constant graphon, then $\tilde{H}^* = \tilde{F}^* = \{\tilde{h}^*_{u \to r}\}$, i.e., both sets contain a single element.*

2.2.3. *Multiplicity.* We next explore whether $\tilde{H}^*$ can contain multiple paths. Theorem 2.8 provides a concrete criterion. In particular, $\tilde{H}^*$ contains multiple paths if and only if there exist $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ such that

$$\tilde{I}_{1,T}(\tilde{u}, \tilde{r}) = I_{1,T}(u, r^{\sigma_{\boldsymbol{v}_1}}) = I_{1,T}(u, r^{\sigma_{\boldsymbol{v}_2}}), \qquad \tilde{h}^*_{u \to r^{\sigma_{\boldsymbol{v}_1}}} \neq \tilde{h}^*_{u \to r^{\sigma_{\boldsymbol{v}_2}}}. \tag{2.24}$$

We thus need to determine whether these two conditions can be satisfied simultaneously.

We begin by focussing on the latter condition: Can $\tilde{F}^*$ (defined in Theorem 2.8) contain multiple paths? The answer is yes: even though for any $\tilde{h}, \tilde{g} \in \tilde{F}^*$ we have $\tilde{u} = \tilde{h}_0 = \tilde{g}_0$ and $\tilde{r} = \tilde{h}_T = \tilde{g}_T$, this does not necessarily imply that $\tilde{h}_t = \tilde{g}_t$ for $t \in (0, T)$. To see why, we consider the following simple example. Let $u, r \in \mathscr{W}$ be such that

$$u(x, y) = r(x, y) = \begin{cases} 1 & \text{if } x \leqslant \frac{1}{2},\, y \leqslant \frac{1}{2}, \\ 0 & \text{if } x \geqslant \frac{1}{2},\, y \geqslant \frac{1}{2}, \\ \frac{1}{2} & \text{otherwise,} \end{cases} \tag{2.25}$$

and $\sigma \in \mathcal{M}$ be such that

$$\sigma(x) = \begin{cases} x + \frac{1}{2} & \text{if } x \leqslant \frac{1}{2}, \\ x - \frac{1}{2} & \text{if } x > \frac{1}{2}. \end{cases} \tag{2.26}$$

Recalling the definition of $f_{a \to b}^*$ from Lemma 2.5, we have

$$h_{u \to r}^*(x, y, t) = \begin{cases} f_{1 \to 1}^*(t) & \text{if } x \leqslant \frac{1}{2}, \, y \leqslant \frac{1}{2}, \\ f_{0 \to 0}^*(t) & \text{if } x \geqslant \frac{1}{2}, \, y \geqslant \frac{1}{2}, \\ f_{\frac{1}{2} \to \frac{1}{2}}^*(t) & \text{otherwise}, \end{cases} \tag{2.27}$$

and

$$h_{u \to r^\sigma}^*(x, y, t) = \begin{cases} f_{1 \to 0}^*(t) & \text{if } x \leqslant \frac{1}{2}, \, y \leqslant \frac{1}{2}, \\ f_{0 \to 1}^*(t) & \text{if } x \geqslant \frac{1}{2}, \, y \geqslant \frac{1}{2}, \\ f_{\frac{1}{2} \to \frac{1}{2}}^*(t) & \text{otherwise}. \end{cases} \tag{2.28}$$

The three (x, y)-coordinate paths of $h_{u \to r}^*$ (solid) and $h_{u \to r^\sigma}^*$ (dashed) are illustrated in Figure 5 for $\lambda = \mu = 1$ and $T = 3$. It is easiest to see that $\tilde{h}_{u \to r}^* \neq \tilde{h}_{u \to r^\sigma}^*$ by looking at their values when $t = \frac{3}{2}$: $h_{u \to r^\sigma}^*(\cdot, \cdot, \frac{3}{2})$ is a constant graphon while $h_{u \to r}^*(\cdot, \cdot, \frac{3}{2})$ is not.
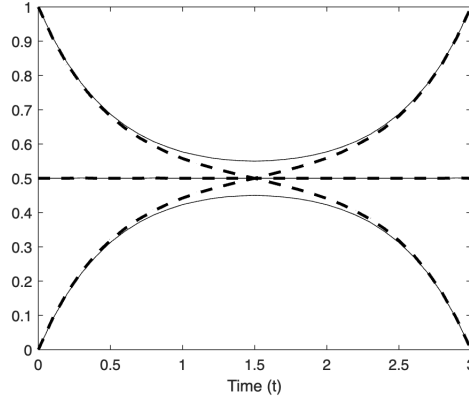
Figure 5. The (x, y)-coordinate paths of $h_{u \to r}^*$ (solid) and $h_{u \to r^\sigma}^*$ (dashed) for $\lambda = \mu = 1$, $T = 3$, with u, r and σ defined in (2.25) and (2.26).

While the above example demonstrates that $\tilde{F}^*$ can contain multiple elements, it does not imply that $\tilde{H}^*$ can contain multiple elements. Indeed, the next proposition implies that in the above example $\tilde{H}^*$ contains a single element: $\tilde{H}^* = \{\tilde{h}_{u \to r}^*\}$.

Proposition 2.10. [Uniqueness] *If there exist $\sigma_1, \sigma_2 \in \mathcal{M}$ such that, for any $a < b$ and $x \in [0, 1]$,*

$$u^{\sigma_2}(a, x) \leqslant u^{\sigma_2}(b, x), \qquad r^{\sigma_1}(a, x) \leqslant r^{\sigma_1}(b, x), \tag{2.29}$$

then $\tilde{H}^* = \{\tilde{h}_{u^{\sigma_2} \to r^{\sigma_1}}^*\}$.

u

	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$
$\frac{1}{5}$	u_{11}	b	a
$\frac{2}{5}$	b	a	u_{23}
$\frac{2}{5}$	a	u_{23}	b

r

	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$
$\frac{1}{5}$	r_{11}	c	d
$\frac{2}{5}$	c	c	r_{23}
$\frac{2}{5}$	d	r_{23}	d

r^σ

	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$
$\frac{1}{5}$	r_{11}	d	c
$\frac{2}{5}$	d	d	r_{23}
$\frac{2}{5}$	c	r_{23}	c

Figure 6. An example of u, r and σ for which $\tilde{h}^*_{u\to r} \neq \tilde{h}^*_{u\to r^\sigma}$ and $\tilde{I}(\tilde{h}^*_{u\to r}) = \tilde{I}(\tilde{h}^*_{u\to r^\sigma})$.

Is there always a single optimising path, i.e., does $\tilde{H}^*$ always contain a single element? To answer this question we consider the graphons u, r and r^σ illustrated in Figure 6. Let $A_1 = [0, \frac{1}{5})$, $A_2 = [\frac{1}{5}, \frac{3}{5})$ and $A_3 = [\frac{3}{5}, 1]$. For constants $a, b, c, d, u_{11}, u_{23}, r_{11}, r_{23} \in [0, 1]$, let

$$u(x,y) = \begin{cases} u_{11} & \text{if } (x,y) \in A_1^2, \\ a & \text{if } (x,y) \in A_2^2 \cup A_1 \times A_3 \cup A_3 \times A_1, \\ b & \text{if } (x,y) \in A_3^2 \cup A_1 \times A_2 \cup A_2 \times A_1, \\ u_{23} & \text{otherwise,} \end{cases} \qquad (2.30)$$

and

$$r(x,y) = \begin{cases} r_{11} & \text{if } (x,y) \in A_1^2, \\ c & \text{if } (x,y) \in A_2^2 \cup A_1 \times A_3 \cup A_3 \times A_1, \\ d & \text{if } (x,y) \in A_3^2 \cup A_1 \times A_2 \cup A_2 \times A_1, \\ r_{23} & \text{otherwise,} \end{cases} \qquad (2.31)$$

and let $\sigma \in \mathscr{M}$ be such that

$$\sigma(x) = \begin{cases} x & \text{if } x \in A_1, \\ x + \frac{2}{5} & \text{if } x \in A_2, \\ x - \frac{2}{5} & \text{if } x \in A_3. \end{cases} \qquad (2.32)$$

Note that $I(h^*_{u\to r}) = I_{1,T}(u,r) = I_{1,T}(u,r^\sigma) = I(h^*_{u\to r^\sigma})$, regardless of the values of $a, b, c, d, u_{11}, u_{23}, r_{11}, r_{23}$. However, to ensure that both conditions in (2.24) are satisfied, we need to select these parameters carefully. The next proposition tells us how we can do this.

Proposition 2.11. [Non-uniqueness] *Suppose that u, r, σ are given by* (2.30), (2.31), (2.32), *and set*

$$a = c = 0, \quad b = d = \varepsilon, \quad u_{11} = u_{23} = r_{11} = r_{23} = 1. \qquad (2.33)$$

Then, for $\varepsilon, T > 0$ sufficiently small, $\tilde{H}^* = \{\tilde{h}^*_{u\to r}, \tilde{h}^*_{u\to r^\sigma}\}$ and $\tilde{h}^*_{u\to r} \neq \tilde{h}^*_{u\to r^\sigma}$.

Through this example we are led to conclude that if the process begins near $\tilde{u}$ at time 0 and is conditioned to end at $\tilde{r}$ at time T, then it may take one of two equally likely paths. Note

that, in view of the counting and inverse counting lemmas [16, Lemmas 10.23 and 10.32], by specifying graphons $\tilde{u}$ and $\tilde{r}$ at times $t = 0$ and $t = T$ we are in effect specifying the subgraph density $t(H, \tilde{f}_{n,t})$ at times $t = 0$ and $t = T$ for every simple graph H. By these same lemmas, Proposition 2.11 shows that, for some subgraphs H, the subgraph density $t \mapsto t(H, \tilde{f}_{n,t})$ may take one of two equally likely paths from $t(H, \tilde{u})$ to $t(H, \tilde{r})$.

3. Proof of the two-point LDP

In this section we prove Theorem 1.2. We settle the lower semi-continuity of the rate function $I_{1,t}(\tilde{u}, \cdot)$ in Section 3.1, the upper bound of the LDP in Section 3.2, and the lower bound of the LDP in Section 3.3.

Abbreviate $\tilde{\mathbb{P}}_{n,t}^{\tilde{u}}(\cdot) := \mathbb{P}(\tilde{f}_{n,t} \in \cdot \mid \tilde{f}_{n,0} = \tilde{u})$ and $\mathbb{P}_{n,t}^u(\cdot) := \mathbb{P}(f_{n,t} \in \cdot \mid f_{n,0} = u)$. As before, let $\mathscr{W}^{(I)} \subseteq \mathscr{W}$ denote the space of block graphons with $I^2 \in \mathbb{N}$ blocks. Also define the ε-balls

$$\tilde{\mathbb{B}}_\square(\tilde{h}, \varepsilon) := \{\tilde{g} \in \tilde{\mathscr{W}} : \delta_\square(\tilde{h}, \tilde{g}) \leqslant \varepsilon\},$$

$$\mathbb{B}_\square(\tilde{h}, \varepsilon) := \{g \in \mathscr{W} : \delta_\square(\tilde{h}, \tilde{g}) \leqslant \varepsilon\}, \qquad (3.1)$$

$$\mathbb{B}_\square(h, \varepsilon) := \{g \in \mathscr{W} : d_\square(h, g) \leqslant \varepsilon\}.$$

Write $\tilde{\mathscr{W}}_n$ to denote the set of empirical graphons with n vertices.

We first state two properties of $I_{1,t}$, uniformity and convexity, that are needed along the way and are straightforward to verify.

Lemma 3.1. *For every $t > 0$ and $u, h \in [0, 1]$,*

$$I_{1,t}(u, h) \leqslant \max\{-\log p_{01,t}, -\log(1 - p_{11,t})\} < \infty. \qquad (3.2)$$

Moreover, for $\eta, \varepsilon > 0$, let

$$\Delta_I(\eta, \varepsilon) := \max_{u,h \in [0,1], u' \in [u-\eta, u+\eta], h' \in [h-\varepsilon, h+\varepsilon]} |I_{1,t}(u, h) - I_{1,t}(u', h')|. \qquad (3.3)$$

Then $\lim_{\eta, \varepsilon \downarrow 0} \Delta_I(\eta, \varepsilon) = 0$.

Lemma 3.2. *For every $t > 0$ and $u, h \in [0, 1]$, $u \mapsto I_{1,t}(u, h)$ is convex and $h \mapsto I_{1,t}(u, h)$ is strictly convex. Moreover, for every $u, h \in \mathscr{W}$ and $A, B \subseteq [0, 1]$,*

$$\frac{1}{\text{Leb}(A \times B)} \int_{A \times B} \mathrm{d}x \, \mathrm{d}y \, I_{1,t}(u(x, y), h(x, y)) \geqslant I_{1,t}\big(\bar{u}(A \times B), \bar{h}(A \times B)\big), \qquad (3.4)$$

where

$$\bar{u}(A \times B) := \frac{1}{\text{Leb}(A \times B)} \int_{A \times B} \mathrm{d}x \, \mathrm{d}y \, u(x, y), \qquad (3.5)$$

and $\bar{h}(A \times B)$ is defined similarly.

3.1. **Lower semi-continuity.** We first establish that $I_{1,t}(\tilde{u}, \cdot)$ is a good rate function, i.e., $I_{1,t}(\tilde{u}, \cdot) \not\equiv \infty$ and $\tilde{x} \mapsto \tilde{I}_{1,t}(\tilde{u}, \tilde{x})$ has compact level sets. Because $\tilde{\mathscr{W}}$ is compact, it suffices to show that $\tilde{x} \mapsto \tilde{I}_{1,t}(\tilde{u}, \tilde{x})$ is lower semi-continuous. We will in fact show that $(\tilde{u}, \tilde{x}) \mapsto \tilde{I}_{1,t}(\tilde{u}, \tilde{x})$ is lower semi-continuous, because this stronger property is needed below.

Lemma 3.3. *Suppose that $\lim_{n \to \infty} \delta_\square(\tilde{u}_n, \tilde{u}) = 0$ and $\lim_{n \to \infty} \delta_\square(\tilde{h}_n, \tilde{h}) = 0$. Then*

$$\liminf_{n \to \infty} \tilde{I}_{1,t}(\tilde{u}_n, \tilde{h}_n) \geqslant \tilde{I}_{1,t}(\tilde{u}, \tilde{h}). \qquad (3.6)$$

Proof. Let $\mathcal{A}_k$ be the set of all k-set partitions of $[0,1]$. In particular, $\mathcal{A}_k$ includes any $\{A_i\}_{i=1}^k$ such that $A_i \subseteq [0,1]$, $A_i \cap A_j = \varnothing$ for $i,j \in [k]$ with $i \neq j$, and $\cup_{i=1}^k A_i = [0,1]$. Let $(\sigma_n)_{n \in \mathbb{N}}$ be any sequence of elements in $\mathscr{M}$. Recalling (1.19)–(1.20) and applying Lemmas 3.1–3.2, we obtain, for any $k \in \mathbb{N}$,

$$
\begin{aligned}
&\liminf_{n \to \infty} I_{1,t}(u_n, h_n^{\sigma_n}) \\
&\geqslant \liminf_{n \to \infty} \sup_{\{A_i\}_{i=1}^k \in \mathcal{A}_k} \sum_{i,j=1}^k \mathrm{Leb}(A_i \times A_j) I_{1,t}\big(\bar{u}_n(A_i \times A_j), \overline{h_n^{\sigma_n}}(A_i \times A_j)\big) \\
&\geqslant \liminf_{n \to \infty} \sup_{\{A_i\}_{i=1}^k \in \mathcal{A}_k} \sum_{i,j=1}^k \mathrm{Leb}(A_i \times A_j) \hspace{3cm} (3.7) \\
&\qquad \times \big[I_{1,t}\big(\bar{u}(A_i \times A_j), \overline{h^{\sigma_n}}(A_i \times A_j)\big) - \Delta_I\big(d_\square(u_n, u), d_\square(h_n, h)\big) \big] \\
&\geqslant \inf_{\sigma \in \mathscr{M}} \sup_{\{A_i\}_{i=1}^k \in \mathcal{A}_k} \sum_{i,j=1}^k \mathrm{Leb}(A_i \times A_j) I_{1,t}\big(\bar{u}(A_i \times A_j), \overline{h^\sigma}(A_i \times A_j)\big).
\end{aligned}
$$

The proof is complete once we show that

$$
\lim_{k \to \infty} \inf_{\sigma \in \mathscr{M}} \sup_{\{A_i\}_{i=1}^k \in \mathcal{A}_k} \sum_{i,j=1}^k \mathrm{Leb}(A_i \times A_j) I_{1,t}\big(\bar{u}(A_i \times A_j), \overline{h^\sigma}(A_i \times A_j)\big) = \inf_{\sigma \in \mathscr{M}} I_{1,t}(u, h^\sigma). \quad (3.8)
$$

1. First we establish (3.8) when u and h are *block graphons*, i.e., $u \in \mathscr{W}^{(I)}$ and $h \in \mathscr{W}^{(J)}$. Let $0 = a_0 < a_1 < \cdots < a_I = 1$ and $0 = b_0 < b_1 < \cdots < b_J = 1$ denote their block endpoints, $A_i = [a_{i-1}, a_i)$ and $B_i = [b_{i-1}, b_i)$ their block intervals, and $\{u_{ij}\}_{1 \leqslant i,j \leqslant I}$ and $\{h_{\ell m}\}_{1 \leqslant \ell,m \leqslant J}$ their block values. (For example, if $x \in A_i$ and $y \in A_j$, then $u(x,y) = u_{ij}$.) For $i \in [I]$ and $j \in [J]$, let

$$
C_{J(i-1)+j} := \big\{ x \in [0,1] \colon x \in A_i, \sigma^{-1}(x) \in B_j \big\}, \quad (3.9)
$$

Suppose that $k = IJ$. Then (3.9) defines a sequence of sets $C_1, \ldots, C_k$. We have

$$
\begin{aligned}
&\sum_{i,j=1}^k \mathrm{Leb}(C_i \times C_j) \, I_{1,t}\big(\bar{u}(C_i \times C_j), \overline{h^\sigma}(C_i \times C_j)\big) \\
&\qquad\qquad = \sum_{i,j \in [I], \, \ell,m \in [J]} \mathrm{Leb}(C_{J(i-1)+\ell} \times C_{J(j-1)+m}) \, I_{1,t}(u_{ij}, h_{\ell m}) \quad (3.10) \\
&\qquad\qquad = \int_{[0,1]^2} \mathrm{d}x \, \mathrm{d}y \, I_{1,t}\big(u(x,y), h^\sigma(x,y)\big) = I_{1,t}(u, h^\sigma),
\end{aligned}
$$

where the first and second equality are obtained by observing that u and h^σ are constant on $C_i \times C_j$. Since this holds for any $\sigma \in \mathscr{M}$ and $k \geqslant IJ$ (for $k > IJ$ simply take $C_i = \varnothing$ for all $i > IJ$), we have established (3.8).

To explain the above in a bit more detail, suppose that each point $x \in [0,1]$ is a vertex. Because u and h are block graphons with I and J blocks, respectively, we can think of vertices as being of type $1, \ldots, I$ at time 0 and of type $1, \ldots, J$ at time t. We would like $C_{J(i-1)+j}$ to contain all vertices of type i at time 0 and of type j at time t. It is clear that this means that $x \in A_i$. However, because we have applied an arbitrary permutation σ to h to get h^σ, the types of all the vertices at time t have been mixed up. Nonetheless, we know

that vertex x is of type j at time t if it maps to block j in h when the permutation j is undone (i.e., $\sigma^{-1}(x) \in B_j$). Now, $C_{J(i-1)+j}$ contains all vertices that are of type i at time 0 and of type j at time t. Hence, to arrive at (3.10), simply note that the density of edges between the vertices in $C_{J(i-1)+j}$ and $C_{J(i'-1)+j'}$ at time 0 is $u_{ii'}$, and that the density of edges between the vertices in $C_{J(i-1)+j}$ and $C_{J(i'-1)+j'}$ at time t is $h_{jj'}$.

2. Next we establish (3.8) when u and h are not block graphons by relying on a limiting argument. For $\ell \in \mathbb{N}$ and $g \in \mathscr{W}$, let $\widehat{g}^{(\ell)} \in \mathscr{W}^{(\ell)}$ be the block graphon such that if $i, j \in [\ell]$ and $(x, y) \in [\frac{i-1}{\ell}, \frac{i}{\ell}) \times [\frac{j-1}{\ell}, \frac{j}{\ell}) =: B_{ij}^{(\ell)}$, then

$$\widehat{g}^{(\ell)}(x, y) = \ell^2 \int_{B_{i,j}^{(\ell)}} \mathrm{d}x' \, \mathrm{d}y' \, g(x', y'). \tag{3.11}$$

Applying Lemma 3.1, we have, for any $\sigma \in \mathscr{M}$,

$$\begin{aligned} &|I_{1,t}(u, h^\sigma) - I_{1,t}(\widehat{u}^{(\ell)}, (\widehat{h}^{(\ell)})^\sigma)| \\ &\leqslant \int_{[0,1]^2} \mathrm{d}x \, \mathrm{d}y \, \Delta_I\big(u(x,y) - \widehat{u}^{(\ell)}(x,y), h(x,y) - (\widehat{h}^{(\ell)})^\sigma(x,y)\big) \\ &\leqslant \Delta_I(\varepsilon, \eta) + 2\max\{-\log p_{01,t}, -\log(1 - p_{11,t})\} \\ &\quad \times \mathrm{Leb}\Big\{(x,y) \in [0,1]^2: \ |u(x,y) - \widehat{u}^{(\ell)}(x,y)| \geqslant \varepsilon \ \text{or} \ |h(x,y) - \widehat{h}^{(\ell)}(x,y)| \geqslant \eta\Big\}. \end{aligned} \tag{3.12}$$

Because $\widehat{u}^{(\ell)} \to u$ and $\widehat{h}^{(\ell)} \to h$ in L^2 as $\ell \to \infty$ ([6, Proposition 2.6]), for any $\varepsilon, \eta > 0$ the second term in the right-hand side of (3.12) tends to 0 as $\ell \to \infty$. Letting $\varepsilon, \eta \downarrow 0$ and applying Lemma 3.1 once more, we obtain

$$\lim_{\ell \to \infty} \inf_{\sigma \in \mathscr{M}} I_{1,t}(\widehat{u}^{(\ell)}, (\widehat{h}^{(\ell)})^\sigma) = \inf_{\sigma \in \mathscr{M}} I_{1,t}(u, h^\sigma). \tag{3.13}$$

Noting that convergence in L^2 implies convergence in the cut distance, and using the fact that we have already established (3.8) for block graphons, we find

$$\begin{aligned} &\lim_{k \to \infty} \inf_{\sigma \in \mathscr{M}} \sup_{\{A_i\}_{i=1}^k \in \mathcal{A}_k} \sum_{i,j=1}^k \mathrm{Leb}(A_i \times A_j) I_{1,t}\big(\overline{u}(A_i \times A_j), \overline{h^\sigma}(A_i \times A_j)\big) \\ &= \lim_{\ell \to \infty} \lim_{k \to \infty} \inf_{\sigma \in \mathscr{M}} \sup_{\{A_i\}_{i=1}^k \in \mathcal{A}_k} \sum_{i,j=1}^k \mathrm{Leb}(A_i \times A_j) I_{1,t}\Big(\overline{\widehat{u}^{(\ell)}}(A_i \times A_j) \overline{(\widehat{h}^{(\ell)})^\sigma}(A_i \times A_j)\Big) \\ &= \lim_{\ell \to \infty} \inf_{\sigma \in \mathscr{M}} I_{1,t}\big(\widehat{u}^{(\ell)}, (\widehat{h}^{(\ell)})^\sigma\big) = \inf_{\sigma \in \mathscr{M}} I_{1,t}(u, h^\sigma), \end{aligned} \tag{3.14}$$

which completes the proof of (3.8).   □

3.2. Upper bound.
We start by observing that

$$\mathbb{P}(\tilde{f}_{n,t} \in \cdot \mid \tilde{f}_{n,0} = \tilde{u}_n) = \mathbb{P}(\tilde{f}_{n,t} \in \cdot \mid f_{n,0} = u_n). \tag{3.15}$$

Indeed, due to the fact that the dynamics is homogeneous, the outcome of $\tilde{f}_{n,t}$ is independent of the specific representative of $\tilde{u}_n$. We first establish the upper bound when $u \in \mathscr{W}^{(I)}$ for some $I \in \mathbb{N}$, i.e., the limiting initial graphon has a block structure. Afterwards we can use a limiting argument to obtain the upper bound for $u \in \mathscr{W}$, which we will not spell out.

Lemma 3.4. *Suppose that $u \in \mathscr{W}^{(I)}$ for some $I \in \mathbb{N}$, and $u_n^\eta \in \mathbb{B}_\square(u, \eta)$ for all $\eta > 0$ and n large enough. Then*

$$\lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}(\tilde{f}_{n,t} \in \tilde{C} \mid f_{n,t} = u_n^\eta) \leqslant - \inf_{\tilde{x} \in \tilde{C}} \tilde{I}_{1,t}(\tilde{u}, \tilde{x}) \tag{3.16}$$

for any closed set $\tilde{C} \subseteq \tilde{\mathscr{W}}$.

Proof. By [6, Lemma 4.1], it suffices to prove that, for all $\tilde{h} \in \tilde{\mathscr{W}}$,

$$\lim_{\varepsilon \downarrow 0} \lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}(\tilde{f}_{n,t} \in \tilde{\mathbb{B}}_\square(\tilde{h}, \varepsilon) \mid \tilde{f}_{n,0} = \tilde{u}_n^\eta) \leqslant -\tilde{I}_{1,t}(\tilde{u}, \tilde{h}), \tag{3.17}$$

which is equivalent to

$$\lim_{\varepsilon \downarrow 0} \lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}(f_{n,t} \in \mathbb{B}_\square(\tilde{h}, \varepsilon) \mid f_{n,0} = u_n^\eta) \leqslant -\tilde{I}_{1,t}(\tilde{u}, \tilde{h}), \tag{3.18}$$

by which we have transferred the problem from $\tilde{\mathscr{W}}$ to $\mathscr{W}$. Note that to get (3.18) we have applied (3.15) to replace $\tilde{f}_{n,0} = \tilde{u}_n^\eta$ by $f_{n,0} = u_n^\eta$ in the condition. Since (3.15) holds for any u_n^η in the equivalence class $\tilde{u}_n^\eta$, we may assume that there exists a u in the equivalence class $\tilde{u}$ such that $d_\square(u_n^\eta, u) < \eta$ for all n large enough. The proof consists of 6 steps.

1. In contrast to $\tilde{\mathbb{B}}_\square(\tilde{h}, \varepsilon)$, whose elements cling tightly to $\tilde{h}$, the elements of $\mathbb{B}_\square(\tilde{h}, \varepsilon)$ are scattered throughout $\mathscr{W}$. We therefore need a systematic method of collecting these elements. To this end we recall a version of Szemerédi's regularity lemma [6, Theorem 3.1], which states that for any $\varepsilon > 0$ there exist a constant $C(\varepsilon) < \infty$ and a set $\mathscr{W}(\varepsilon) \subseteq \mathscr{W}$ with $|\mathscr{W}(\varepsilon)| \leqslant C(\varepsilon)$ such that for any $f \in \mathscr{W}$ there exist $\phi \in \mathscr{M}$ and $g \in \mathscr{W}(\varepsilon)$ satisfying $d_\square(f^\phi, g) < \varepsilon$, and that for any $g \in \mathscr{W}(\varepsilon)$ there exists a $J \in \mathbb{N}$ such that $g \in \mathscr{W}^{(J)}$. Thus, if we let

$$\mathbb{B}_\square(\mathscr{W}(\varepsilon), \varepsilon) = \left\{ f \in \mathscr{W} : \min_{g \in \mathscr{W}(\varepsilon)} d_\square(g, f) \leqslant \varepsilon \right\}, \tag{3.19}$$

then

$$\{f_{n,t} \in \mathbb{B}_\square(\tilde{h}, \varepsilon)\} \subseteq \{f_{n,t} \in \mathbb{B}_\square(\tilde{h}, \varepsilon)\} \cap \left(\bigcup_{\sigma_n \in \mathscr{M}_n} \{f_{n,t}^{\sigma_n} \in \mathbb{B}_\square(\mathscr{W}(\varepsilon), \varepsilon)\} \right)$$
$$= \bigcup_{g \in \mathscr{W}(\varepsilon)} \bigcup_{\sigma_n \in \mathscr{M}_n} \{f_{n,t} \in \mathbb{B}_\square(\tilde{h}, \varepsilon)\} \cap \{f_{n,t}^{\sigma_n} \in \mathbb{B}_\square(g, \varepsilon)\}, \tag{3.20}$$

where $\mathscr{M}_n$ is the set of permutations of the n intervals of length $1/n$ in $[0, 1]$. Because $\mathscr{W}(\varepsilon)$ is finite, it is enough to show that

$$\lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}^{u_n^\eta} \left(\bigcup_{\sigma_n \in \mathscr{M}_n} \{f_{n,t} \in \mathbb{B}_\square(\tilde{h}, \varepsilon)\} \cap \{f_{n,t}^{\sigma_n} \in \mathbb{B}_\square(g, \varepsilon)\} \right)$$
$$\leqslant -\tilde{I}_{1,t}(u, \tilde{h}) + E(\varepsilon) \qquad \forall g \in \mathscr{W}(\varepsilon), \tag{3.21}$$

where $E(\varepsilon)$ must vanish as $\varepsilon \downarrow 0$. Note that the event in (3.21) is empty when $\delta_\square(\tilde{g}, \tilde{h}) > 2\varepsilon$. We thus only need to establish (3.21) when $\delta_\square(\tilde{g}, \tilde{h}) \leqslant 2\varepsilon$. Observe that the left-hand side of

(3.21) is at most

$$
\lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}^{u_n^\eta} \Big(\bigcup_{\sigma_n \in \mathscr{M}_n} \{ f_{n,t}^{\sigma_n} \in \mathbb{B}_\square(g, \varepsilon) \} \Big)
$$
$$
\leqslant \lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \max_{\sigma_n \in \mathscr{M}_n} \log \mathbb{P}^{u_n^\eta} \left(f_{n,t}^{\sigma_n} \in \mathbb{B}_\square(g, \varepsilon) \right) \tag{3.22}
$$

because $\log(n!) = o(\binom{n}{2})$. To bound the right-hand side of (3.22), we show that we can replace $\mathscr{M}_n$ by a finite set $\mathcal{T} = \mathcal{T}(I, J, \eta)$ (whose cardinality does not depend on n) without incurring a significant error.

2. We construct the set $\mathcal{T} = \mathcal{T}(I, J, \eta)$ as in the proof of [10, Lemma 3.3]. Recall that $u \in \mathscr{W}^{(I)}$ and $g \in \mathscr{W}^{(J)}$, and write $0 = a_0 < a_1 < \cdots < a_I = 1$ and $0 = b_0 < b_1 < \cdots < b_J = 1$ to denote their block endpoints. Define the intervals $A_i = [a_{i-1}, a_i)$ and $B_j = [b_{j-1}, b_j)$. Let

$$
V := \left\{ (v_{i,j})_{i \in [I], j \in [J]} \colon v_{ij} \in (0, 1), \sum_{j \in [J]} v_{ij} = \mathrm{Leb}(A_i), \sum_{i \in [I]} v_{ij} = \mathrm{Leb}(B_j) \right\}, \tag{3.23}
$$

and for $\boldsymbol{v} \in V$ define

$$
A_{ij} := \left[a_i + \sum_{k=1}^{j-1} v_{i,k}, a_i + \sum_{k=1}^{j} v_{i,k} \right), \qquad B_{ji} := \left[b_j + \sum_{k=1}^{i-1} v_{k,j}, b_j + \sum_{k=1}^{i} v_{k,j} \right). \tag{3.24}
$$

Pick $\tau_{\boldsymbol{v}} \in \mathscr{M}$ satisfying (see Figure 7)

$$
\tau_{\boldsymbol{v}}(A_{ij}) = B_{ij}, \quad \forall\, i \in [I], j \in [J]. \tag{3.25}
$$

Concretely, this means that we choose $\tau_{\boldsymbol{v}}$ such that if $x \in A_{ij}$, then

$$
\tau_{\boldsymbol{v}}(x) = \left(x - a_i + \sum_{k=1}^{j-1} v_{ik} \right) + \left(b_j + \sum_{k=1}^{i-1} v_{kj} \right). \tag{3.26}
$$

The map $\tau_{\boldsymbol{v}}$ can be understood as follows. For a vertex $v \in [n]$ and a set $A \subseteq [0, 1]$, write $v \rightsquigarrow A$ when $[\frac{v-1}{n}, \frac{v}{n}) \subseteq A$. Refer to v such that $v \rightsquigarrow A_i$ as a type-i vertex. The interval A_{ij} contains roughly $n v_{ij}$ type-i vertices, which are the only type-i vertices that get mapped onto the interval B_j. Thus, under the map $\tau_{\boldsymbol{v}}$, B_j contains roughly $n v_{ij}$ type-i vertices. Note also that, after $\tau_{\boldsymbol{v}}$ has been applied, the labels of type-i vertices inside each block are sorted in increasing order.

3. We have now introduced all the objects that are needed to construct the set $\mathcal{T}$. We have the set V and a mapping τ that relates elements of V to permutations. As $\mathcal{T}$ must be *finite* while V is uncountably infinite, we cannot let $\mathcal{T}$ be simply the image of V under τ. Instead, we construct a finite subset $\bar{V}$ of V such that any element of V is close to an element of $\bar{V}$ (exploiting the compactness of V), after which we let $\mathcal{T}$ be the image of $\bar{V}$ under τ. Concretely, we let $\bar{V} \subseteq V$ be a finite set such that for any $\boldsymbol{u} \in V$ there exists a $\boldsymbol{v} \in \bar{V}$ with

$$
\| \boldsymbol{u} - \boldsymbol{v} \|_\infty < \frac{\eta}{2IJ}. \tag{3.27}
$$
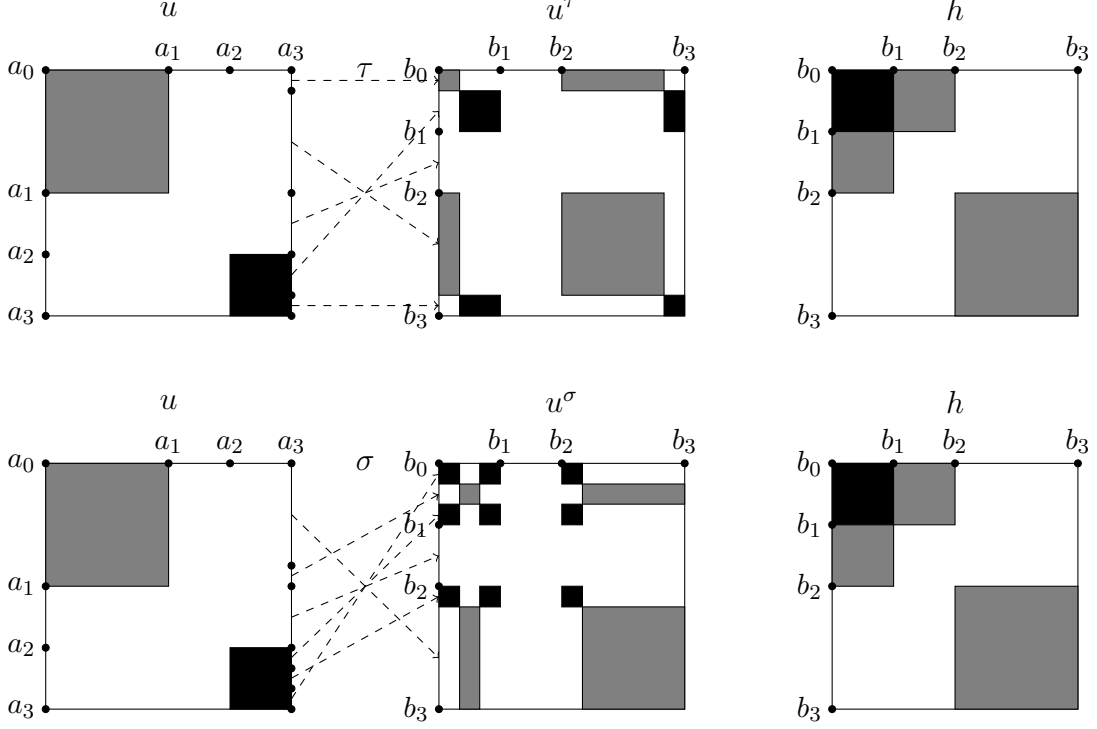
Figure 7. Illustration of the map τ in (3.25).

After that we put $\mathcal{T} := \{\tau_{\boldsymbol{v}}\colon \boldsymbol{v} \in \bar{V}\}$. It should be noted that if $\sigma \in \mathcal{M}$ and $C_{ij}(\sigma) := \{v \in [n]\colon v \leadsto A_i, \sigma_n(v) \leadsto B_j\}$, then for any $\sigma_n \in \mathcal{M}_n$ there exists $\tau \in \mathcal{T}$ such that

$$\frac{1}{n}\sum_{i,j}|C_{ij}(\sigma_n) - C_{ij}(\tau)| < \eta, \tag{3.28}$$

provided n is sufficiently large. In other words, for any $\sigma_n \in \mathcal{M}_n$ there exists a permutation $\tau \in \mathcal{T}$ that maps approximately the same proportion of type-i vertices to the interval B_j (see Figure 7 for an illustration). 'Note that we require n to be large to account for boundary effects, i.e., for $v \in [n]$ such that $[\frac{v-1}{n}, \frac{v}{n})$ is not contained in a single A_i or such that $\tau([\frac{v-1}{n}, \frac{v}{n}))$ is not contained in a single B_j.

4. Let $\sigma_n^0 \in \mathcal{M}_n$ be a permutation that permutes the blocks B_j only, and sorts the different vertices within B_j in ascending order of their original label. Formally this means that σ_n^0 satisfies the following properties:

- If $\sigma_n(u) \leadsto B_j$, then $\sigma^0 \circ \sigma_n(u) \leadsto B_j$.
- If $\sigma_n(u), \sigma_n(v) \leadsto B_j$ and $u \leadsto A_{i_1}$, $v \leadsto A_{i_2}$ with $i_1 < i_2$, then $(\sigma_n^0 \circ \sigma_n)(u) < (\sigma_n^0 \circ \sigma_n)(v)$.
- If $\sigma_n(u), \sigma_n(v) \leadsto B_j$ and $u, v \leadsto A_i$ with $u < v$, then $(\sigma_n^0 \circ \sigma_n)(u) < (\sigma_n^0 \circ \sigma_n)(v)$.

Observe that, because $g = g^{\sigma_n^0}$,

$$\mathbb{P}^{u_n^\eta}(f_{n,t}^{\sigma_n} \in \mathbb{B}_\square(g, \varepsilon)) = \mathbb{P}^{u_n^\eta}\left(f_{n,t}^{\sigma_n \circ \sigma_n^0} \in \mathbb{B}_\square(g, \varepsilon)\right) = \mathbb{P}^{(u_n^\eta)^{\sigma_n \circ \sigma_n^0}}(f_{n,t} \in \mathbb{B}_\square(g, \varepsilon)), \tag{3.29}$$

where the last equality follows from the fact that applying $\sigma_n \circ \sigma_0$ at time 0 is equivalent to applying it at time t. Now, by (3.27), for any $\sigma_n \in \mathcal{M}_n$ there exists a $\tau \in \mathcal{T}$ such that

$$d_\square\big((u_n^\eta)^{\sigma_n \circ \sigma_n^0}, u^\tau\big) \leqslant d_\square\big((u_n^\eta)^{\sigma_n \circ \sigma_n^0}, u^{\sigma_n \circ \sigma_n^0}\big) + d_\square\big(u^{\sigma_n \circ \sigma_n^0}, u^\tau\big) \leqslant 2\eta. \tag{3.30}$$

Consequently, we have derived the upper bound

$$\lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \max_{\sigma_n \in \mathcal{M}_n} \log \mathbb{P}^{u_n^\eta}\left(\{f_{n,t}^{\sigma_n} \in \mathbb{B}_\square(g, \varepsilon)\}\right)$$

$$\leqslant \lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \max_{\tau \in \mathcal{T}} \sup_{\bar{u}_n^\eta \in \mathbb{B}_\square(u^\tau, 2\eta)} \mathbb{P}^{\bar{u}_n^\eta}(f_{n,t} \in \mathbb{B}_\square(g, \varepsilon)) \tag{3.31}$$

for every $\varepsilon > 0$ and $g \in \mathcal{W}$.

5. To further bound the right-hand side of (3.31), let $\{u_{ij}\}_{i,j \in [I]}$ and $\{g_{\ell m}\}_{\ell,m \in [J]}$ be the block values of u and g, respectively (for example, if $x \in A_i$ and $y \in A_j$, then $u(x, y) = u_{ij}$). We assume without loss of generality that $v_{ij} > 0$ for all i, j (recall (3.23); the i, j with $v_{ij} = 0$ can be ignored). Abbreviate

$$\eta_{i\ell jm} := \frac{2\eta}{v_{i\ell} v_{jm}}, \qquad \varepsilon_{i\ell jm} := \frac{\varepsilon}{v_{i\ell} v_{jm}}. \tag{3.32}$$

Observe that, for each $i, j \in [I]$ and $k, \ell \in [J]$, if $d_\square(\bar{u}_n^\eta, u^\tau) \leqslant 2\eta$, then

$$\int_{A_{ik} \times A_{j\ell}} \mathrm{d}x \, \mathrm{d}y \, \bar{u}_n(x, y) \in [u_{ij} - \eta_{i\ell jm}, u_{ij} + \eta_{i\ell jm}], \tag{3.33}$$

while if $f_{n,t} \in \mathbb{B}_\square(g, \varepsilon)$, then

$$\int_{A_{ik} \times A_{j\ell}} \mathrm{d}x \, \mathrm{d}y \, f_{n,t}(x, y) \in [g_{k\ell} - \varepsilon_{i\ell jm}, g_{k\ell} + \varepsilon_{i\ell jm}]. \tag{3.34}$$

Because the rectangle $A_{i\ell} \times A_{jm}$ represents $n^2 v_{i\ell} v_{jm}$ independently evolving edges, we have

$$\limsup_{n \to \infty} \frac{1}{n^2 v_{i\ell} v_{jm}} \log \mathbb{P}\Bigg(\int_{A_{i\ell} \times A_{jm}} \mathrm{d}x \, \mathrm{d}y \, f_{n,t}(x, y) \in [g_{\ell m} - \varepsilon_{i\ell jm}, g_{\ell m} + \varepsilon_{i\ell jm}] \, \Bigg|$$

$$\int_{A_{i\ell} \times A_{jm}} \mathrm{d}x \, \mathrm{d}y \, f_{n,0}(x, y) \in [u_{ij} - \eta_{i\ell jm}, u_{ij} + \eta_{i\ell jm}] \Bigg) \tag{3.35}$$

$$\leqslant - \inf_{\bar{u} \in [u_{ij} - \eta_{i\ell jm}, u_{ij} + \eta_{i\ell jm}], \, x \in [g_{\ell m} - \varepsilon_{i\ell jm}, g_{\ell m} + \varepsilon_{i\ell jm}]} I_{1,t}(\bar{u}, x),$$

while this upper bound must be multiplied by $\frac{1}{2}$ when $(i, \ell) = (j, m)$ (to avoid double counting). This leads to

$$\limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}^{\bar{u}_n^\eta}(f_{n,t} \in \mathbb{B}_\square(g, \varepsilon))$$

$$\leqslant - \sum_{i\ell jm} v_{i\ell} v_{jm} \inf_{\bar{u} \in [u_{ij} - \eta_{i\ell jm}, u_{ij} + \eta_{i\ell jm}], \, x \in [g_{k\ell} - \varepsilon_{ikj\ell}, g_{k\ell} + \varepsilon_{ikj\ell}]} I_{1,t}(\bar{u}, x) \tag{3.36}$$

$$\leqslant - I_{1,t}(u^\tau, g) + \sum_{i\ell jm} v_{i\ell} v_{jm} \Delta_I\left(\eta_{i\ell jm} \wedge 1, \varepsilon_{i\ell jm} \wedge 1\right),$$

where $\Delta_I(\eta, \varepsilon)$ is defined in (3.3). Regarding the last inequality, note that because we are dealing with block graphons the integral in the definition of $I_{1,t}$ can be expressed as a sum with weights given by v_{ij}. Set $\gamma > 0$. If $v_{i\ell} v_{jm} > \gamma$, then

$$v_{i\ell} v_{jm} \, \Delta_I \left(\eta_{i\ell jm} \wedge 1, \varepsilon_{i\ell jm} \wedge 1 \right) \leqslant \Delta_I \left(\tfrac{2\eta}{\gamma}, \tfrac{\varepsilon}{\gamma} \right), \tag{3.37}$$

whereas if $v_{i\ell} v_{jm} \leqslant \gamma$ then, by Lemma 3.1,

$$v_{i\ell} v_{jm} \Delta_I \left(\eta_{i\ell jm}, \tfrac{\varepsilon}{\gamma} \right) \leqslant C\gamma \tag{3.38}$$

with $C := \max\{-\log p_{01,t}, -\log(1 - p_{11,t}\}$. Consequently, for any $\boldsymbol{v} \in V$ and $\gamma > 0$, we have

$$\sum_{i\ell jm} v_{i\ell} v_{jm} \, \Delta_I \left(\eta_{i\ell jm} \wedge 1, \varepsilon_{i\ell jm} \wedge 1 \right) \leqslant I^2 J^2 \left(C\gamma + \Delta_I \left(\tfrac{2\eta}{\gamma}, \tfrac{\varepsilon}{\gamma} \right) \right) =: E(\gamma, \eta, \varepsilon). \tag{3.39}$$

Combining the above formulas, we arrive at

$$\begin{aligned}
\lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \max_{\tau \in \mathcal{T}} \sup_{\bar{u}_n^\eta \in \mathbb{B}_\square(u^\tau, 2\eta)} & \mathbb{P}^{\bar{u}_n^\eta}(f_{n,t} \in \mathbb{B}_\square(g, \varepsilon)) \\
& \leqslant - \lim_{\eta \downarrow 0} \min_{\tau \in \mathcal{T}} \left[I_{1,t}(u^\tau, g) + E(\gamma, \eta, \varepsilon) \right] \\
& \leqslant - \inf_{\phi \in \mathcal{M}} I_{1,t}(u^\phi, g) + E(\gamma, 0, \varepsilon) = -\tilde{I}_{1,t}(\tilde{u}, \tilde{g}) + E(\gamma, 0, \varepsilon).
\end{aligned} \tag{3.40}$$

Picking $\gamma = \varepsilon^{1/2}$, we can apply Lemma 3.1, to obtain

$$E(\varepsilon) := E(\varepsilon^{1/2}, 0, \varepsilon) \downarrow 0, \qquad \varepsilon \downarrow 0. \tag{3.41}$$

6. We can now finally prove (3.21). Recall that we only need to consider $g \in \mathscr{W}(\varepsilon)$ such that $\tilde{\delta}_\square(\tilde{g}, \tilde{h}) \leqslant 2\varepsilon$. In view of (3.31), (3.40) and (3.41), it is enough to show that

$$-\tilde{I}_{1,t}(\tilde{u}, \tilde{g}) \leqslant -\tilde{I}_{1,t}(\tilde{u}, \tilde{h}) + E(2\varepsilon) \tag{3.42}$$

for all $\tilde{g} \in \mathscr{W}$ such that $\tilde{\delta}_\square(\tilde{g}, \tilde{h}) \leqslant 2\varepsilon$. Without loss of generality we may assume that $d_\square(g, h) \leqslant 2\varepsilon$. Following a similar line of reasoning as above, we get

$$-I_{1,t}(u^\sigma, g) \leqslant -I_{1,t}(u^\sigma, h) + E(2\varepsilon) \qquad \forall \, \sigma \in \mathcal{M}, \tag{3.43}$$

which implies (3.42). $\qquad \square$

3.3. Lower bound.

To establish the lower bound, it suffices to prove that

$$\lim_{\varepsilon \downarrow 0} \lim_{\eta \downarrow 0} \liminf_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}\big(\tilde{f}_{n,t} \in \tilde{\mathbb{B}}_\square(\tilde{h}, \varepsilon) \mid \tilde{f}_{n,0} = \tilde{u}_n^\eta\big) \geqslant -\tilde{I}_{1,t}(\tilde{u}, \tilde{h}) \qquad \forall \, \tilde{h} \in \tilde{\mathscr{W}}. \tag{3.44}$$

For any $\beta > 0$ there exists a $\phi(\beta) \in \mathcal{M}$ such that

$$\tilde{I}_{1,t}(\tilde{u}, \tilde{h}) \geqslant I_{1,t}(u, h^\phi) - \beta. \tag{3.45}$$

Because $\mathbb{B}_\square(h^\phi, \varepsilon) \subseteq \mathbb{B}_\square(\tilde{h}, \varepsilon)$ for any $\phi \in \mathcal{M}$, picking $h = h^{\phi(\beta)}$ and letting $\beta \downarrow 0$, we see that (3.44) follows once we show that

$$\lim_{\varepsilon \downarrow 0} \lim_{\eta \downarrow 0} \liminf_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}\big(f_{n,t} \in \mathbb{B}_\square(h, \varepsilon) \mid f_{n,0} = u_n^\eta\big) \geqslant -I_{1,t}(u, h) \qquad \forall \, u, h \in \mathscr{W}. \tag{3.46}$$

The proof comes in 5 steps and is constructed around a series of technical lemmas (Lemmas 3.5–3.8 below).

1. To prove (3.46), we first introduce some notation. As before, we work with block graphons. For $k \in \mathbb{N}$ and $i, j \in [k]$, let $B_{i,j}^{(k)} := [\frac{i-1}{k}, \frac{i}{k}) \times [\frac{j-1}{k}, \frac{j}{k})$. For $g \in \mathcal{W}$, we let $\widehat{g}_k \in \mathcal{W}^{(k)}$ be defined at the bottom-left corner points of $B_{i,j}^{(k)}$ by

$$\widehat{g}_k\left(\tfrac{i-1}{k}, \tfrac{j-1}{k}\right) := k^2 \int_{B_{i,j}^{(k)}} \mathrm{d}x \, \mathrm{d}y \, g(x, y), \tag{3.47}$$

and for $(x, y) \in B_{i,j}^{(k)}$ as

$$\widehat{g}_k(x, y) := \widehat{g}_k\left(\tfrac{i-1}{k}, \tfrac{j-1}{k}\right). \tag{3.48}$$

We settle (3.46) by using a Cramér-transform-type argument, i.e., we rely on a particular change of measure. Concretely, for $z, x \in [0, 1]$, let

$$\tau_t(z, x) := \underset{v \in \mathbb{R}}{\arg\max} \left[vx - J_{t,v}(z) \right], \tag{3.49}$$

where $J_{t,v}(z)$ is the function defined in (1.17). The idea is to use τ_t to describe the probability that particular edges are active at time t when $f_{n,t}$ is conditioned to be close to h. To that end, abbreviate $\theta_{k,t}(x, y) := \exp(\tau_t(\widehat{u}_k(x, y), \widehat{h}_k(x, y)))$, let

$$\begin{aligned} \alpha_{k,t}(x, y) &:= \frac{p_{11,t} \, \theta_{k,t}(x, y)}{1 - p_{11,t} + p_{11,t} \, \theta_{k,t}(x, y)}, \\ \beta_{k,t}(x, y) &:= \frac{p_{01,t} \, \theta_{k,t}(x, y)}{1 - p_{01,t} + p_{01,t} \, \theta_{k,t}(x, y)}, \end{aligned} \tag{3.50}$$

and for $\eta > 0$ put

$$q_{k,n}^{\eta}(x, y) := u_n^{\eta}(x, y)\alpha_{k,t}(x, y) + [1 - u_n^{\eta}(x, y)]\beta_{k,t}(x, y). \tag{3.51}$$

For $i, j \in [k]$, let

$$q_k^{\eta}(i, j, n) := \frac{1}{\mathrm{Leb}(B_{i,j}^{(n)})} \int_{B_{i,j}^{(n)}} \mathrm{d}x \, \mathrm{d}y \, q_{n,k}^{\eta}(x, y). \tag{3.52}$$

We can informally interpret $q_k^{\eta}(i, j, n)$ as an estimate of the probability that edge (i, j) is active at time t given that $f_{n,t}$ is close to h. Note that $q_k^{\eta}(i, j, n)$ depends not only on whether edge (i, j) is initially active (dependence on n), but also on the proportion of other 'nearby' edges that are initially active (dependence on k).

2. In the following lemma we show that $q_{k,n}^{\eta} \in \mathcal{W}$ converges to h in an approropriate limit.

Lemma 3.5. *For every $t > 0$,*

$$\lim_{k \to \infty} \lim_{\eta \downarrow 0} \lim_{n \to \infty} d_{\square}(h, q_{k,n}^{\eta}) = 0. \tag{3.53}$$

Proof. First note that

$$d_{\square}(q_{k,n}^{\eta}, h) \leqslant d_{\square}(q_{k,n}^{\eta}, \widehat{h}_k) + d_{\square}(h, \widehat{h}_k). \tag{3.54}$$

We next analyse $d_{\square}(q_{k,n}^{\eta}, \widehat{h}_k)$. From (3.49), we find the first-order condition

$$\widehat{h}_k(x, y) = \widehat{u}_k(x, y) \, \alpha_{k,t}(x, y) + [1 - \widehat{u}_k(x, y)] \, \beta_{k,t}(x, y) \qquad \forall \, x, y \in [0, 1]. \tag{3.55}$$

Using (3.55), we obtain (rewrite the supremum over subsets $S, T \subseteq [0,1]$ in (1.2) as a supremum over functions $a, b \colon [0,1] \to [0,1]$)

$$d_\square(q_{k,n}^\eta, \widehat{h}_k)$$

$$= \sup_{a,b} \int_{[0,1]^2} dx\, dy\, a(x) b(y) \left[\alpha_{k,t}(x,y)[u_n^\eta(x,y) - \widehat{u}_k(x,y)] + \beta_{k,t}(x,y)[\widehat{u}_k(x,y) - u_n^\eta(x,y)] \right]$$

$$\leqslant \sup_{a,b} \int_{[0,1]^2} dx\, dy\, a(x) b(y) \left[\alpha_{k,t}(x,y)[u_n^\eta(x,y) - u(x,y)] + \beta_{k,t}(x,y)[u(x,y) - u_n^\eta(x,y)] \right]$$

$$+ \int_{[0,1]^2} dx\, dy\, |u(x,y) - \widehat{u}_k(x,y)|$$

$$\leqslant 2k^2 d_\square(u, u_n^\eta) + \int_{[0,1]^2} dx\, dy\, |u(x,y) - \widehat{u}_k(x,y)|,$$

(3.56)

where we note that $\alpha_{k,t}(x,y), \beta_{k,t}(x,y) \in [0,1]$, and use the triangle inequality in combination with the fact that the pair $(\widehat{u}_k, \widehat{h}_k)$ can take at most k^2 different values (corresponding to their values on the interior of $B_{i,j}^{(k)}$). The claim now follows from (3.54) and (3.56), because $d_\square(u, u_n^\eta) \downarrow 0$, $\widehat{h}_k \to h$ in L^2 and $\widehat{u}_k \to u$ in L^2 (see [6, Proposition 2.6]). □

3. The next step is to construct a random variable H_n on simple graphs with n vertices by declaring that, for every $i, j \in [n]$ with $i < j$, vertices i and j are connected by an edge with probability $q_{n,k}^\eta(i,j)$ defined in (3.52). Let $\mathbb{P}_{n,k,h}^\eta$ denote the law of H_n. Note that if $f \in \mathscr{W}_n$, then

$$\mathbb{P}_{n,k,h}^\eta(\{f\}) = \prod_{1 \leqslant i < j \leqslant n} q_k^\eta(i,j,n)^{f_{ij}} [1 - q_k^\eta(i,j,n)]^{1-f_{ij}},$$

$$\mathbb{P}_{n,t}^\eta(\{f\}) = \prod_{1 \leqslant i < j \leqslant n} p_{11,t}^{u_n^\eta(i,j) f_{ij}} p_{01,t}^{[1-u_n^\eta(i,j)] f_{ij}}$$

$$\times (1 - p_{11,t})^{u_n^\eta(i,j)(1-f_{ij})} (1 - p_{01,t})^{[1-u_n^\eta(i,j)](1-f_{ij})},$$

(3.57)

where f_{ij} denotes the value of $f(x,y)$ on the interior of $B_{i,j}^{(n)}$.

Lemma 3.6. *For fixed $k \in \mathbb{N}$,*

$$\lim_{\eta \downarrow 0} \lim_{n \to \infty} \frac{1}{\binom{n}{2}} \int_{\mathscr{W}_n} \left(\log \frac{d\mathbb{P}_{n,k,h}^\eta}{d\mathbb{P}_{n,t}^\eta} \right) d\mathbb{P}_{n,k,h}^\eta = I_{1,t}(\widehat{u}_k, \widehat{h}_k).$$

(3.58)

Proof. Let $n_\ell = \ell k$ with $\ell \in \mathbb{N}$, and note that, for all $i, j \in [n_\ell]$,

$$B_{i,j}^{(n_\ell)} \subseteq B_{\lceil i/k \rceil, \lceil j/k \rceil}^{(k)}.$$

(3.59)

Rewriting (3.51) as a product, and applying (3.52) and (3.57), we obtain that, for any $f \in \mathscr{W}_n$,

$$\frac{\mathbb{P}_{n,k,h}^\eta(f)}{\mathbb{P}_{n,t}^\eta(f)} = \prod_{1 \leqslant i < j \leqslant n} \bar{F}_{k,i,j}(1,1)^{u_n^\eta(i,j) f_{ij}} \bar{F}_{k,i,j}(1,0)^{[1-u_n^\eta(i,j)] f_{ij}}$$

(3.60)

$$\times \bar{F}_{k,i,j}(0,1)^{u_n^\eta(i,j)(1-f_{ij})} \bar{F}_{k,i,j}(0,0)^{[1-u_n^\eta(i,j)](1-f_{ij})},$$

where, for $\ell, m \in \{0, 1\}$,

$$F_{k,x,y}(\ell, m) = \frac{(\theta_{k,t}(x,y))^\ell}{1 - p_{m1,t} + p_{m1,t}\, \theta_{k,t}(x,y)}, \quad \bar{F}_{k,i,j}(\ell, m) = \frac{(\bar{\theta}_{k,t}(i,j))^\ell}{1 - p_{m1,t} + p_{m1,t}\, \bar{\theta}_{k,t}(i,j)}, \tag{3.61}$$

with

$$\bar{\theta}_{k,t}(i,j) := \theta_{k,t}\left(\tfrac{i-1}{k}, \tfrac{j-1}{k}\right). \tag{3.62}$$

From (3.60), elementary computations yield

$$\lim_{\eta\downarrow 0} \lim_{n\to\infty} \frac{1}{\binom{n}{2}} \int_{\mathscr{W}_n} \left(\log \frac{\mathrm{d}\mathbb{P}^\eta_{n,k,h}}{\mathrm{d}\mathbb{P}^\eta_{n,t}}(f) \right) \mathrm{d}\mathbb{P}^\eta_{n,k,h}(f)$$

$$= \lim_{\eta\downarrow 0} \lim_{n\to\infty} \frac{1}{\binom{n}{2}} \sum_{1\leqslant i<j\leqslant n} \int_{\mathscr{W}} \left[u_n^\eta(i,j) f_{ij} \log \bar{F}_{k,i,j}(1,1) + [1 - u_n^\eta(i,j)] f_{ij} \log \bar{F}_{k,i,j}(1,0) \right.$$

$$\left. + u_n^\eta(i,j)(1 - f_{ij}) \log \bar{F}_{k,i,j}(0,1) + [1 - u_n^\eta(i,j)](1 - f_{ij}) \log \bar{F}_{k,i,j}(0,0) \right] \mathrm{d}\mathbb{P}^\eta_{n,k,h}(f)$$

$$= \lim_{\eta\downarrow 0} \lim_{n\to\infty} \frac{1}{\binom{n}{2}} \sum_{1\leqslant i<j\leqslant n} \left[\tau_t\big(\widehat{u}_k(i,j), \widehat{h}_k(i,j)\big) \int_{\mathscr{W}} f_{ij}\, \mathrm{d}\mathbb{P}^\eta_{n,k,h}(f) + u_n^\eta(i,j) \log \bar{F}_{k,i,j}(0,1) \right.$$

$$\left. + [1 - u_n^\eta(i,j)] \log \bar{F}_{k,i,j}(0,0) \right]$$

$$= \lim_{\eta\downarrow 0} \lim_{n\to\infty} \frac{1}{\binom{n}{2}} \sum_{1\leqslant i<j\leqslant n} \left[u_n^\eta(i,j) \left\{ \tau_t\big(\widehat{u}_k(i,j), \widehat{h}_k(i,j)\big)\, p_{11,t} \bar{F}_{k,i,j}(1,1) + \log \bar{F}_{k,i,j}(0,1) \right\} \right.$$

$$\left. + [1 - u_n^\eta(i,j)] \left\{ \tau_t\big(\widehat{u}_k(i,j), \widehat{h}_k(i,j)\big)\, p_{01,t} \bar{F}_{k,i,j}(1,0) + \log \bar{F}_{k,i,j}(0,0) \right\} \right]. \tag{3.63}$$

Using that $\widehat{u}_k(x,y)$ is constant on the interior of $B_{i,j}^{(k)}$ for all $1 \leqslant i, j \leqslant k$, in combination with the fact that $\lim_{\eta\downarrow 0} \lim_{n\to\infty} d_\square(u_n^\eta, u) = 0$, we obtain that (3.63) equals

$$\int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y \left[\widehat{u}_k(x,y) \left\{ \tau_t\big(\widehat{u}_k(x,y), \widehat{h}_k(x,y)\big)\, p_{11,t} F_{k,x,y}(1,1) + \log F_{k,x,y}(0,1) \right\} \right.$$

$$\left. + [1 - \widehat{u}_k(x,y)] \left\{ \tau_t\big(\widehat{u}_k(x,y), \widehat{h}_k(x,y)\big)\, p_{01,t} F_{k,x,y}(1,0) + \log F_{k,x,y}(0,0) \right\} \right]. \tag{3.64}$$

Applying (3.55), we see that (3.64) equals

$$\int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y \left[\tau_t(\widehat{u}_k(x,y), \widehat{h}_k(x,y)) \widehat{h}_k(x,y) + \widehat{u}_k(x,y) \log F_{k,x,y}(0,1) \right.$$

$$\left. + [1 - \widehat{u}_k(x,y)] \log F_{k,x,y}(0,0) \right] = I_{1,t}(\widehat{u}_k, \widehat{h}_k), \tag{3.65}$$

which settles the claim in (3.58) along subsequences n_ℓ of the form ℓk with $\ell \in \mathbb{N}$. Straightforward reasoning gives the same along full sequences: the resulting discrepancies correspond to sets of vanishing Lebesgue measure (see the proof of [6, Proposition 2.6] for a similar argument). $\qquad\square$

4. Two further lemmas are needed.

Lemma 3.7. *For every $t > 0$,*

$$\lim_{k \to \infty} I_{1,t}(\widehat{u}_k, \widehat{h}_k) = I_{1,t}(u, h). \tag{3.66}$$

Proof. The claim follows from (3.3) and the fact that $\widehat{u}_k \xrightarrow{L^2} u$ and $\widehat{h}_k \xrightarrow{L^2} h$ as $k \to \infty$. $\quad\square$

Lemma 3.8. *For fixed $k \in \mathbb{N}$ and $\varepsilon > 0$,*

$$\lim_{n \to \infty} \mathbb{P}^\eta_{n,k,h}(\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon)) = 1. \tag{3.67}$$

Proof. The claim follows from the same argument as in the proof of [6, Lemmas 5.6 and 5.8–5.11]. $\quad\square$

5. We are now ready to prove (3.46). For fixed $k \in \mathbb{N}$ and $\varepsilon > 0$,

$$\mathbb{P}^\eta_{n,t}(\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon))$$

$$= \int_{\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon)} \mathrm{d}\mathbb{P}^\eta_{n,t} = \int_{\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon)} \exp\left(-\log \frac{\mathrm{d}\mathbb{P}^\eta_{n,k,h}}{\mathrm{d}\mathbb{P}^\eta_{n,t}}\right) \mathrm{d}\mathbb{P}^\eta_{n,k,h} \tag{3.68}$$

$$= \mathbb{P}^\eta_{n,k,h}(\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon)) \frac{1}{\mathbb{P}^\eta_{n,k,h}(\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon))} \int_{\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon)} \exp\left(-\log \frac{\mathrm{d}\mathbb{P}^\eta_{n,k,h}}{\mathrm{d}\mathbb{P}^\eta_{n,t}}\right) \mathrm{d}\mathbb{P}^\eta_{n,k,h}.$$

Therefore, by Jensen's inequality,

$$\log \mathbb{P}^\eta_{n,t}(\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon)) \geq \log \mathbb{P}^\eta_{n,k,h}(\mathbb{B}_\square(q_{n,k}, \varepsilon))$$

$$- \frac{1}{\mathbb{P}^\eta_{n,k,h}(\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon))} \int_{\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon)} \left(\log \frac{\mathrm{d}\mathbb{P}^\eta_{n,k,h}}{\mathrm{d}\mathbb{P}^\eta_{n,t}}\right) \mathrm{d}\mathbb{P}^\eta_{n,k,h}. \tag{3.69}$$

By Lemma 3.8, $\mathbb{P}_{n,k,h}(\mathbb{B}_\square(q_{n,k}, \varepsilon)) \to 1$, which implies that

$$\lim_{\eta \downarrow 0} \liminf_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}^\eta_{n,t}(\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon)) \geq -\lim_{\eta \downarrow 0} \lim_{n \to \infty} \frac{1}{\binom{n}{2}} \int \left(\log \frac{\mathrm{d}\mathbb{P}^\eta_{n,k,h}}{\mathrm{d}\mathbb{P}^\eta_{n,t}}\right) \mathrm{d}\mathbb{P}^\eta_{n,k,h}. \tag{3.70}$$

According to Lemma 3.6, the right-hand side equals $-I_{1,t}(\widehat{u}_k, \widehat{h}_k)$. Since

$$\begin{aligned}
&\lim_{k \to \infty} \lim_{\eta \downarrow 0} \lim_{n \to \infty} d_\square(q^\eta_{n,k}, h) = 0 && \text{(by Lemma 3.5)}, \\
&\lim_{k \to \infty} I_{1,t}(\widehat{u}_k, \widehat{h}_k) = I_{1,t}(u, h) && \text{(by Lemma 3.7)},
\end{aligned} \tag{3.71}$$

if we let $k \to \infty$, then we obtain from (3.70) that

$$\lim_{\eta \downarrow 0} \liminf_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}^\eta_{n,t}(\mathbb{B}_\square(h, \eta)) \geq -I_{1,t}(u, h), \tag{3.72}$$

where we use that $\mathbb{B}_\square(q^\eta_{n,k}, \varepsilon) \supseteq \mathbb{B}_\square(h, \eta)$ for $0 < \eta < \varepsilon$ and n large enough. $\quad\square$

Collecting the results in Sections 3.1–3.3, we see that we have completed the proof of Theorem 1.2.

4. Proofs of the multi-point and sample-path LDPs

4.1. **Proof of the multi-point LDP.** The objective of this section to prove Theorem 1.3 with the help of Theorem 1.2. The structure aligns with Section 3: lower semi-continuity, lower bound, upper bound.

4.1.1. *Lower semi-continuity.* For $\tilde{h}, \tilde{g} \in \tilde{\mathscr{W}}^{(\ell)}$, let

$$\delta_\square^\ell(\tilde{h}, \tilde{g}) := \max_{i \in [\ell]} \delta_\square(\tilde{h}_i, \tilde{g}_i). \tag{4.1}$$

By Lemma 3.3, for any sequence $(\tilde{h}_n)_{n \in \mathbb{N}}$ in $\tilde{\mathscr{W}}^{|j|+1}$ such that $\lim_{n \to \infty} \delta_\square^{|j|+1}(\tilde{h}_n, \tilde{h}) = 0$,

$$\liminf_{n \to \infty} \tilde{I}_j(\tilde{h}_n) = \liminf_{n \to \infty} \sum_{i=1}^{|j|} \tilde{I}_{1, t_i - t_{i-1}}(\tilde{h}_{n,i-1}, \tilde{h}_{n,i}) \geqslant \sum_{i=1}^{|j|} \tilde{I}_{1, t_i - t_{i-1}}(\tilde{h}_{i-1}, \tilde{h}_i) = \tilde{I}_j(\tilde{h}), \quad (4.2)$$

which settles the lower semi-continuity of $\tilde{I}_j$.

4.1.2. *Lower bound.* Let $\tilde{h} = (\tilde{h}_i)_{i=0}^{|j|} \in \tilde{\mathscr{W}}^{|j|+1}$ with $\tilde{h}_0 \equiv \tilde{u}$, and let $\boldsymbol{\eta} = (\eta_i)_{i=0}^{|j|} \in (0,1)^{|j|+1}$. Define

$$\tilde{\mathbb{B}}_\square^{|j|}(\tilde{h}, \boldsymbol{\eta}) = \left\{ \tilde{g} \in \tilde{\mathscr{W}}^{|j|+1} \colon \delta_\square(\tilde{h}_i, \tilde{g}_i) \leqslant \eta_i \quad \forall i \in [|j|] \right\}. \tag{4.3}$$

For n large enough such that $\tilde{u}_n \in \tilde{\mathbb{B}}_\square(\tilde{h}_0, \eta_0)$, the Markov property yields

$$\frac{1}{\binom{n}{2}} \log(\tilde{\mu}_n \circ p_j^{-1})\big(\tilde{\mathbb{B}}_\square^{|j|}(h, \boldsymbol{\eta})\big)$$
$$\geqslant \frac{1}{\binom{n}{2}} \sum_{i=1}^{|j|} \inf_{\tilde{w} \in \tilde{\mathbb{B}}_\square(\tilde{h}_{i-1}, \eta_{i-1}) \cap \tilde{\mathscr{W}}_n} \log \mathbb{P}\Big(\tilde{f}_{n, t_i} \in \tilde{\mathbb{B}}_\square(\tilde{h}_i, \eta_i) \,\Big|\, \tilde{f}_{n, t_{i-1}} = \tilde{w}\Big). \tag{4.4}$$

Applying Theorem 1.2, we obtain

$$\lim_{\eta_i \downarrow 0} \lim_{\eta_{i-1} \downarrow 0} \liminf_{n \to \infty} \frac{1}{\binom{n}{2}} \inf_{\tilde{w} \in \tilde{\mathbb{B}}_\square(\tilde{h}_{i-1}, \eta_{i-1}) \cap \tilde{\mathscr{W}}_n} \log \mathbb{P}\Big(\tilde{f}_{n, t_i} \in \tilde{\mathbb{B}}_\square(\tilde{h}_i, \eta_i) \,\Big|\, \tilde{f}_{n, t_{i-1}} = \tilde{w}\Big)$$
$$\geqslant -\lim_{\eta_i \downarrow 0} \inf_{\tilde{x} \in \tilde{\mathbb{B}}_\square(\tilde{h}_i, \eta_i)} I_{1, t_i - t_{i-1}}(\tilde{h}_{i-1}, \tilde{x}) = -I_{1, t_i - t_{i-1}}(\tilde{h}_{i-1}, \tilde{h}_i). \tag{4.5}$$

Combining (4.4) and (4.5), we get

$$\lim_{\eta_{|j|} \downarrow 0} \cdots \lim_{\eta_0 \downarrow 0} \liminf_{n \to \infty} \frac{1}{\binom{n}{2}} \log(\tilde{\mu}_n \circ p_j^{-1})\big(\tilde{\mathbb{B}}_\square^{|j|}(\tilde{h}, \boldsymbol{\eta})\big) \geqslant -\sum_{i=1}^{|j|+1} I(\tilde{h}_{i-1}, \tilde{h}_i) = -\tilde{I}_j(\tilde{h}), \tag{4.6}$$

from which the desired lower bound follows.

4.1.3. *Upper bound.* Following similar arguments as above, we obtain

$$\lim_{\eta_{|j|} \downarrow 0} \cdots \lim_{\eta_0 \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log(\tilde{\mu}_n \circ p_j^{-1})\big(\mathbb{B}_\square^{|j|}(h, \boldsymbol{\eta})\big) \leqslant -I_j(h). \tag{4.7}$$

To achieve this, we need a 'local-to-global transference' result. But because $|j|$ is finite and $\tilde{\mathscr{W}}$ is compact, we can use the ideas in [6, Lemma 4.1]. $\qquad \square$

4.2. **Proof of the sample-path LDP.** We are now ready to prove the sample-path LDP in Theorem 1.4.

Proof. Consider a path $\tilde{h} \in \tilde{\mathscr{W}} \times [0, T]$. Applying Theorem 1.3 and the Dawson-Gärtner projective limit LDP [11, Theorem 4.6.1], we obtain a sample-path LDP in the pointwise topology (for which we use the label $\mathcal{X}$) with rate $\binom{n}{2}$ and with rate function

$$\tilde{I}_\mathcal{X}(\tilde{h}) := \sup_{0 = t_0 < t_1 < \cdots < t_k \leqslant T} \sum_{i=1}^{k} \tilde{I}_{1, t_i - t_{i-1}}(\tilde{h}_{t_{i-1}}, \tilde{h}_{t_i}), \tag{4.8}$$

where $\tilde{h} \in \tilde{\mathcal{W}} \times [0, T]$. There are two major challenges we need to overcome to establish Theorem 1.4:

(I) Prove exponential tightness to strengthen the topology (Lemma 4.1 below).
(II) Show that $\tilde{I}_{\mathcal{X}}(\tilde{h}) = \tilde{I}(\tilde{h})$ (Lemma 4.2 below).

(I) Recall the definition of D in (1.12). We say that the sequence of probability measures $(\mu_n)_{n \in \mathbb{N}}$ on D is exponentially tight when for every $\alpha < \infty$ there exists a compact set $K_\alpha \subseteq D$ such that

$$\limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mu_n(K_\alpha^c) < -\alpha. \tag{4.9}$$

Lemma 4.1. $(\tilde{\mu}_n)_{n \in \mathbb{N}}$ *is exponentially tight.*

Proof. By [12, Theorem 4.1] (with $\beta = 1$ in the notation used in [12] and n replaced by $\binom{n}{2}$), the compactness of $\tilde{\mathcal{W}}$ and the Markov property, it suffices to show that for each $\eta, \xi > 0$ there exist random variables $\gamma_n(\eta, \xi)$, satisfying

$$\mathbb{E}\left[e^{\binom{n}{2}\xi\delta_\square(\tilde{f}_{n,\eta'}, \tilde{u})} \mid f_{0,n} = u_n\right] \leq \mathbb{E}[e^{\gamma_n(\eta,\xi)}], \qquad 0 \leq \eta' \leq \eta, \; u_n \in \mathcal{W}_n, \tag{4.10}$$

such that, for each $\xi > 0$,

$$\lim_{\eta \downarrow 0} \limsup_{n \to \infty} \frac{1}{\binom{n}{2}} \log \mathbb{E}[e^{\gamma_n(\eta,\xi)}] = 0. \tag{4.11}$$

To construct $\gamma_n(\eta, \xi)$, let E denote the total number of edges that change (i.e., go from active to inactive or from inactive to active) somewhere in the time interval $[0, \eta]$. Then, given $f_{n,0} = u_n$, we have

$$\delta_\square(\tilde{f}_{n,\eta'}, \tilde{u}) \leq \int_{[0,1]^2} \mathrm{d}x \, \mathrm{d}y \, |f_{n,\eta'}(x,y) - u(x,y)| \leq E \binom{n}{2}^{-1}, \tag{4.12}$$

where the last inequality holds for all $0 \leq \eta' \leq \eta$. Next observe that, because all edges evolve independently, for any $u_n \in \mathcal{W}_n$ the random variable E given $f_{n,0} = u_n$ is stochastically dominated by $Y = \mathrm{Bin}(\binom{n}{2}, 1 - e^{-\max\{\lambda,\mu\}\eta})$. Thus, if we let $\gamma_n(\eta, \xi) = \xi Y$, then (4.10) holds and

$$\frac{1}{\binom{n}{2}} \log \mathbb{E}[e^{\gamma_n(\eta,\xi)}] = \log\left(e^{-\max\{\lambda,\mu\}\eta} + e^\xi(1 - e^{-\max\{\lambda,\mu\}\eta})\right) \downarrow 0, \qquad \eta \downarrow 0, \tag{4.13}$$

which finishes the proof. $\qquad \square$

(II) The following identity holds.

Lemma 4.2. $\tilde{I}(\tilde{h}) = \tilde{I}_{\mathcal{X}}(\tilde{h})$ *for all* $\tilde{h} \in \tilde{\mathcal{W}} \times [0, T]$.

Proof. Recall the definition of $I(h)$ in (1.26), and that $\mathcal{AC}$ is the set of functions h on $\mathcal{W} \times [0, T]$ such that $t \mapsto h_t(x, y)$ is absolutely continuous for almost all $(x, y) \in [0, 1]^2$.

1. We first establish that

$$\tilde{I}_{\mathcal{X}}(\tilde{h}) \equiv \inf_{h \sim \tilde{h}} \sup_{0 = t_0 < t_1 < \cdots < t_k \leq T} \sum_{i=1}^{k} I_{1, t_i - t_{i-1}}(h_{t_{i-1}}, h_{t_i}) \leq \tilde{I}(\tilde{h}). \tag{4.14}$$

It is enough to show that $I_\mathcal{X}(h) \leqslant I(h)$ for all $h \sim \tilde{h}$. If $h \notin \mathcal{AC}$, then $I(h) = \infty$ by definition, whereas if $h \in \mathcal{AC}$, then due the convexity of $I_{1,t}$ we obtain $I_\mathcal{X} \leqslant I(h)$ by applying Jensen's inequality.

2. We now establish the reverse inequality. Similarly as above, it is enough to show that $I_\mathcal{X}(h) \geqslant I(h)$ for all $h \sim \tilde{h}$.

a. Suppose that $h \in \mathcal{AC}$. Then $h'_t(x,y) := \frac{\partial}{\partial s} h_s(x,y)\big|_{s=t}$ exists for almost all $x, y \in [0,1]^2$. Letting $t_i - t_{i-1} = \Delta t$ for all i, we have

$$
\begin{aligned}
I_\mathcal{X}(h) &\geqslant \sum_{i=1}^{T/\Delta t} I_{1,\Delta t}(h_{(i-1)\Delta t}, h_{i\Delta t}) \\
&= \sum_{i=1}^{T/\Delta t} \int_0^1 \mathrm{d}x \int_0^x \mathrm{d}y \\
&\qquad \sup_{v \in \mathbb{R}} \Bigg[v\big(h_{(i-1)\Delta t}(x,y) + h'_{(i-1)\Delta t}(x,y)\Delta t\big) \\
&\qquad\qquad\qquad - h_{(i-1)\Delta t}(x,y) \log\left[\mu\Delta t + \mathrm{e}^v(1 - \mu\Delta t)\right] \\
&\qquad\qquad\qquad - [1 - h_{(i-1)\Delta t}(x,y)]\log(1 - \lambda + \mathrm{e}^v \lambda\Delta t) + o(\Delta t) \Bigg] \\
&= \sum_{i=1}^{T/\Delta t} \int_0^1 \mathrm{d}x \int_0^x \mathrm{d}y \sup_{v \in \mathbb{R}} \Bigg[vh'_{(i-1)\Delta t}(x,y) - h_{(i-1)\Delta t}(x,y)\mu(\mathrm{e}^{-v} - 1) \\
&\qquad\qquad\qquad - [1 - h_{(i-1)\Delta t}(x,y)]\lambda(\mathrm{e}^v - 1) + o(1) \Bigg]\Delta t \\
&\to \int_0^T \mathrm{d}t \int_0^1 \mathrm{d}x \int_0^x \mathrm{d}y \sup_{v \in \mathbb{R}} \Bigg[vh'_t(x,y) - h_t(x,y)\mu(\mathrm{e}^{-v} - 1) - [1 - h_t(x,y)]\lambda(\mathrm{e}^v - 1) \Bigg] \\
&=: I(h), \qquad \Delta t \downarrow 0.
\end{aligned}
$$
(4.15)

This implies that if $h \in \mathcal{AC}$, then $I_\mathcal{X}(h) \geqslant I(h)$.

b. It is now enough to show that if $h \sim \tilde{h}$ and $h \notin \mathcal{AC}$, then there exists a $\widehat{h} \sim \tilde{h}$ such that

$$
I(\widehat{h}) = I_\mathcal{X}(h)
$$
(4.16)

(in many cases we can take $h = \widehat{h}$). The following argument is sketchy, but follows standard reasonings.

Suppose that $h \notin \mathcal{AC}$. By definition this means that the set of $(x,y) \in [0,1]^2$ for which there exist $\delta_{x,y} > 0$ and $\{s_{1,x,y}^k < t_{1,x,y}^k \leqslant \cdots \leqslant s_{\ell_k,x,y}^k < t_{\ell_k,x,y}^k\}$ with

$$
\lim_{k \to \infty} \sum_{i=1}^{\ell_k} (t_{i,x,y}^k - s_{i,x,y}^k) = 0, \qquad \lim_{k \to \infty} \sum_{i=1}^{\ell_k} |h_{t_{i,x,y}^k}(x,y) - h_{s_{i,x,y}^k}(x,y)| \geqslant \delta_{x,y},
$$
(4.17)

has positive Lebesgue measure. We distinguish between a number of cases.

• Suppose that there exist $\{s_1^k < t_1^k \leqslant \cdots \leqslant s_{\ell_k}^k < t_{\ell_k}^k\}$ independent of (x, y) such that the set of $(x, y) \in [0, 1]^2$ for which

$$\lim_{k \to \infty} \sum_{i=1}^{\ell_k} (t_i^k - s_i^k) = 0, \qquad \lim_{k \to \infty} \sum_{i=1}^{\ell_k} |h_{t_i^k} - h_{s_{\ell_i}^k}| \geqslant \delta \text{ for some } \delta > 0, \qquad (4.18)$$

has positive Lebesgue measure. Then, following arguments similar to those in the proof of [11, Lemma 5.1.6], we get $I_{\mathcal{X}}(h) = I(h) = \infty$.

• Suppose that no sequence satisfying (4.18) exists. Roughly speaking, the paths $t \mapsto h_t(x, y)$ that are not absolutely continuous fall into two categories: those that contain 'steps' and those that contain 'holes'.

(i) We say that a path $t \mapsto h_t(x, y)$ contains a step when there exists a $t \in [0, T]$ such that if $t_k \uparrow t$ and $t_k' \downarrow t$, then $h_{t_k}(x, y) \to c$ and $h_{t_k'}(x, y) \to d$ with $c \neq d$.
(ii) We say that a path $t \mapsto h_t(x, y)$ contains a hole when there exists a $t \in [0, T]$ such that if $t_k \to t$ (with $t_k \neq t$ for all k), then $h_{t_k} \to c \neq h_t$.
(If the above limits do not exists, then the arguments below can be easily adapted.)

▸ Suppose that the set of (x, y) such that $t \mapsto h_t(x, y)$ contains a step of size $\gamma > 0$ has positive Lebesgue measure $\beta > 0$. Then

$$I_{\mathcal{X}}(h) \geqslant \sum_{i=1}^{T/\Delta t} I_{1, \Delta t}(h_{(i-1)\Delta t}, h_{i \Delta t}) \geqslant \beta \gamma \log \frac{\gamma}{1 - e^{-\max\{\lambda, \mu\} \Delta t}} \to \infty, \qquad \Delta t \downarrow 0. \qquad (4.19)$$

where the last inequality follows from a similar reasoning as in the proof of Lemma 4.1. This implies $I_{\mathcal{X}}(h) = I(h) = \infty$.

▸ Suppose that the set of (x, y) such that $t \mapsto h_t(x, y)$ contains a hole has positive Lebesgue measure.

(i) We say that $t \mapsto h_t(x, y)$ has a hole at time t if $h_{t_k}(x, y) \to c \neq h_t(x, y)$ for any $t_k \to t$ (with $t_k \neq t$ for all k).
(ii) We say that $t \mapsto \widehat{h}_t(x, y)$ has this hole filled in if $\widehat{h}_t(x, y) = c$ and $\widehat{h}_t(x, y) = h_t(x, y)$ otherwise.

Construct $\widehat{h} \in \mathscr{W} \times [0, T]$ from h by filling in all the holes. Since there exists no sequence $\{s_{1, x, y}^k < t_{1, x, y}^k \leqslant \cdots \leqslant s_{\ell_k, x, y}^k < t_{\ell_k, x, y}^k\}$ satisfying (4.18), a positive Lebesgue measure of holes cannot occur simultaneously. Thus, $\widehat{h}_t(x, y) = h_t(x, y)$ almost everywhere, which implies that $I_{\mathcal{X}}(h) = I_{\mathcal{X}}(\widehat{h})$. In addition, because $\mathscr{W}$ was constructed by taking the quotient with respect to almost sure equivalence, we also have $\widehat{h} \sim \tilde{h}$. The fact that $I(\widehat{h}) = I_{\mathcal{X}}(\widehat{h})$ now follows from the arguments above. □

Lemmas 4.1–4.2 in combination with Theorem 1.3 complete the proof of Theorem 1.4. □

5. Proofs: Applications

5.1. **Application 1.** The following lemma is the time-varying equivalent of [7, Theorem 4.1 and Proposition 4.2] (see also [19, Theorem 2.7]). Let

$$\phi_T(H, r) := \inf \left\{ \tilde{I}_{1,T}(u, h) : h \in \mathscr{W}, t(H, h) \geqslant r \right\}, \qquad (5.1)$$

let F^+ be the set of minimisers, and let $\tilde{F}^+$ be the image of F^+ in $\tilde{\mathscr{W}}$.

Lemma 5.1. *Fix a constant graphon u. Let H be a d-regular graph for some $d \in \mathbb{N}\backslash\{1\}$. Suppose that $\lim_{n\to\infty} \delta_\square(\tilde{u}_n, \tilde{u}) = 0$ and $u\,p_{11,T} + (1-u)p_{01,T} < r < 1$. Then*

$$\lim_{n\to\infty} \frac{1}{\binom{n}{2}} \log \mathbb{P}\big(t(H, \tilde{f}_{n,T}) \geqslant r \mid \tilde{f}_{n,0} = \tilde{u}_n\big) = \phi_T(H, r). \tag{5.2}$$

Moreover, $\tilde{F}^+$ is non-empty and compact and, for each $\varepsilon > 0$, there exists a positive constant $C(H, r, \lambda, \mu, T, u, \varepsilon)$ such that

$$\mathbb{P}\Big(\delta_\square(\tilde{f}_{n,T}, \tilde{F}^+) \geqslant \varepsilon \;\Big|\; t(H, \tilde{f}) \geqslant r, \tilde{f}_{n,0} = \tilde{u}_n\Big) \leqslant \mathrm{e}^{-Cn^2}, \qquad n \in \mathbb{N}. \tag{5.3}$$

Proof. Observe that $x \mapsto I_{1,T}(u, x)$ is uniformly continuous (by Lemma 3.1), $h \mapsto I_{1,T}(u, h)$ is constant under measure-preserving bijections when u is a constant graphon, and $h \mapsto I_{1,T}(u, h) = \tilde{I}_{1,T}(\tilde{u}, \tilde{h})$ is lower semi-continuous (by Lemma 3.3). Therefore the claim follows via the same line of argument used to prove [6, Theorems 6.1–6.2], where we apply Theorem 1.3 in place of [6, Theorem 5.2]. $\qquad\square$

The proof of Theorem 2.1 borrows various elements from that of [19, Theorem 1.1].

Proof of Theorem 2.1. We distinguish between the two cases.

(i) Suppose that $(r^d, I_{1,T}(u, r))$ lies on the convex minorant of $x \mapsto I_{1,T}(u, x^{1/d})$. Applying the generalised version of Hölder's inequality derived in [13], we obtain that, for any $f \in \mathscr{W}$,

$$t(H, f) = \int_{[0,1]^3} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\, f(x, y)f(y, z)f(z, x) \leqslant \left(\int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\, f^d(x, y)\right)^{e(H)/d} = \|f\|_d^{e(H)}. \tag{5.4}$$

Abbreviate $\psi(x) := I_{1,T}(u, x^{1/d})$, and let $\widehat{\psi}$ be the convex minorant of ψ. Then by Jensen's inequality we have

$$I_{1,T}(u, f) = \int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\, \psi(f(x, y)^d) \geqslant \int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\, \widehat{\psi}(f(x, y)^d)$$

$$\geqslant \widehat{\psi}\left(\int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\, f(x, y)^d\right) = \widehat{\psi}(\|f\|_d^d). \tag{5.5}$$

Consequently, by (5.4), if $t(H, f) \geqslant r^{e(H)}$, then $\|f\|_d^d \geqslant r^d$, while, by (5.5), if $\|f\|_d^d \geqslant r^d$, then $I_{1,T}(u, f) \geqslant \widehat{\psi}(r^d) = I_{1,T}(u, r)$, where the last equality follows from the condition that $(r^d, I_{1,T}(u, r))$ lies on the convex minorant of ψ. Because ψ is strictly increasing on the interval $(u\,p_{11,T} + (1-u)\,p_{01,T}, 1]$ and $\widehat{\psi}$ is not linear in any neighbourhood of r^d, equality can occur if and only if $f \equiv r$. Thus, $\tilde{F}^+ = \{\tilde{r}\}$. The claim now follows from Lemma 5.1.

(ii) Suppose that $(r^d, I_{1,T}(u, r))$ does not lie on the convex minorant of $x \mapsto I_{1,T}(u, x^{1/d})$. Then there necessarily exist $0 \leqslant r_1 < r < r_2 \leqslant 1$ such that the point $(r^2, I_{1,T}(u, r))$ lies strictly above the line segment joining $(r_1^d, I_{1,t}(u, r_1))$ and $(r_2^d, I_{1,T}(u, r_2))$. Following the method set out in [19, Lemma 3.4], we can use this fact to construct a graphon r_ε with $I_{1,T}(u, r_\varepsilon) < I_{1,T}(u, r)$. (We refer the reader to [19] for the specific details of this construction.) Thus, again using the fact that ψ is strictly increasing on the interval

$[u\,p_{11,T} + (1-u)\,p_{01,T}, 1]$, we conclude that $\tilde{F}^+$ contains no constant graphons. Let $\tilde{C} \subseteq \mathscr{W}$ be the set of constant graphons. Since $\tilde{F}^+$ and $\tilde{C}$ are disjoint and compact, we have $\delta_\square(\tilde{F}^+, \tilde{C}) > 0$. The result now follows by applying Lemma 5.1 with $\varepsilon = \frac{1}{2}\delta_\square(\tilde{F}^+, \tilde{C})$. $\qquad\square$

Proof of Proposition 2.2. We derive an explicit expression for $I_{1,T}$. Recall from (1.17)–(1.18) that

$$I_{1,T}(u, r) = \sup_{v \in \mathbb{R}}[vr - J_{T,v}(u)] \tag{5.6}$$

with

$$J_{T,v}(u) = u\log(1 - p_{11,T} + \mathrm{e}^v p_{11,T}) + (1-u)\log(1 - p_{01,T} + \mathrm{e}^v p_{01,T}). \tag{5.7}$$

Due to the convexity of $v \mapsto vr - J_{T,v}(u)$, we obtain the maximiser in the right-hand side of (5.6) by setting the partial derivative with respect to v equal to 0. This yields

$$0 = r - u\frac{p_{11,T}\mathrm{e}^v}{1 - p_{11,T} + \mathrm{e}^v p_{11,T}} - (1-u)\frac{\mathrm{e}^v p_{01,T}}{1 - p_{01,T} + \mathrm{e}^v p_{01,T}}, \tag{5.8}$$

which implies that

$$\begin{aligned}
0 = {} & \mathrm{e}^{2v}(1-r)p_{11,T}p_{01,T} \\
& + \mathrm{e}^v\big[(u-r)p_{11,T}(1-p_{01,T}) + (1-u-r)p_{01,T}(1-p_{11,T})\big] \\
& - r(1-p_{11,T})(1-p_{01,T}).
\end{aligned} \tag{5.9}$$

We obtain $I_{1,T}(u, r)$ by substituting into (5.7)

$$\mathrm{e}^v = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \tag{5.10}$$

where a, b, c can be read from the three lines in (5.9). It is easily verified that if $r > u$ and $T \downarrow 0$, then $a = (1-r)\lambda T + O(T^2)$, $b = u - r + O(T)$, and $c = -r\mu T + O(T^2)$, so that

$$\mathrm{e}^v = \frac{u - r}{(1-r)\lambda T} + O(1). \tag{5.11}$$

Setting $x^{1/d} = r$, we therefore have

$$I_{1,T}(u, x^{1/d}) = (x^{1/d} - u)\log(1/T) + O(1). \tag{5.12}$$

Because $x \mapsto x^{1/d}$ is concave, for T sufficiently small the point $r^d \mapsto I_{1,T}(u, r)$ cannot lie on the convex minorant of $x \mapsto I_{1,T}(u, x^{1/d})$. The fact that, for such T, $G_n(T)$ is SB now follows from Theorem 2.1. $\qquad\square$

Proof of Proposition 2.3. If $u = 0$, then

$$I_{1,T}(0, x^{1/2}) = x^{1/2}\log\frac{x^{1/2}}{p_{01,T}} + (1 - x^{1/2})\log\frac{1 - x^{1/2}}{1 - p_{01,T}}. \tag{5.13}$$

Calling $F(x)$ all terms that do not depend on T, we compute

$$\begin{aligned}
I_{1,T}(0, x^{1/2}) &= -x^{1/2}\log p_{01,T} - (1 - x^{1/2})\log(1 - p_{01,T}) + F(x), \\
\frac{\partial I_{1,T}(0, x^{1/2})}{\partial x} &= -\tfrac{1}{2}x^{-1/2}\log p_{01,T} + \tfrac{1}{2}x^{-1/2}\log(1 - p_{01,T}) + F'(x), \\
\frac{\partial^2 I_{1,T}(0, x^{1/2})}{\partial x^2} &= \tfrac{1}{4}x^{-3/2}\log p_{01,T} - \tfrac{1}{4}x^{-3/2}\log(1 - p_{01,T}) + F''(x).
\end{aligned} \tag{5.14}$$

The last line (5.14) is an increasing function of $p_{01,T}$, which itself is an increasing function of T (recall (1.10)). Thus, $r^d \mapsto I_{1,T}(0,r)$ lies on the convex minorant of $x \mapsto I_{1,T}(0, x^{1/d})$, then the same is true for all $T' > T$. Theorem 2.1 now yields the desired result, i.e., when $G_n(T)$ is SB also $G_n(T')$ is SB. The same argument applies when $u = 1$, but in this case $p_{11,T}$ is a decreasing function of T. $\qquad\square$

5.2. **Application 2.**

Proof of Theorem 2.4. Recall that $\tilde{H}$ is the set of all paths in $\mathscr{W} \times [0,T]$ that start at $\tilde{u}$ and end at $\tilde{r}$. Since $\tilde{H}$ is not compact, we first demonstrate that we can restrict our search for elements of $\tilde{H}^*$ to a compact set. To do this, we note that, by Lemma 2.7, for any u, r in the equivalence classes $\tilde{u}, \tilde{r}$,

$$\begin{aligned} I(h_{u \to r}^*) = I_{1,T}(u,r) &\leqslant \max\{I_{1,T}(0,1), I_{1,T}(1,0)\} \\ &= \max\left\{\log \frac{1}{p_{01,T}}, \log \frac{1}{1-p_{11,T}}\right\} =: K < \infty. \end{aligned} \tag{5.15}$$

Thus, no paths with rate strictly greater than K can be an element of $\tilde{H}^*$. Let

$$\tilde{H}_\eta^{(K)} := \left\{\tilde{h} \in \mathscr{W} \times [0,T] : \tilde{I}(\tilde{h}) \leqslant K, \; \tilde{h}_0 = \tilde{u}, \; \delta_\square(\tilde{h}(T), \tilde{r}) \leqslant \eta\right\}. \tag{5.16}$$

The fact that $\tilde{H}_\eta^{(K)}$ is compact follows from a similar line of reasoning as the one used in the proof of Lemma 4.1. Since $\tilde{H}^* \subseteq \tilde{H}_0^{(K)}$, $\tilde{H}_0^{(K)}$ is compact, and $\tilde{I}$ is lower semi-continuous, $\tilde{I}$ must attain its minimum on $\tilde{H}_0^{(K)}$. Thus, $\tilde{H}^*$ is non-empty. By the lower semi-continuity of $\tilde{I}$, $\tilde{H}^*$ is also closed (and hence compact).

Fix $\varepsilon > 0$ and let

$$\tilde{H}_{\eta,\varepsilon}^{\geqslant} := \{\tilde{h} \in \tilde{H}_\eta^{(K)} : \; \delta_\square^\infty(\tilde{h}, \tilde{H}^*) \geqslant \varepsilon\}. \tag{5.17}$$

Then, for the same reasons as above, $\tilde{H}_{\eta,\varepsilon}^{\geqslant}$ is compact for all $\eta, \varepsilon \geqslant 0$. Define

$$I_1 := \inf_{\tilde{h} \in \tilde{H}} I(\tilde{h}), \qquad I_2 := \inf_{\tilde{h} \in \tilde{H}_{0,\varepsilon}^{\geqslant}} I(\tilde{h}), \tag{5.18}$$

By Theorem 1.4, we have

$$\begin{aligned} \lim_{\eta \downarrow 0} \limsup_{n \to \infty} &\frac{1}{\binom{n}{2}} \log \mathbb{P}^{u_n}\left(\delta_\square^\infty(\tilde{f}_n, \tilde{H}^*) \geqslant \varepsilon \;\Big|\; \tilde{f}_{n,t} \in \tilde{H}_\eta^{(K)}\right) \\ &\leqslant \lim_{\eta \downarrow 0}\left[\inf_{\tilde{h} \in \tilde{H}_\eta^{(K)}} \tilde{I}(\tilde{h}) - \inf_{\tilde{h} \in \tilde{H}_{\eta,\varepsilon}^{\geqslant}} I(\tilde{h})\right] = I_1 - I_2. \end{aligned} \tag{5.19}$$

The proof is complete once we are able show that $I_1 < I_2$. Now, clearly, $I_1 \leqslant I_2$. If $I_1 = I_2$, then the compactness of $\tilde{H}_{\eta,\varepsilon}^{\geqslant}$ implies that there exists a $\tilde{h} \in \tilde{H}_{\eta,\varepsilon}^{\geqslant}$ satisfying $\tilde{I}(\tilde{h}) = I_2$. However, this means that $\tilde{h} \in \tilde{H}^*$ and hence $\tilde{H}_{0,\varepsilon}^{\geqslant} \cap \tilde{H}^* \neq \varnothing$, which is a contradiction. $\quad\square$

Proof of Lemma 2.5. We first demonstrate that $f_{u \to r}^* \in [0,1] \times [0,T]$ is in the set of minimisers of I. By the contraction principle, it is enough to show that $I(f_{u \to r}^*) = I_{1,T}(u,r)$. Applying Theorem 1.3 and the contraction principle, we have

$$I_{1,T}(u,r) = \min_{s \in [0,1]} \left[I_{1,t}(u,s) - I_{1,T-t}(s,r)\right] = I_{1,t}(u, f_{u \to r}^*(t)) + I_{1,T-t}(f_{u \to r}^*(t), r). \tag{5.20}$$

Consequently,

$$I_{1,T}(u,r) = \sup_{k \in \mathbb{N}} \sup_{0=t_0 < t_1 \cdots < t_k = T} \sum_{i=1}^{k} I_{1,t_i-t_{i-1}}\big(f^*_{u \to r}(t_{i-1}), f^*_{u \to r}(t_i)\big) = I(f^*_{u \to r}). \tag{5.21}$$

We next prove that $f^*_{u \to r}$ is the unique minimiser of $I(f)$ conditional on $f(0) = u$ and $f(T) = r$. Because $I(f^*_{u \to r}) = I_{1,T}(u,r)$, it suffices to establish that, for any $u, r \in [0,1]$, $t < T$ and $s \neq f^*_{u \to r}(t)$,

$$I_{1,t}(u,s) + I_{1,T-t}(s,r) > I_{1,T}(u,r). \tag{5.22}$$

To establish (5.22), note that from the definition of $f^*_{u \to r}(t)$ and the fact that both $x \mapsto I_{1,t}(u,x)$ and $x \mapsto I_{1,T-t}(x,r)$ are continuously differentiable we get

$$\frac{\partial I_{1,t}(u,x)}{\partial x}\bigg|_{x=f^*_{u \to r}(t)} = -\frac{\partial I_{1,T-t}(x,r)}{\partial x}\bigg|_{x=f^*_{u \to r}(t)}. \tag{5.23}$$

Combine this with the fact that, for any $x \in [0,1]$, $u \mapsto I_{1,T}(u,x)$ is convex and $r \mapsto I_{1,T}(x,r)$ is strictly convex (see Lemma 3.2), to get

$$I_{1,t}(u,s) + I_{1,T-t}(s,r) > I_{1,t}(u, f^*_{u \to r}(t)) + (s - f^*_{u \to r}(t)) \frac{\partial I_{1,t}(u,x)}{\partial x}\bigg|_{x=f^*_{u \to r}(t)}$$

$$+ I_{1,T-t}(f^*_{u \to r}(t), r) + (s - f^*_{u \to r}(t)) \frac{\partial I_{1,T-t}(x,r)}{\partial x}\bigg|_{x=f^*_{u \to r}(t)}$$

$$= I_{1,t}(u, f^*_{u \to r}(t)) + I_{1,T-t}(f^*_{u \to r}(t), r) = I_{1,T}(u,r), \tag{5.24}$$

from which we conclude that indeed $f^*_{u \to r}$ is the unique minimiser of $I(f)$. $\qquad\square$

The next lemma, whose proof is standard and is omitted, can be used to compute $f^*_{u \to r}$ (see [20] for a related method). Let $\tau_t(u,r) = v$ be the unique solution of

$$0 = r - u \frac{p_{11,t}e^v}{1 - p_{11,t} + e^v p_{11,t}} - (1-u) \frac{e^v p_{01,t}}{1 - p_{01,t} + e^v p_{01,t}}, \tag{5.25}$$

which amounts to solving a quadratic equation (recall (5.10)).

Lemma 5.2. $f^*_{u \to r}(t) = s$ is the unique solution of

$$e^{\tau_t(u,s)}\big(1 - p_{01,T-t} + e^{\tau_{T-t}(s,r)} p_{01,T-t}\big) = 1 - p_{11,T-t} + e^{\tau(s,r)} p_{11,T-t}. \tag{5.26}$$

Proof of Lemma 2.7. Let

$$(h^*_{u \to r})'(x,y,t) = \frac{\partial h^*_{u \to r}(x,y,t)}{\partial t}. \tag{5.27}$$

We have

$$\int_0^T dt \int_{[0,1]^2} dx\, dy\, \mathcal{L}\big(h^*_{u \to r}(x,y,t), (h^*_{u \to r})'(x,y,t)\big)$$

$$= \int_{[0,1]^2} dx\, dy \int_0^T dt\, \mathcal{L}\big(h^*_{u \to r}(x,y,t), (h^*_{u \to r})'(x,y,t)\big) = I_{1,T}(u,r), \tag{5.28}$$

where we apply Lemma 2.5 to obtain the last equality. By the contraction principle, $h^*_{u \to r}$ is in the set of minimisers of $I(h)$ subject to the conditions $h_0 = u$ and $h_T = r$. To show that

$h^*_{u \to r}$ is the unique minimiser, note that if $h_0 = u$, $h_T = r$ and $h \neq h^*_{u \to r}$, then, by (2.14) and Lemma 2.5, there exist $\varepsilon, \beta > 0$ such that

$$\text{Leb}\left\{(x, y) \in [0,1]^2 \colon I(h(x,y)) \geq I_{1,T}(u(x,y), r(x,y)) + \varepsilon\right\} > \beta. \qquad (5.29)$$

Consequently, $I(h) \geq I(h^*_{u \to r}) + \varepsilon\beta$, which implies that indeed $h^*_{u \to r}$ is unique. $\qquad \square$

Proof of Theorem 2.8. Suppose that $u \in \mathscr{W}^{(I)}$ and $r \in \mathscr{W}^{(J)}$ for some $I, J \in \mathbb{N}$. Let $0 = a_0 < a_1 < \cdots < a_I = 1$ and $0 = b_0 < b_1 < \cdots < b_J = 1$ be their block end points, and $A_i := (a_{i-1}, a_i]$ and $B_i := (b_{i-1}, b_i]$ their block intervals. Recall the function $\alpha \colon \mathscr{M} \mapsto [0,1]^{I \times J}$ defined in (2.19) and the (compact) set V defined in (2.20). For any $\sigma, \phi \in \mathscr{M}$ with $\alpha(\sigma) = \alpha(\phi)$ we have $I_{1,T}(u, r^\phi) = I_{1,T}(u, r^\sigma)$, which implies

$$\inf_{\sigma \in \mathscr{M}} I_{1,t}(u, r^\sigma) = \min_{\boldsymbol{v} \in V} I_{1,T}(u, r^{\sigma_{\boldsymbol{v}}}), \qquad (5.30)$$

where $\sigma_{\boldsymbol{v}}$ is any element of $\mathscr{M}$ with $\alpha(\sigma_{\boldsymbol{v}}) = \boldsymbol{v} \in V$. Because $\boldsymbol{v} \mapsto I_{1,T}(u, r^{\sigma_{\boldsymbol{v}}})$ is continuous and V is compact, the set of minimisers $V^* \subseteq V$ of (5.30) is also compact. Suppose that $\tilde{h} \in \tilde{H}^*$. Then, by Lemma 2.7 and the compactness of V^*, there exist $\boldsymbol{v}^* \in V^*$ and a sequence $(h^{[i]})_{i \in \mathbb{N}}$ of representatives of $\tilde{h}$ such that

$$\lim_{i \to \infty} I(h^{[i]}) = I_{1,T}(u, r^{\sigma_{\boldsymbol{v}^*}}), \qquad \lim_{i \to \infty} \alpha(\sigma^{[i]}) = \boldsymbol{v}^*, \qquad (5.31)$$

where $\sigma^{[i]}$ is any element of $\mathscr{M}$ with $h_T^{[i]} = r^{\sigma^{[i]}}$.

Suppose that $\tilde{h} \notin \tilde{F}^*$. Due to the compactness of V, also $\tilde{F}^*$ is compact. Thus, there exist $\varepsilon > 0$ and $t \in (0, T)$ such that

$$\delta_\square\left(\tilde{h}_t, \tilde{h}^*_{u \to r^{\sigma_{\boldsymbol{v}^*}}}(\cdot, \cdot, t)\right) \geq \varepsilon. \qquad (5.32)$$

By (5.31),

$$\lim_{i \to \infty} \delta^\infty_\square(\tilde{h}^*_{h_0^{[i]} \to h_T^{[i]}}, \tilde{h}^*_{u \to r^{\sigma_{\boldsymbol{v}^*}}}) \to 0, \qquad (5.33)$$

which implies that, for any $0 < \varepsilon' < \varepsilon$, there exist i^* such that if $i > i^*$ then

$$\delta_\square\left(\tilde{h}_t, \tilde{h}^*_{h_0^{[i]} \to h_T^{[i]}}(\cdot, \cdot, t)\right) \geq \varepsilon'. \qquad (5.34)$$

Now, using the fact that the cut distance is bounded above by the L^1 distance, for $i > i^*$ we obtain

$$\text{Leb}\left\{(x, y) \in [0,1]^2 \colon \left|h_t^{[i]}(x, y) - f^*_{h_0^{[i]}(x,y) \to h_T^{[i]}(x,y)}(t)\right| > \varepsilon'\right\} \geq \varepsilon'. \qquad (5.35)$$

Let u_{ij} and $r_{\ell m}$ denote the values of the block constants (i.e., if $(x, y) \in A_i \times A_j$, then $u(x, y) = u_{ij}$, and likewise for $r_{\ell m}$). Let

$$\kappa_t(\varepsilon) := \min_{i,j \in [I], \ell, m \in [J], \beta \colon |\beta| \geq \varepsilon} \Big[I_{1,t}\big(u_{ij}, f^*_{u_{ij} \to r_{\ell m}}(t) + \beta\big) + I_{1,T-t}\big(f^*_{u_{ij} \to r_{\ell m}}(t) + \beta, r_{\ell m}\big)$$

$$- I_{1,T}(u_{ij}, r_{\ell m})\Big] > 0,$$

$$(5.36)$$

where the last inequality follows from Lemma 2.5. By the contraction principle, we know that

$$I(h^{[i]}) \geq I_{1,t}\big(u, h_t^{[i]}\big) + I_{1,T-t}\big(h_t^{[i]}, h_T^{[i]}\big). \qquad (5.37)$$

Now suppose $i > i^*$ and evaluate the integral

$$\int_{[0,1]^2} \mathrm{d}x\,\mathrm{d}y\,\left[I_{1,t}\big(u(x,y), h_t^{[i]}(x,y)\big) + I_{1,T-t}\big(h_t^{[i]}(x,y), h_T^{[i]}(x,y)\big)\right] \tag{5.38}$$

by distinguishing two sets: the (x,y) identified in (5.35) and the rest. By (5.35), the first set has Lebesgue measure at least ε', and by (5.36) the integrand has value at least

$$I_{1,T}\big(u(x,y), h_T^{[i]}(x,y)\big) + \kappa_t(\varepsilon'), \tag{5.39}$$

while for the second set, the integrand has value at least

$$I_{1,T}\big(u(x,y), h_T^{[i]}(x,y)\big). \tag{5.40}$$

Combining the above observations, we find that the integral in (5.38) satisfies the lower bound

$$I_{1,t}\big(u, h_t^{[i]}\big) + I_{1,T-t}\big(h_t^{[i]}, h_T^{[i]}\big) \geqslant I_{1,T}(u, r^{\sigma v*}) + \varepsilon'\kappa_t(\varepsilon'). \tag{5.41}$$

This is in contradiction with (5.31). Thus, we have established (2.22). A similar argument yields (2.23). $\qquad\square$

Proof of Proposition 2.10. First, note that if $a < b$ and $c < d$, then

$$I_{1,T}(a,c) + I_{1,T}(b,d) < I_{1,T}(a,d) + I_{1,T}(b,c). \tag{5.42}$$

Combining (2.29) and (5.42) with the observation that $I_{1,T}(u^\sigma, r^\phi) = I_{1,T}(u, r^{\phi \circ \sigma^{-1}})$, we see that $r^{\phi \circ \sigma^{-1}}$ is the unique representative of $\tilde{r}$ that minimises $I_{1,T}(u, \cdot)$. Equivalently, if $u \in \mathscr{W}^{(I)}$ and $r \in \mathscr{W}^{(J)}$ for some $I, J \in \mathbb{N}$, then $\boldsymbol{v}^* = \alpha(\phi \circ \sigma^{-1})$ is the unique element of V that minimises $\boldsymbol{v} \mapsto I_{1,T}(u, r^{\sigma v})$, and the claim follows directly from Theorem 2.8. If not, then the same follows after we consider a sequence of block approximations. $\qquad\square$

Proof of Proposition 2.11. Suppose that u, r, σ are given by (2.30), (2.31) and (2.32). Let $\boldsymbol{v}_1 = \alpha(\mathtt{id})$ and $\boldsymbol{v}_2 = \alpha(\sigma)$, where $\mathtt{id} \in \mathscr{M}$ is the identity, i.e., $r = r^{\mathtt{id}}$. By Theorem 2.8, the claim has been proven once we we have shown that: (i) $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are the unique minimisers of $\boldsymbol{v} \mapsto I_{1,T}(u, r^{\sigma v})$ (recall that $\sigma_{\boldsymbol{v}}$ is any $\sigma \in \mathscr{M}$ with $\alpha(\sigma) = \boldsymbol{v}$); (ii) $\tilde{h}^*_{u \to r^{\sigma v_1}, T} \neq \tilde{h}^*_{u \to r^{\sigma v_2}, T}$.

(i) Let $a = c = 0$, $b = d = \varepsilon$ and $u_{11} = u_{23} = r_{11} = r_{23} = 1$. Suppose $\varepsilon > 0$ is small but fixed and $T << \varepsilon$. Following the same arguments as in the proof of Proposition 2.2, we find that, for $T \downarrow 0$,

$$I_{1,T}(a', b') = |a' - b'| \log(1/T) + O(1), \qquad a', b' \in [0,1]. \tag{5.43}$$

For $a', b' \in [0,1]$, we then have

$$\begin{aligned} I_{1,T}(a', a') &\sim I_{1,T}(b', b'), & I_{1,T}(0, \varepsilon) &\sim I_{1,T}(\varepsilon, 0), \\ I_{1,T}(\varepsilon, 1) &\sim I_{1,T}(1, \varepsilon), & I_{1,T}(0, 1) &\sim I_{1,T}(1, 0), \end{aligned} \tag{5.44}$$

and

$$I_{1,T}(a', a') << I_{1,T}(0, \varepsilon) << I_{1,T}(\varepsilon, 1) < I_{1,T}(0, 1). \tag{5.45}$$

Observe that

$$I_{1,T}(u, r^{\sigma v_1}) = I_{1,T}(u, r^{\sigma v_2}) = \tfrac{4}{25} I_{1,T}(0, \varepsilon) + \tfrac{4}{25} I_{1,T}(\varepsilon, 0) + o(1), \tag{5.46}$$

where here we apply $\lim_{T \to 0} I_{1,T}(a', a') = 0$ which reflects the fact that edges are increasingly unlikely to change (i.e., go from active to inactive or vice-versa) between times 0 and T as

$T \to 0$. Note that above expression for $I_{1,T}(u, r^{\sigma v_1})$ and $I_{1,T}(u, r^{\sigma v_2})$ contains none of the much larger terms $I_{1,T}(\varepsilon, 1)$, $I_{1,T}(1, \varepsilon)$, $I_{1,T}(0, 1)$, $I_{1,T}(1, 0)$ listed in (5.45). Further note that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are the only elements of V whose corresponding rate function does not incur any of these much larger terms. Below we use this fact to establish *(i)*. In particular, we first show that if $\boldsymbol{v}$ is a minimiser of $\boldsymbol{v} \mapsto I_{1,T}(u, r^{\sigma v})$, then it must be close to either $\boldsymbol{v}_1$ or $\boldsymbol{v}_2$. Afterwards we show that if $\boldsymbol{v}$ is close to $\boldsymbol{v}_1$ ($\boldsymbol{v}_2$), then by moving still closer to $\boldsymbol{v}_1$ ($\boldsymbol{v}_2$) we strictly decrease the rate.

Recall that V^* denotes the set of $\boldsymbol{v} \in V$ that minimise $\boldsymbol{v} \mapsto I_{1,T}(u, r^{\sigma v})$. Suppose that $\boldsymbol{v} \in V^*$. Let $z_1 = \frac{1}{5} - v_{11}$. Observing that $v_{12} + v_{13} = z_1$, $v_{22} + v_{32} \geqslant \frac{1}{5}$ and $v_{23} + v_{33} \geqslant \frac{1}{5}$, we obtain

$$I_{1,T}(u, r^{\sigma v}) > \int_{A_2 \times A_3 \cup A_3 \times A_2} dx\, dy\, I_{1,T}\big(u(x,y), r^{\sigma v}(x,y)\big) \geqslant \tfrac{2z_1}{5} I_{1,T}(1, \varepsilon), \qquad (5.47)$$

which, by (5.43), (5.46) and the fact that $\boldsymbol{v} \in V^*$, implies that if $\varepsilon \leqslant 1/5$ then for T sufficiently small,

$$z_1 \leqslant \frac{4\varepsilon}{5(1 - \varepsilon)} \leqslant \varepsilon. \qquad (5.48)$$

Now suppose that $v_{11} \leqslant \varepsilon$ and $v_{22} > v_{32}$. Let $z_2 = \frac{2}{5} - v_{22}$. Observing that $v_{33} > \frac{1}{5} - \varepsilon$, we have

$$I_{1,T}\big(u, r^{\sigma v}\big) > \int_{A_2 \times A_3 \cup A_3 \times A_2} dx\, dy\, I_{1,T}\big(u(x,y), r^{\sigma v}(x,y)\big) \geqslant 2z_2\big(\tfrac{1}{5} - \tfrac{1}{2}\varepsilon\big) I_{1,T}(1, \varepsilon), \quad (5.49)$$

which, by the same reasoning as above, implies that if $\varepsilon < 1/40$ then for T sufficiently small

$$z_2 \leqslant \frac{4\varepsilon}{5(1 - 5\varepsilon)(1 - \varepsilon)} \leqslant \varepsilon. \qquad (5.50)$$

Similarly, if $v_{11} \leqslant \varepsilon$ and $v_{22} > v_{32}$ and $z_3 = \frac{2}{5} - v_{33}$, then when $\varepsilon < 1/40$ for T sufficiently small $z_3 \leqslant \varepsilon$. Thus, we have shown that if $\boldsymbol{v} \in V^*$, $v_{22} > v_{32}$ and $\varepsilon < 1/40$, then

$$\|\boldsymbol{v} - \boldsymbol{v}_1\|_\infty \leqslant \varepsilon \qquad (5.51)$$

for T sufficiently small. The same arguments can be used to show that if $\boldsymbol{v} \in V^*$, $v_{22} < v_{32}$, and $\varepsilon < 1/40$, then $\|\boldsymbol{v} - \boldsymbol{v}_2\|_\infty \leqslant \varepsilon$ for T sufficiently small. It can also be easily shown that if $v_{22} = v_{32}$, then $\boldsymbol{v}$ cannot be a minimiser of $\boldsymbol{v} \mapsto I_{1,T}(u, r^{\sigma v})$. Thus, we have shown that $\boldsymbol{v}$ must be close to either $\boldsymbol{v}_1$ or $\boldsymbol{v}_2$.

Next suppose that $\|\boldsymbol{v} - \boldsymbol{v}_1\|_\infty \leqslant \varepsilon$. We show that by moving still closer to $\boldsymbol{v}_1$ we strictly decrease the value of the associated rate. To that end, first suppose that $v_{11} < \frac{1}{5}$. Then, because $v_{11} + v_{12} + v_{13} = \frac{1}{5}$, either $v_{12} > 0$ or $v_{13} > 0$. In addition, because $v_{11} + v_{21} + v_{31} = \frac{1}{5}$, either $v_{21} > 0$ or $v_{31} > 0$. Suppose, for instance, that $v_{12} > 0$ and $v_{31} > 0$, and let $\eta \leqslant \min\{v_{12}, v_{31}\}$. Define $\boldsymbol{v}'$ as

$$v'_{11} = v_{11} + \eta, \qquad v'_{32} = v_{32} + \eta, \qquad v'_{12} = v_{12} - \eta, \qquad v'_{31} = v_{31} - \eta, \qquad (5.52)$$

and $v'_{ij} = v_{ij}$ otherwise. Let

$$D(\boldsymbol{v}, \boldsymbol{v}', x, y) := I_{1,T}\big(u(x,y), r^{\sigma v}(x,y)\big) - I_{1,T}\big(u(x,y), r^{\sigma v'}(x,y)\big). \qquad (5.53)$$

From elementary considerations, we obtain

$$\int_{A_1^2} \mathrm{d}x \, \mathrm{d}y \, D(\boldsymbol{v}, \boldsymbol{v}', x, y) = \tfrac{2\eta \log(1/T)}{5} + O(\eta + \eta\varepsilon \log(1/T)),$$

$$\int_{A_1 \times A_2 \cup A_2 \times A_1} \mathrm{d}x \, \mathrm{d}y \, D(\boldsymbol{v}, \boldsymbol{v}', x, y) = \tfrac{4\eta \log(1/T)}{5} + O(\eta + \eta\varepsilon \log(1/T)),$$

$$\int_{A_1 \times A_3 \cup A_3 \times A_1} \mathrm{d}x \, \mathrm{d}y \, D(\boldsymbol{v}, \boldsymbol{v}', x, y) = O(\eta + \eta\varepsilon \log(1/T)),$$

$$\int_{A_2^2} \mathrm{d}x \, \mathrm{d}y \, D(\boldsymbol{v}, \boldsymbol{v}', x, y) = -\tfrac{4\eta \log(1/T)}{5} + O(\eta + \eta\varepsilon \log(1/T)), \tag{5.54}$$

$$\int_{A_2 \times A_3 \cup A_3 \times A_2} \mathrm{d}x \, \mathrm{d}y \, D(\boldsymbol{v}, \boldsymbol{v}', x, y) = O(\eta + \eta\varepsilon \log(1/T)),$$

$$\int_{A_3^2} \mathrm{d}x \, \mathrm{d}y \, D(\boldsymbol{v}, \boldsymbol{v}', x, y) = 0.$$

Consequently, it is possible to choose ε and T sufficiently small to ensure that $I_{1,T}(u, r^{\sigma \boldsymbol{v}'}) < I_{1,T}(u, r^{\sigma \boldsymbol{v}})$. Moreover, it is possible to select ε and T such that this inequality holds for any $\boldsymbol{v} \in V$ such that $\|\boldsymbol{v} - \boldsymbol{v}_1\|_\infty \leqslant \varepsilon$, $v_{11} < 1/5$, $v_{12} > 0$, and $v_{31} > 0$. Following the same arguments, we see that the same is true when $v_{11} < 1/5$, $v_{12} > 0$, $v_{21} > 0$, $\eta \leqslant \min\{v_{12}, v_{21}\}$ and $\boldsymbol{v}'$ is defined as

$$v'_{11} = v_{11} + \eta, \qquad v'_{22} = v_{22} + \eta, \qquad v'_{12} = v_{12} - \eta, \qquad v'_{21} = v_{21} - \eta \tag{5.55}$$

or when $v_{11} < 1/5$, $v_{13} > 0$, $v_{31} > 0$, $\eta \leqslant \min\{v_{13}, v_{31}\}$ and $\boldsymbol{v}'$ is defined as

$$v'_{11} = v_{11} + \eta, \qquad v'_{33} = v_{33} + \eta, \qquad v'_{13} = v_{13} - \eta, \qquad v'_{31} = v_{31} - \eta \tag{5.56}$$

or when $v_{11} < 1/5$, $v_{13} > 0$, $v_{21} > 0$, $\eta \leqslant \min\{v_{13}, v_{21}\}$, and $\boldsymbol{v}'$ is defined as

$$v'_{11} = v_{11} + \eta, \qquad v'_{23} = v_{23} + \eta, \qquad v'_{13} = v_{13} - \eta, \qquad v'_{31} = v_{31} - \eta. \tag{5.57}$$

Note that for each of these cases, after replacing $\boldsymbol{v}$ by $\boldsymbol{v}'$ we decrease $z_1 = \tfrac{1}{5} - v_{11}$ and strictly decrease the associated rate. We can hence conclude that if ε and T are sufficiently small, $\|\boldsymbol{v} - \boldsymbol{v}_1\|_\infty \leqslant \varepsilon$ and $\boldsymbol{v} \in V^*$, then $v_{11} = \tfrac{1}{5}$. Using similar arguments, we can verify that if $\|\boldsymbol{v} - \boldsymbol{v}_1\|_\infty \leqslant \varepsilon$, $v_{11} = \tfrac{1}{5}$ and $v_{22} < \tfrac{2}{5}$, then $I_{1,T}(u, r^{\sigma \boldsymbol{v}_2}) < I_{1,T}(u, r^{\sigma \boldsymbol{v}})$ for ε and T sufficiently small. Thus, if $\boldsymbol{v} \in V^*$ and $\|\boldsymbol{v} - \boldsymbol{v}_1\|_\infty \leqslant \varepsilon$, then $\boldsymbol{v} = \boldsymbol{v}_1$ when ε and T are sufficiently small. Repeating these arguments when $\|\boldsymbol{v} - \boldsymbol{v}_2\|_\infty \leqslant \varepsilon$, we have established (i).

(ii) The proof focusses on the diagonal blocks A_1^2, A_2^2, A_3^2. We have

$$h^*_{u \to r^{\sigma \boldsymbol{v}_1}}(x, y, t) = \begin{cases} f^*_{1 \to 1}(t), & (x, y) \in A_1^2, \\ f^*_{0 \to 0}(t), & (x, y) \in A_2^2, \\ f^*_{\varepsilon \to \varepsilon}(t), & (x, y) \in A_3^2, \end{cases} \tag{5.58}$$

and

$$h^*_{u \to r^{\sigma \boldsymbol{v}_2}}(x, y, t) = \begin{cases} f^*_{1 \to 1}(t), & (x, y) \in A_1^2, \\ f^*_{0 \to \varepsilon}(t), & (x, y) \in A_2^2, \\ f^*_{\varepsilon \to 0}(t), & (x, y) \in A_3^2. \end{cases} \tag{5.59}$$

Observe that, for almost all values of $t \in (0, T)$, $f_{0 \to 0}^*(t)$ is different from $f_{1 \to 1}^*(t)$, $f_{0 \to \varepsilon}^*(t)$, and $f_{\varepsilon \to 0}^*(t)$. Fix one such value of t, and let

$$C^* := \min \left\{ |f_{0 \to 0}^*(t) - f_{1 \to 1}^*(t)|, |f_{0 \to 0}^*(t) - f_{0 \to \varepsilon}^*(t)|, |f_{0 \to 0}^*(t) - f_{\varepsilon \to 0}^*(t)| \right\} > 0. \quad (5.60)$$

For $\phi \in \mathscr{M}$ and $i = 1, 2, 3$, let $L_i^\phi = \{\phi(x) \in A_2 \colon x \in A_i\}$. Since $u, r \in \mathscr{W}^{(3)}$ and $\mathrm{Leb}(A_2) = \frac{2}{5}$, for any $\phi \in \mathscr{M}$ there exists an i such that $\mathrm{Leb}(L_i^\phi) \geqslant \frac{2}{15}$. Consequently,

$$
\begin{aligned}
&d_\square \left(h_{u \to r^\sigma v_1}^* (\cdot, \cdot, t), \left(h_{u \to r^\sigma v_2}^* (\cdot, \cdot, t) \right)^\phi \right) \\
&\qquad \geqslant \left| \int_{L_i^\phi \times L_i^\phi} \mathrm{d}x \, \mathrm{d}y \, \left[h_{u \to r^\sigma v_1}^* (x, y, t) - (h_{u \to r^\sigma v_2}^*)^\phi (x, y, t) \right] \right| \geqslant C^* \left(\tfrac{2}{15} \right)^2.
\end{aligned}
\quad (5.61)
$$

Since this bound is uniform in ϕ, we have $\delta_\square (\tilde{h}_{u \to r^\sigma v_1}^* (\cdot, \cdot, t), \tilde{h}_{u \to r^\sigma v_2}^* (\cdot, \cdot, t)) > 0$. $\qquad \square$

References

[1] Athreya, S. & den Hollander, F., & Röllin, A. (2020). Graphon-valued stochastic processes from population genetics. To appear in: *Annals of Applied Probability. arXiv preprint*: arXiv:1908.06241.

[2] Borgs, C., Chayes, J., Lovász, L., Sós, V., & Vesztergombi, K. (2008). Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Mathematics* **219**, 1801–1851.

[3] Borgs, C., Chayes, J., Lovász, L., Sós, V., & Vesztergombi, K. (2012). Convergent sequences of dense graphs II: Multiway cuts and statistical physics, *Annals of Mathematics* **176**, 151–219.

[4] Borgs, C., Chayes, J., Gaudio, J., Petti, S., & Sen, S. (2020). A large deviation principle for block models. *arXiv preprint*: arXiv:2007.14508.

[5] Černý, J. & Klimovsky, A. (2018). Markovian dynamics of exchangeable arrays. *arXiv preprint*: arXiv:1810.1316a5.

[6] Chatterjee, S. (2017). *Large Deviations for Random Graphs.* Lecture Notes in Mathematics **2197**. Springer, New York NY, USA.

[7] Chatterjee, S. & Varadhan, S.R.S. (2011). The large deviation principle for the Erdős-Rényi random graph. *European Journal of Combinatorics* **32**, 1000–1017.

[8] Crane, H. (2016). Dynamic random networks and their graph limits. *Annals of Applied Probability* **26**, 691–721.

[9] Crane, H. (2017). Exchangeable graph-valued Feller processes. *Probability Theory and Related Fields* **168**, 849–899.

[10] Dhara, S. & Sen, S. (2019). Large deviation for uniform graphs with given degrees. *arXiv preprint*: arXiv:1904.07666.

[11] Dembo, A. & Zeitouni, O. (1998). *Large Deviations Techniques and Applications.* Springer, New York NY, USA.

[12] Feng, J. & Kurtz, T. (2006). *Large Deviations for Stochastic Processes.* American Mathematical Society, Providence RI, USA.

[13] Finner, H. (1992). A generalization of Hölder's inequality and some probability inequalities. *Annals of Probability* **20**, 1893–1901.

[14] den Hollander, F. (2000). *Large Deviations*, Fields Institute Monographs 14, American Mathematical Society, Providence RI, USA

[15] Liggett, T.M. (1985). *Interacting Particle Systems.* Grundlehren der mathematischen Wissenschaften 276, Spinger-Verlag, New York.

[16] Lovász, L. (2012). *Large Networks and Graph Limits.* American Mathematical Society, Providence RI, USA.

[17] Lovász, L. & Szegedy, B. (2006). Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B* **96**, 933–957.

[18] Lovász, L. & Szegedy, B. (2007). Szemerédi's lemma for the analyst. *Geometric And Functional Analysis (GAFA)* **17**, 252–270.

[19] Lubetzky, E. & Zhao, Y. (2015). On replica symmetry of large deviations in random graphs. *Random Structures & Algorithms* **47**, 109–146.

[20] Mandjes, M. (1999). Rare event analysis of the state frequencies of a large number of Markov chains. *Stochastic Models* **15**, 577–592.

[21] Markering, M. (2020). *The large deviation principle for inhomogeneous Erdős-Rényi random graphs.* Bachelor thesis, Leiden University.

[22] Ráth, B. (2012). Time evolution of dense multigraph limits under edge-conservative preferential attachment dynamics. *Random Structures & Algorithms* **41**, 365–390.

[23] Shwartz, A. & Weiss, A. (1995). *Large Deviations for Performance Analysis.* Chapman & Hall, London, United Kingdom.

KORTEWEG-DE-VRIES INSTITUUT, UNIVERSITEIT VAN AMSTERDAM, PO BOX 94248, 1090 GE AMSTERDAM, THE NETHERLANDS

Email address: `pbraunsteins@gmail.com`

MATHEMATISCH INSTITUUT, UNIVERSITEIT LEIDEN, PO BOX 9512, 2300 RA LEIDEN, THE NETHERLANDS

Email address: `denholla@math.leidenuniv.nl`

KORTEWEG-DE VRIES INSTITUUT, UNIVERSITEIT VAN AMSTERDAM, PO BOX 94248, 1090 GE AMSTERDAM, THE NETHERLANDS

Email address: `M.H.R.Mandjes@uva.nl`