Received: 26 August 2016 | Revised: 09 April 2019 | Accepted: 31 March 2020 | Published on: 17 November 2020

DOI: 10.1002/rsa.20973

RESEARCH ARTICLE

WILEY

Dynamic concentration of the triangle-free process

Tom Bohman[1] | **Peter Keevash**[2]

[1]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania

[2]Mathematical Institute, University of Oxford, Oxford, United Kingdom

Correspondence
Tom Bohman, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213.
Email: tbohman@math.cmu.edu

Funding information
This research was supported by the NSF Grants, DMS-1001638 and DMS-1100215 (T.B.). ERC Grants, 647678 and 239696. EPSRC Grant, EP/G056730/1 (P.K.).

Abstract

The triangle-free process begins with an empty graph on n vertices and iteratively adds edges chosen uniformly at random subject to the constraint that no triangle is formed. We determine the asymptotic number of edges in the maximal triangle-free graph at which the triangle-free process terminates. We also bound the independence number of this graph, which gives an improved lower bound on the Ramsey numbers $R(3,t)$: we show $R(3,t) > (1/4 - o(1))t^2/\log t$, which is within a $4 + o(1)$ factor of the best known upper bound. Our improvement on previous analyses of this process exploits the self-correcting nature of key statistics of the process. Furthermore, we determine which bounded size subgraphs are likely to appear in the maximal triangle-free graph produced by the triangle-free process: they are precisely those triangle-free graphs with density at most 2.

KEYWORDS

dynamic concentration, Ramsey numbers, triangle-free

1 | INTRODUCTION

Constrained random graph processes provide both an interesting class of random graph models and a natural source for constructions in graph theory. Although the dependencies introduced by the constraints make such processes difficult to analyze, the evidence to date suggests that they are particularly useful for producing graphs of interest for certain extremal problems. Here we consider the triangle-free random graph process, which is defined by sequentially adding edges, starting with the empty graph, chosen uniformly at random subject to the constraint that no triangle is formed. Formally, let $G(0)$ be the empty graph on n vertices. At stage i we have a graph $G(i)$; we denote its edge set by $E(i)$, and let $O(i)$ be the set of pairs xy that are *open*, in that $G(i) \cup \{xy\}$ has no triangle. We obtain $G(i+1)$ from $G(i)$ by adding a uniformly random pair from $O(i)$.

This process was introduced by Bollobás and Erdős (see [9]), and first analyzed by Erdős et al. [12], using a differential equations method introduced by Ruciński and Wormald [23] for the analysis of the constrained graph process known as the "d-process." One motivation for their work was that their analysis of the triangle-free process led to the best lower bound on the Ramsey number $R(3,t)$ known at that time. The Ramsey number $R(s,t)$ is the least number n such that any graph on n vertices contains a complete graph with s vertices or an independent set with t vertices. In general, very little is known about these numbers, even approximately. The upper bound $R(3,t) = O(t^2/\log t)$ was obtained by Ajtai et al. [1], but for many years the best known lower bound, due to Erdős [11], was $\Omega(t^2/\log^2 t)$. The order of magnitude was finally determined by Kim [17], who showed that $R(3,t) = \Omega(t^2/\log t)$. He employed a semi-random construction that is loosely related to the triangle-free process, thus leaving open the question of whether the triangle-free process itself achieves this bound; this was conjectured by Spencer [25] and proved by Bohman [5]. There is now a large literature on the general H-free process, obtained by replacing "triangle" by any fixed graph H in the definition; see [8, 10, 19–22, 28–33]. However, the theory is still very much in its early stages: we conjectured that our lower bound for H strictly 2-balanced, given in [8], gives the correct order of magnitude for the length of the process, but so far this has only been proved for some special graphs (cycles [21, 22, 29], K_4 [30], and the diamond [20]).

In this paper, we specialize to the triangle-free process, where we can now give an asymptotically optimal analysis. Our improvement on previous analyses of this process exploits the self-correcting nature of key statistics of the process. For a treatment of self-correction in a simpler context see [6]. The methods that we use to establish self-correction of the triangle-free process build on the ideas used recently by Bohman et al. [7] for an analysis of the triangle-removal process. Furthermore, the results of this paper have also been obtained independently and simultaneously by Fiz Pontiveros et al. [13]; their proof also exploits self-correction, but is different to ours in some important ways (particularly in the methodologies for establishing self-correction and the analysis of the early part of the process, and also including many subtle differences, such as the definitions of the ensemble of key statistics that can be mutually controlled throughout the process).

Let G be the maximal triangle-free graph at which the triangle-free process terminates.

Theorem 1.1. *With high probability, every vertex of G has degree* $(1+o(1))\sqrt{\frac{1}{2}n\log n}$. *Thus the number of edges in G is* $\left(\frac{1}{2\sqrt{2}}+o(1)\right)(\log n)^{1/2}n^{3/2}$ *with high probability.*

We also obtain the following bound on the size of any independent set in G.

Theorem 1.2. *With high probability, G has independence number at most* $(1+o(1))\sqrt{2n\log n}$.

An immediate consequence is the following new lower bound on Ramsey numbers. The best known upper bound is $R(3,t) < (1+o(1))t^2/\log t$, due to Shearer [24].

Theorem 1.3. $R(3,t) > \left(\frac{1}{4}-o(1)\right)t^2/\log t$.

These results are predicted by a simple heuristic. The graph $G(i)$ that we get after i steps of the triangle-free process should closely resemble the Erdős-Rényi random graph $G_{n,p}$ with $i = n^2p/2$, with the exception that $G_{n,p}$ should have many triangles while $G(i)$ has none.

In addition to Theorems 1.1 and 1.2 we show that this heuristic extends to all small subgraph counts; in particular, we answer the folklore question (brought to our attention by Joel Spencer) of which subgraphs appear in G. The *density* of a graph H with $V_H \neq \emptyset$ is $d(H) = \frac{|E_H|}{|V_H|}$. The *maximum density* $m(H)$ of H is the maximum of $d(H')$ over nonempty subgraphs H' of H.

Theorem 1.4. *Let H be a nonempty triangle-free graph.*

(i) If $m(H) \leq 2$ *then* $\mathbb{P}(H \subseteq G) = 1 - o(1)$.
(ii) If $m(H) > 2$ *then* $\mathbb{P}(H \subseteq G) = o(1)$.

Thus, the small subgraphs that are likely to appear in G are exactly the same as the triangle-free subgraphs that appear in $G_{n,p}$ when $p = \Theta(n^{-1/2} \log^{1/2} n)$.

Note that the lower bound on $R(3, t)$ given by the triangle-free process is nonconstructive; for an explicit construction of a triangle-free graph on $\Theta(t^{3/2})$ vertices with independence number less than t see Alon [2]. Alon et al. [3] gave a construction that can be applied to G to produce a *regular* Ramsey $R(3, t)$ graph, at the cost of a worse constant in the lower bound on $R(3, t)$.

The bulk of this paper is occupied with the analysis required for the lower bound in Theorem 1.1. To prove this, we in fact prove much more generally that we can "track" several ensembles of "extension variables" for most of the process; this is formalized as Theorem 2.13. The proof of Theorem 2.13 is outlined in the next section, then implemented over the four following sections. In Section 3 we present some coupling and union bound estimates that are needed throughout the paper, and also prove Theorem 1.4, assuming Theorem 2.13. In Sections 4 to 6, we prove Theorem 2.13 via a self-correcting analysis of three ensembles of random variables. Section 7 is mostly occupied by the proof of Theorem 1.2; it also contains the proof of the upper bound in Theorem 1.1, which is similar and easier. We conclude with some brief remarks in Section 8.

2 | OVERVIEW OF LOWER BOUND

In this section we outline the proof of the lower bound in Theorem 1.1. We are guided throughout by the heuristic that $G(i)$ should resemble $G_{n,p}$ with $i = n^2 p/2$. Before proceeding with the outline of the proof we mention a consequence of this heuristic that is central to the entire argument. We introduce a time parameter t that is a rescaling of the number of steps i, defined by

$$t = in^{-3/2}.$$

For intuition, it is helpful to think of t as a continuous parameter, as it takes values less than $\sqrt{\log n}$, which is negligible compared with the polynomial scalings of the key statistics of the process.

Note that

$$p = 2tn^{-1/2}.$$

We define $Q(i)$ to be the number of open *ordered* pairs in $G(i)$. (So $Q(i) = 2|O(i)|$.) This variable is crucial to our understanding of the process. We have $Q(0) = n^2 - n$, and the process ends exactly when $Q(i) = 0$. How do we expect $Q(i)$ to evolve? If $G(i)$ resembles $G_{n,p}$ then for any pair uv we should have

$$\mathbb{P}(uv \in O(i)) \approx \left(1 - p^2\right)^{n-2} \approx e^{-np^2} = e^{-4t^2}.$$

We set $q(t) = e^{-4t^2} n^2$ and expect to have

$$Q(i) \approx q(t)$$

for most of the evolution of the process. This is exactly what we prove.

2.1 | Strategy

We use dynamic concentration inequalities for a carefully chosen ensemble of random variables associated with the process. We aim to show $V(i) \approx v(t)$ for all variables V in the ensemble, for some smooth function $v(t)$, which we refer to as the *scaling* of V. Here $V(i)$ denotes the value of V after i steps of the process, and we scale time as $t = in^{-3/2}$. For each V we define a *tracking variable* $\mathcal{T}V(i)$ and aim to show that $\mathcal{D}V(i) = V(i) - \mathcal{T}V(i)$ satisfies $|\mathcal{D}V(i)| < \delta_V(t)v(t)$, for some error functions $\delta_V(t)$. We use $\mathcal{T}V(i)$ rather than $v(t)$ so that we can isolate variations in V from variations in other variables that have an impact on V.

The improvement to earlier analysis of the process comes from "self-correction," that is, the mean-reverting properties of the system of variables. We take $\delta_V(t) = f_V(t) + 2g_V(t)$, where we think of $f_V(t)$ as the "main error term" and $g_V(t)$ as the "martingale deviation term." We usually have $g_V \ll f_V$, but there are some exceptions when t is small and hence $f_V(t)$ is too small. We require $g_V(t)v(t)$ to be "approximately nonincreasing" in t, in that $g_V(t')v(t') = O(g_V(t)v(t))$ for all $t' \geq t$.[1] We define the *critical window*

$$W_V(i) = [(f_V(t) + g_V(t))v(t), (f_V(t) + 2g_V(t))v(t)].$$

We aim to prove the *trend hypothesis* for V, which is the following statement[2]

$$\mathcal{Z}V(i) := |\mathcal{D}V(i)| - \delta_V(t)v(t) \text{ is a supermartingale when } |\mathcal{D}V(i)| \in W_V(i). \tag{1}$$

The trend hypothesis will follow from the *variation equation* for $\delta_V(t)$, which balances the changes in $\mathcal{D}V(i)$ and $\delta_V(t)v(t)$. Since errors can transfer from one variable to another, each variation equation is a differential inequality that can involve many of the error functions.

We aim to track the process up to the time

$$t_{\max} = \frac{1}{2}\sqrt{(1/2 - \varepsilon)\log n},$$

where $\varepsilon > 0$ is a constant, fixed throughout the paper, that can be arbitrarily small. More precisely, we will define a stopping time I as the first step i at which we have failure of various events (defined below), which include the event that V satisfies its required bounds. It will suffice to show that $I > i_{\max} := t_{\max}n^{3/2}$ with high probability.

One way that $I \leq i_{\max}$ can occur is when there exists $i^* = I \leq i_{\max}$ and some variable V where $\mathcal{D}V(i^*)$ is too large. In this situation, $\mathcal{D}V$ enters W_V from below at some[3] step $i' < i^*$, stays in $W_V(i)$ for $i' \leq i < i^*$ then goes above $W_V(i^*)$ at step i^*. During this time $\mathcal{Z}V(i)$ is a supermartingale, with $\mathcal{Z}V(i'-1) \leq -g_V(t')v(t')$ and $\mathcal{Z}V(i^*) \geq 0$, so we have an increase of at least $g_V(t')v(t')$ against the drift of the supermartingale. Then we use Freedman's martingale inequality [14], which is as follows.

Lemma 2.1 (Freedman). *Suppose $(X(i))_{i\geq 0}$ is a supermartingale with respect to the filtration $\mathcal{F} = (\mathcal{F}_i)_{i\geq 0}$. Suppose that $X(i+1) - X(i) \leq B$ for all i and define $V(j) = \sum_{i=1}^{j} \mathrm{Var}(X(i) \mid \mathcal{F}_{i-1})$. Then for any $a, v > 0$ we have*

---

[1]There will be one exceptional type of variable, the vertex degrees, for which this does not hold.

[2]We will only give the analysis for "upper critical window," that is, we consider $\mathcal{D}V(i)$ positive; the case of $\mathcal{D}V(i)$ negative can be treated in exactly the same way with reversed signs. We also remark that we need to "freeze" $\mathcal{Z}V(i)$ if V becomes "bad" (see (13) in Section 2.5).

[3]We will be able to assume a certain lower bound $i' > i_V$ via coupling arguments given in Section 3, and also that V is "good" (see Section 2.5).

$$\mathbb{P}\left(\exists i \text{ such that } X(i) \geq X(0) + a \text{ and } V(i) \leq v\right) \leq \exp\left(-\frac{a^2}{2(v + Ba)}\right).$$

To apply Freedman's inequality, we let $\mathcal{F} = (\mathcal{F}_i)_{i \geq 0}$ be the natural filtration for the triangle-free process, in which each $\mathcal{F}_i$ consists of all events determined by the choice of the first i edges, and we estimate

$$\text{Var}_V(i) := \text{Var}(\mathcal{Z}V(i) \mid \mathcal{F}_{i-1}) \quad \text{and} \quad N_V(i) := |\mathcal{Z}V(i+1) - \mathcal{Z}V(i)|.$$

Since $g_V(t)v(t)$ is approximately nonincreasing (unless V is a vertex degree variable), to obtain the required estimate $|\mathcal{D}V(i)| < \delta_V(t)v(t)$ with subpolynomial failure probability, it suffices to have the following two bounds, which together we call the *boundedness hypothesis*:

$$g_V(t)^2 v(t)^2 = \omega\left(\text{Var}_V(i)(n \log n)^{3/2}\right), \tag{2}$$

$$g_V(t)v(t) = \omega\left(N_V(i) \log n\right). \tag{3}$$

The lower bound of Theorem 1.1 will follow from Theorem 2.13, in which we show $I > i_{\max}$ with high probability, so every variable in our ensembles satisfies the required estimate for all $i < i_{\max}$; in particular $Q(i) > 0$, so the process persists at least to step $i_{\max}$. The proof of Theorem 2.13 is by a union bound over a polynomial number of events, each of which has subpolynomial failure probability (for brevity, we say these events hold "whp," meaning "with high probability"). We divide these events into four groups, which are treated successively over the next four sections: first events not analyzed by the critical window method described in this section, and then critical window events for three types of variables. The above discussion proves that for each variable V the required critical window event holds whp under the trend and boundedness hypotheses. For ease of reference we formulate this as a lemma, in which I_V denotes the first $i \geq i_V$ (the "activation step" for V, see Definition 2.9) at which the required estimate on V fails (we let $I_V = \infty$ if there is no such step).

Lemma 2.2. *For any variable V and step $i_V \geq 1$, if $|\mathcal{D}V(i_V)| < \delta_V(t_V)v(t_V)$ and the trend and boundedness hypotheses for V hold for all $i_V \leq i < I$ then whp we do not have $I = I_V \leq i_{\max}$.*

2.2 | Variables

All definitions are with respect to the graph $G(i)$ of edges at step i of the triangle-free process. Sometimes we use a variable name to also denote the set that it counts, for example, $Q(i)$ is the number of ordered open pairs, and also denotes the set of ordered open pairs. We usually omit (i) and (t) from our notation, for example, Q means $Q(i)$ and q means $q(t)$. We use capital letters for variable names and the corresponding lower case letter for the scaling. We express scalings using the (approximate) edge density and open pair density; these are respectively

$$p = 2in^{-2} = 2tn^{-1/2} \quad \text{and} \quad \hat{q} = e^{-4t^2}.$$

The next most important variable in our analysis, after the variable Q defined above, is the variable Y_{uv} which, for a fixed pair of vertices uv, is the number of vertices w such that uw is an open pair and vw is an edge. It is natural that Y_{uv} should play an important role in this analysis, as it directly controls the evolution in the number $Q(i)$ of ordered open pairs: if uv is the edge selected at step $i+1$ then

$$Q(i) - Q(i+1) = 2(1 + Y_{uv} + Y_{vu}).$$

Similarly, we have the following expression, used throughout the paper, for the probability (conditional on the history of the process up to step i) that any particular open pair in $Q(i)$ is not open in $Q(i+1)$:

$$\mathbb{P}(uv \notin Q(i+1) \mid \mathcal{F}_i, uv \in Q(i)) = 2(1 + Y_{uv} + Y_{vu})/Q. \tag{4}$$

From the heuristics (which we will prove) $Y \approx y = 2t\hat{q}n^{1/2}$ and $Q \approx q = \hat{q}n^2$ we can approximate edge-closure probabilities by

$$4y(t)/q(t) = 8tn^{-3/2} = -n^{-3/2}\hat{q}'(t)/\hat{q}(t), \tag{5}$$

which agrees with the intuition provided by the mean value approximation

$$\hat{q}(t) - \hat{q}(t + n^{-3/2}) \approx -\hat{q}'(t)n^{-3/2}.$$

To control these variables we need to embed them in some larger ensembles of variables that mutually control each other. The motivation for introducing each of the ensembles defined below is as follows: control of the global variables is needed to get good control of Q (better than that implied by control of all Y_{uv}), control of the stacking variables is needed to get good control of Y_{uv}, and controllable variables play a crucial role in our analysis of the stacking variables.

2.2.1 | Global variables

We begin with the variable that we are most interested in understanding: the number of open pairs. We also include two other variables that will allow us to maintain precise control on the number of open pairs.

- $Q = 2|O(i)|$ is the number of ordered open pairs. The scaling is $q = \hat{q}n^2$.
- R is the number of ordered triples with 3 open pairs. The scaling is $r = \hat{q}^3n^3$.
- S is the number of ordered triples abc where ab is an edge and ac, bc are open pairs. The scaling is $s = p\hat{q}^2n^3 = 2t\hat{q}^2n^{5/2}$.

We refer to Q, R, and S as *global* variables.

2.2.2 | Controllable variables

Next we formulate a very general condition under which we can approximate a variable up to a proportional error with polynomial decay. Suppose Γ is a graph, J is a spanning subgraph of Γ and $A \subseteq V_\Gamma$. We refer to (A, J, Γ) as an *extension*. Suppose that $\phi : A \to [n]$ is an injective mapping. We define the *extension variables* $X_{\phi,J,\Gamma}(i)$ to be the number of injective maps $f : V_\Gamma \to [n]$ such that

(i) f restricts to ϕ on A,
(ii) $f(e) \in E(i)$ for every $e \in E_J$ not contained in A, and
(iii) $f(e) \in O(i)$ for every $e \in E_\Gamma \setminus E_J$ not contained in A.

We call (J,Γ) the *underlying graph pair* of $X_{\phi,J,\Gamma}$. We introduce the abbreviations $V = X_{\phi,J,\Gamma}$,

$$n(V) = |V_\Gamma| - |A|, \quad e(V) = e_J - e_{J[A]}, \quad \text{and} \quad o(V) = (e_\Gamma - e_J) - (e_{\Gamma[A]} - e_{J[A]}),$$

which are respectively the numbers of vertices, edges, and open pairs[4] not contained in the base of the extension. The *scaling* of V is a deterministic function of the time t defined by

$$v = x_{A,J,\Gamma} = n^{n(V)} p^{e(V)} \hat{q}^{o(V)},$$

that is, it predicts the evolution of V according to the heuristic that each of the $\sim n^{n(V)}$ injections $f : V_\Gamma \to [n]$ satisfying (i) should independently satisfy (ii) for each $e \in E_J \setminus E_{J[A]}$ with probability p and (iii) for each $e \in E_\Gamma \setminus E_{\Gamma[A]}$ with probability $\hat{q}$. This prediction is correct only if there is no subextension that is "dense," in that it has scaling much smaller than 1.

When considering such subextensions $(B, J[B'], \Gamma[B'])$ with $A \subseteq B \subseteq B' \subseteq V_\Gamma$, we denote the scaling by[5]

$$S_B^{B'} = S_B^{B'}(J,\Gamma) = n^{|B'|-|B|} p^{e_{J[B']}-e_{J[B]}} \hat{q}^{(e_{\Gamma[B']}-e_{J[B']})-(e_{\Gamma[B]}-e_{J[B]})}.$$

For example, $S_A^{V_\Gamma} = v$. Note that if $A \subseteq B \subseteq B' \subseteq B'' \subseteq V_\Gamma$ then $S_B^{B''} = S_{B'}^{B''} S_B^{B'}$.

Let $t' \geq 1$. We say that V is *controllable at time* t' if $o(V) > 0$ (i.e., at least one pair not contained in the base is open) and for $1 \leq t \leq t'$ and $A \subsetneq B \subseteq V_\Gamma$ we have

$$S_A^B(J,\Gamma) \geq n^{\delta'}, \tag{6}$$

where $\delta' > 0$ is a fixed global parameter much smaller than ε (see (8) below for the parameter hierarchy).

The *controllable ensemble* is the collection of variables $X_{\phi,J,\Gamma}$ controllable at time 1 such that $|V_\Gamma| \leq M^3$, where $M = 3/\varepsilon$ (see (10)).

Remark 2.3. The proof that we can track the controllable ensemble (up to the precision needed for our purposes) is relatively short. In a certain sense, our results on controllable variables can be viewed as a triangle-free process analogue of the concentration on subgraph extensions in $G_{n,p}$ that follows from Kim-Vu polynomial concentration (see Lemma 3.4). A similar analogue should hold for the triangle removal process, and the introduction of this idea would simplify the analysis of the triangle removal process recently given by Bohman et al. [7].

We emphasize that variables in the controllable ensemble may not be controllable at all times, and that when we call a variable "controllable" we mean that it is controllable at a particular time (usually denoted by t).

2.2.3 | Stacking variables

In order to understand the evolution of the global variables Q, R, and S, we now introduce an ensemble of *stacking variables*. The name of this ensemble indicates that its members are obtained by stacking basic building blocks, each of which is a one-vertex extension. We start with two such extensions which

[4]We hope that this will not be confusable with our use of the "little-o" notation $o(1) \to 0$ as $n \to \infty$.
[5]The letter "S" is used for scalings and stacking variables, but we hope that this will not lead to any confusion, as the use is determined by the form of the superscript.

are defined for every ordered pair uv. We have already met the first, Y_{uv}, in our above discussion of the evolution of Q; the second, X_{uv}, is clearly required for understanding the evolution of Y_{uv}, as if w contributes to X_{uv} and we select the edge vw then w will instead contribute to Y_{uv}.

- Y_{uv} is the number of vertices w such that uw is an open pair and vw is an edge. The scaling is $y = 2t\hat{q}n^{1/2}$.
- X_{uv} is the number of vertices w such that uw and vw are open pairs. The scaling is $x = \hat{q}^2 n$.

The other two building blocks are one-vertex "degree" extensions defined for every vertex u.

- X_u is the *open degree* of u, defined as the number of vertices w such that uw is open. The scaling is $x_1 = n\hat{q}$.
- Y_u is the *degree* of u, defined as the number of vertices w such that uw is an edge. The scaling is $y_1 = 2tn^{1/2}$.

We will define stacking variables by composing certain sequences of such one-vertex extensions. We start by setting up notation for describing an arbitrary such variable, although we will only track a subset of the collection of the stacking variables, the M-bounded stacking variables, which will be defined later in the section.

Definition 2.4. We define[6] a *symbol set* $\Sigma^* = \{O, E, Y^O, X^O, Y^I, X^I\}$ and let $\mathcal{S}$ be the set of all nonempty finite sequences π in Σ (i.e., $\pi \in \cup_{m\geq 1}\Sigma^m$) such that

(i) if E occurs then it only does so as the last symbol of π,
(ii) $\pi(1) \notin \{Y^I, X^I\}$,
(iii) there is no j with $\pi(j) = O$ and $\pi(j+1) \in \{Y^I, X^I\}$, except possibly in the last two positions.

For any $\pi \in \mathcal{S}$ and pair of vertices uv (we will only consider $uv \notin E(i)$) we define S^{π}_{uv} according to the following rules. At each step there is an *active rung* (initially uv) and a *last vertex* (initially v). Suppose we have constructed $i-1$ steps of our stacking variable and that we have an active rung xy with last vertex y. If $\pi(i) = O$ ("open") then the next step is an X_y extension, the single open pair in this extension is the new active rung, and the new vertex is the new last vertex. If $\pi(i) = E$ ("edge") then the next step is an Y_y extension and then there is no active rung: the variable terminates here.

Now suppose $\pi(i) \notin \{O, E\}$; that is, suppose $\pi(i)$ indicates an X or Y extension on the active rung. The superscript indicates the *direction* of this extension. For Y it determines whether we add Y_{xy} or Y_{yx}, and the new open pair becomes the active rung. For X it determines which of the two new open pairs becomes the active rung. In both cases, a superscript of O (for "outer") indicates that the new active rung is incident with the last vertex, y, while a superscript of I (for "inner") indicates that the next active rung is not incident with y (i.e., it is incident with x).

We think of S^{π}_{uv} as counting injections ψ from $V(S^{\pi}_{uv}) := \{\alpha_u, \alpha_v, \alpha_1, \ldots, \alpha_{|\pi|}\}$ to $[n]$ such that $\psi(\alpha_u) = u$, $\psi(\alpha_v) = v$ and each $\psi(\alpha_j)$ is a vertex that plays the role in the extension defined by $\pi(j)$ for $j = 1, \ldots, |\pi|$, that is, $S^{\pi}_{uv} = X_{\phi,J,\Gamma}$ is the extension variable with $V(\Gamma) = V(S^{\pi}_{uv})$, $A = \{\alpha_u, \alpha_v\}$, $\phi(\alpha_u) = u$, $\phi(\alpha_v) = v$ and (J, Γ) is defined so that edges specified by the extension are mapped to edges of $G(i)$, and likewise for open pairs.

The above distinction between "inner" and "outer" is crucial for understanding what kind of proportional accuracy one should expect in controlling these variables. For an intuitive explanation of

[6]Each symbol represents a certain extension (as described below). We include condition (i) so that the definition makes sense and (ii), (iii) so as to reduce the number of cases in the analysis of stacking variables.

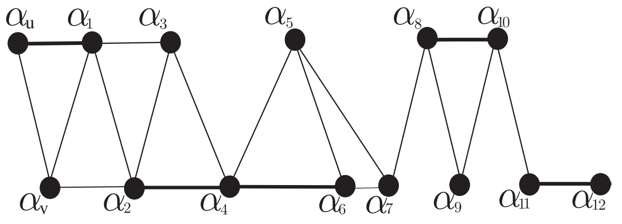

FIGURE 1 The stacking variable S^{π}_{uv} corresponding to $\pi = Y^O X^O X^O Y^O O Y^O X^I O O Y^O O E$. Thick lines represent edges and thin lines represent open pairs. Open pairs with one vertex in each row of the diagram are rungs. There are 3 triangular ladders, namely $\pi[-1;4] = \pi[v;4]$, $\pi[4;7]$, and $\pi[8;10]$, which respectively have sets of turning points $\{\alpha_1, \alpha_2, \alpha_3\}$, $\{\alpha_5\}$, and $\{\alpha_9\}$

this phenomenon, and to clarify the meaning of the definition, we introduce a pictorial representation of stacking variables, in which we think of the vertices of the active rung as the locations of the feet of someone walking on the graph. An outer extension corresponds to moving the other foot to that moved in the previous step, whereas an inner extension corresponds to moving the same foot (the intuition in the latter case is that the variable then "sees less" of the graph and so suffers a less accurate approximation).

In our pictorial representation (see Figure 1), we visualize π as a horizontal strip of two rows ("top" and "bottom"), with vertex labels arranged sequentially from left to right according to the corresponding order in π. We start by assigning α_u to the top and α_v to the bottom. In each step we assign the new vertex so that any pair of vertices meets both rows if and only if it is a rung (this uniquely defines the assignment). The direction superscripts indicate whether the new vertex is added to the same (I) or different (O) row to the last vertex. Conversely, any such drawing determines a unique order $\alpha_1, \dots, \alpha_t$ of vertices, which we call the *stacking order*, from which we can reconstruct π.

We note that the vertex set of any rung is a cutset of the graph Γ associated with S^{π}_{uv}.

The simplest stacking variables are those of length 1, namely the building blocks $S^{X^O}_{uv} = X_{uv}$, $S^{Y^O}_{uv} = Y_{vu}$, $S^{O}_{uv} = X_v$, and $S^{E}_{uv} = Y_v$. The last two examples illustrate the general phenomenon that when $\pi(1) \in \{O, E\}$ we obtain an extension based at the single vertex v, which does not depend on u. While we could denote this variable more simply by S^{π}_{v}, it is convenient to have a unified notation for stacking variables that allows the effective base of the extension to have one or two vertices.

We also introduce some further terminology which is suggested by the faint resemblance between our drawings of stacking variables and ladders. A *triangular ladder* $\pi[x;y]$ of π is a portion of $V(S^{\pi}_{uv})$ cut off by a subsequence $x-2, \dots, y$ of consecutive positions in π where $x, \dots, y$ is a maximal subsequence such that $\pi(j) \notin \{O, E\}$ for all $x \le j \le y$. (In this definition, we adopt the convention $\alpha_u = \alpha_{-1}$ and $\alpha_v = \alpha_0$ so as to allow $x \in \{1, 2\}$.) If $x < i < y$ we say that α_i is a *turning point* if the superscript of $\pi(i+1)$ is O. Note that if α_i is a turning point then it is in at least two rungs. The open pairs containing α_i are $\alpha_{i^-}\alpha_i$ and $\alpha_j\alpha_i$ for all $i+1 \le j \le i^+$, where i^- is the previous turning point (or x if there is none) and i^+ is the next turning point (or y if there is none). If α_i is in the top row (for example) then α_{i^-} and α_j for $i+1 \le j \le i^+$ are consecutive along the bottom row. We note that any stacking variable is a concatenation of some number of triangular ladders and paths of open pairs, possibly ending with a pendant edge.

We refer to an edge or open pair that is a not a rung as a *stringer*.

We do not track all of the stacking variables defined above; instead, we will track a certain finite family (with size bounded as a function of ε). The precise definition of this family is quite subtle, as we need to take account of both size and direction in order to obtain an ensemble that can be controlled mutually with the other ensembles of variables. We will impose a bound on the length of any consecutive subsequence consisting only of symbols with superscript I (which corresponds to the walker keeping one foot fixed). We will also bound the *weight* of $\pi \in \mathcal{S}$, defined by $w(\pi) = w_1(\pi) + w_2(\pi)$, where

$$w_1(\pi) = |\{i \in |\pi| : \pi(i) \in \{O, E\}\}| \text{ and } w_2(\pi) = |\{i \in |\pi| : \pi(i) \in \{X^O, Y^O\}\}|. \tag{7}$$

Now we define the M-bounded stacking variables that constitute our stacking ensemble.

Definition 2.5. We say that a stacking sequence $\pi \in \mathcal{S}$ (see Definition 2.4) is *M-bounded* if[7]

(i) $w(\pi) \leq 2M$, and $w(\pi') < 2M$, where π' is obtained from π by deleting $\pi(|\pi|)$,
(ii) π does not contain any consecutive subsequence of length M using only $\{X^I, Y^I\}$.

We let $\mathcal{S}_M$ be the set of M-bounded stacking sequences.
The *stacking ensemble* is the collection of all variables of the form S^π_{uv} where $\pi \in \mathcal{S}_M$.

We conclude this section with a simple observation on M-bounded stacking sequences.

Lemma 2.6. *If $\pi \in \mathcal{S}_M$ is an M-bounded stacking sequence then the length of π is $|\pi| < 2M^2$.*

Proof. Let $w_3(\pi)$ be the number of maximal consecutive subsequences of π using only $\{Y^I, X^I\}$. By Definition 2.5.i we have $w_3(\pi) \leq 2M - 1$, as any two such sequences are separated by positions that contribute to $w(\pi)$. Furthermore, by Definition 2.5.ii each such subsequence of has length at most $M - 1$. Therefore $|\pi| \leq w(\pi) + (M-1)w_3(\pi) \leq 2M + (2M-1)(M-1) < 2M^2$. ∎

2.3 | Tracking variables

Recall that each variable V has a tracking variable $\mathcal{T}V$ and we track the difference $\mathcal{D}V = V - \mathcal{T}V$, so as to isolate variations in V from other variations in $G(i)$.

The tracking variables are defined as follows. For the global variables we take

$$\mathcal{T}Q = q, \quad \mathcal{T}R = n^3 \cdot (Q/n^2)^3 = Q^3 n^{-3}, \quad \mathcal{T}S = n^3 \cdot 2tn^{-1/2} \cdot (Q/n^2)^2 = 2tn^{-3/2}Q^2.$$

Note that $\mathcal{T}R$ and $\mathcal{T}S$ are chosen so that $\mathcal{D}R$ and $\mathcal{D}S$ isolate the variations in R and S that do not naturally follow from the variation in Q.

If V is a one-vertex extension with a edges and b open pairs not within its base we take

$$\mathcal{T}V = n \cdot (2tn^{-1/2})^a \cdot (Q/n^2)^b.$$

That is, we set $\mathcal{T}X_{uv} = Q^2 n^{-3}$, $\mathcal{T}Y_{uv} = 2tn^{-3/2}Q$, $\mathcal{T}X_u = Qn^{-1}$, and $\mathcal{T}Y_u = pn = 2tn^{1/2}$.

For the stacking variable S^π_{uv} with $|\pi| \geq 2$ we have two cases,[8] depending on the form of π. The first case is that $\pi(|\pi|-1) \neq O$ or $\pi(|\pi|) \in \{O, E\}$. We write $\pi = \pi^- \circ U$, where U is the last element of π, and let

$$\mathcal{T}S^\pi_{uv} = S^{\pi^-}_{uv} \mathcal{T}U.$$

Note that this choice of $\mathcal{T}S^\pi_{uv}$ isolates variations that are not caused by variations in $S^{\pi^-}_{uv}$.

[7]The precise form of this definition will be crucial in Sections 6.6.4 (outer destruction) and 6.6.5 (fan end destruction).
[8]Section 6.3 includes more discussion and motivation of the definition of $\mathcal{T}S^\pi_{uv}$.

The second case is that $\pi(|\pi|-1)=O$ and $\pi(|\pi|)\notin\{O,E\}$ (we must have $|\pi|\geq 2$). We write $\pi=\pi^- OU$, where U is the last element of π, let $\beta=\alpha_{|\pi|-2}$ and

$$\mathcal{T}S_{uv}^{\pi}=\begin{cases}\sum_{f\in S_{uv}^{\pi^-}} X_{f(\beta)}^2\cdot Qn^{-2} & \text{if } U\in\{X^I,X^O\}\\ \sum_{f\in S_{uv}^{\pi^-}} X_{f(\beta)}^2\cdot 2tn^{-1/2} & \text{if } U=Y^I\\ \sum_{f\in S_{uv}^{\pi^-}} X_{f(\beta)}Y_{f(\beta)}\cdot Qn^{-2} & \text{if } U=Y^O,\end{cases}$$

recalling that X_a denotes the open degree of vertex a and Y_b denotes the degree of vertex b.

For a variable V in the controllable ensemble we will only obtain fairly weak approximations, so the precise definition of the tracking variable is not very important; it is convenient for the calculations to isolate the variation due to Q, so we let

$$\mathcal{T}V=n^{n(V)}p^{e(V)}(Qn^{-2})^{o(V)}.$$

2.4 | Error functions and activation times

With the definitions of our variables in hand, we will now introduce some further notation and define the error functions δ_V (recall that we aim to show $V=\mathcal{T}V\pm v\delta_V$ for each variable V in each of the three ensembles). Throughout the paper we fix parameters according to the hierarchy

$$n^{-1}\ll\delta\ll\delta'\ll\varepsilon; \tag{8}$$

the roles of these parameters may be understood by reference to (9) and (10) for ε, to (6) for δ', and to Definition 2.7 for δ. Our asymptotic notation is respect to n, for example, $o(1)$ denotes a quantity that can be made arbitrarily small for n sufficiently large. We track the process until the time $t_{\max}$ at which $\hat{q}(t_{\max})=n^{-1/2+\varepsilon}$; thus

$$t_{\max}=\tfrac{1}{2}\sqrt{(1/2-\varepsilon)\log n}. \tag{9}$$

The constant M that bounds the size of the stacking and variables in the controllable ensemble depends on $t_{\max}$ through ε: we let

$$M=3/\varepsilon. \tag{10}$$

We will now define the error functions δ_V.

Definition 2.7. Write[9]

$$e(t)=\hat{q}(t)^{-1/2}n^{-1/4}\quad\text{and}\quad L=\sqrt{\log n}.$$

Our error functions take the form $\delta_V=f_V+2g_V$, where[10]

$$f_V(t)=c_V\phi_V(t),\quad g_V(t)=c_V\phi_V(t)\cdot\vartheta(t)L^{-1}(1+t^{-e(V)})\ \text{ and}$$

[9]We hope that e will not be confused with the base of natural logarithms; the exponential function is denoted by exp throughout the paper.
[10]We defer the definitions of c_V and ϑ to Definition 2.8.

$$\phi_V = \begin{cases} e & \text{if } V \text{ is a stacking variable,} \\ e^2 & \text{if } V \text{ is a global variable,} \\ e^\delta & \text{if } V \text{ is in the controllable ensemble.} \end{cases}$$

The behavior of the error functions in Definition 2.7 is mainly determined by the functions ϕ_V, and can be understood without reference to the deferred definitions of c_V and ϑ, as the c_V are "constants" (i.e., independent of time; they are polylogarithmic in n) and the function $\vartheta(t)$ is bounded by constants (depending on ε, but not on n). We introduce ϑ and the $t^{-e(V)}$ term in $g_V(t)$ to handle some technicalities that arise for $t = o(1)$ (which is not the most significant regime of the process, but nevertheless exhibits slightly different behavior from the later regime, so our proof must account for this difference). When $t = \Omega(1)$ we have $g_V = O(L^{-1} f_V) = o(f_V)$, whereas if $t = o(1)$ with sublogarithmic decay and $e(V) > 0$ then we have $g_V \gg f_V$. The point of the $t^{-e(V)}$ term is that the dominant term in vg_V as $t \to 0$ does not contain a power of t.

The intuition for taking $\phi_V = e$ for stacking variables is that they include the variables Y_{uv}, which have scaling $y = 2t\hat{q}n^{1/2} = 2te^{-2}$, and which one cannot expect to control to proportional error better than $y^{-1/2}$. Thus e is a natural reference point for discussing approximations. We note for future reference that

$$e \text{ increases from } e(0) = n^{-1/4} \text{ to } e(t_{\max}) = n^{-\varepsilon/2}, \tag{11}$$

so e always has sublogarithmic decay in n. The notation $L = \sqrt{\log n}$ will be convenient as we always have $t \le t_{\max} < L$. We also note for future reference that the density $\hat{q}$ of open pairs is always much large than the density p of edges: we have

$$\hat{q}/p = e^{-2}/2t > n^\varepsilon/2L > n^{\varepsilon/2}. \tag{12}$$

We take $\phi_V = e^2$ for the global variables so that for these variables we can neglect "product" error terms arising from applications of Lemma 2.14. This is well within the theoretical limit on the accuracy for Q, namely $q^{-1/2} = e^{-1}n^{-3/4} \ll e^{-2}$; the "extra room" will be helpful in the coupling arguments in Section 3 for establishing the required estimates for small t. For variables in the controllable ensemble we only require accuracy that decays sublogarithmically, so we take $\phi_V = e^\delta$, where for δ we recall the parameter hierarchy (8).

The constants c_V that appear in Definition 2.7 will be chosen in Definition 2.8 to establish the trend hypotheses (i.e., to show that each $\mathcal{Z}V$ is a supermartingale). We will see that approximation errors migrate in a complex fashion between the variables and so these choices are quite delicate. As we treat each ensemble of variables in turn during the next three sections we will derive inequalities that these constants must (and do) satisfy in order for the trend hypothesis to hold: see the "variation equations" (22), (23), (24), (25), (30), (42), (43), and (44).

We think of the c_V's as "constant" as they do not depend on time (they are all polylogarithmic in n). We specify them now in advance of the analysis, but we will keep the notation general so that it is clear how to choose the constants. We also define the function $\vartheta(t)$ used above. Note that the constants for the stacking variables are chosen very carefully, so that they *decrease* as the length of π increases (corresponding to more accurate approximations for longer extensions), which will be important in Section 6.6.1 (simple destruction), and there is a more substantial decrease for each occurrence of O or E (counted by $w_1(\pi)$), which will be important in Section 6.6.5 (fan end destruction). There is also an adjustment for the case $\pi = O$, as our argument for controlling degree extensions requires a slightly smaller constant for open degree extensions (see Section 3.4).

Definition 2.8. For all variables in the controllable ensemble we take $c_V = 1$. For the global variables we take

$$c_R = L^{40}, \quad c_S = 2L^{40}, \quad c_Q = 4L^{40}.$$

For a stacking variable $V = S^{\pi}_{uv}$, recalling $w_1(\pi)$ from (7), we set

$$c_V = c_\pi = L^{15} 9^{4M^2 - |\pi| - M w_1(\pi)} (2.2)^{-1_{\pi=O}}.$$

Let $K = M^6 = (3/\varepsilon)^6$ and $\vartheta : [0,\infty] \to [1,\infty]$ be any increasing smooth function such that

$$\vartheta(t) = e^{Kt} \text{ for } 0 \le t \le 1, \ \sup_{t\ge 0} |\vartheta(t)| \le 2e^K \text{ and } \sup_{t \ge 0}(|\vartheta'(t)| + |\vartheta''(t)|) < \infty.$$

Recalling from Definition 2.5 and Lemma 2.6 that $w(\pi) \le 2M$ and $|\pi| < 2M^2$, we see that $L^{15} \le c_V \le L^{15} 9^{4M^2}$ for any $V = S^{\pi}_{uv}$.

Next we define the "activation step" i_V at which we start tracking a variable V using the martingale arguments in Section 2.1 (before then we will use the coupling arguments of Section 3). Our definition is uniform across all V bar one technical exception in which the activation step is slightly later than one might expect.

Definition 2.9. For any variable V, the *activation step* i_V is the smallest $i \ge n^{5/4}$ for which $g_V(t) \le L^{-1}$, except that if V is a stacking variable with $e(V) = 1$ we let $i_V = n^{1.26}$.

The *activation time* is $t_V = i_V n^{-3/2}$.

In the following lemma we give some estimates for the activation steps of various variables; we also show that all error functions are $o(1)$ after activation, and justify our earlier informal assertion that the functions vg_V are approximately nonincreasing (unless V is a vertex degree variable). The notation $\widetilde{\Theta}$ denotes approximation up to a factor polylogarithmic in n.

Lemma 2.10. *Let V be any variable in any ensemble with $o(V) > 0$ (i.e., not a vertex degree).*

(i) If $e(V) = 0$ or $V = S$ then $t_V = n^{-1/4}$.
(ii) If V is a stacking variable with $e(V) > 1$ then $t_V = \widetilde{\Theta}(n^{-1/4e(V)})$.
(iii) If V is in the controllable ensemble with $e(V) > 0$ then $t_V = \widetilde{\Theta}(n^{-\delta/4e(V)})$.
(iv) $\delta_V = o(1)$ for all $t \ge t_V$.
(v) $v(t)g_V(t) = O(v(t')g_V(t'))$ whenever $t \ge t'$.

Proof. For (i), we first note that if $e(V) = 0$ then $g_V = \widetilde{\Theta}(\phi_V)$. We have $\phi_V \le e^{\delta} < n^{-\varepsilon\delta/2} \ll L^{-1}$ by (11), so by definition $i_V = n^{5/4}$, that is, $t_V = n^{-1/4}$. Also, $g_S(t) = \widetilde{\Theta}(e^2)(1 + t^{-1}) = \widetilde{\Theta}(n^{-1/2}t^{-1})$ for $t \le 1$, so $g_S(n^{-1/4}) = \widetilde{\Theta}(n^{-1/4}) \ll L^{-1}$, giving $t_S = n^{-1/4}$, as required.

For (ii), we have $g_V(t) = \widetilde{\Theta}(e)(1 + t^{-e(V)}) = \widetilde{\Theta}(n^{-1/4}t^{-e(V)})$ for $t \le 1$, which hits L^{-1} at some $t_V = \widetilde{\Theta}(n^{-1/4e(V)})$; we obtain (iii) similarly from $g_V(t) = \widetilde{\Theta}(e^{\delta})(1 + t^{-e(V)})$.

For (iv), we note that $f_V(t_V) = O(Lg_V(t_V))(1 + t_V^{-e(V)})^{-1}$. If $e(V) = 0$ then $f_V(t_V) = O(Le^{\delta}) = o(1)$. Otherwise, as $g_V(t_V) \le L^{-1}$ by definition of t_V, (i-iii) give $f_V(t_V) = O(Lg_V(t_V))(1 + t_V^{-e(V)})^{-1} = O(t_V^{e(V)}) = \widetilde{O}(n^{-\delta/4}) = o(1)$. The estimate for $t \ge t_V$ follows as $f_V(t)$ and $g_V(t)/\vartheta(t)$ are decreasing in t, and $\vartheta(t)$ is bounded by $2e^K = O(1)$ by Definition 2.8.

Finally, to see (v) we write $h(t) = v(t)g_V(t) = \Theta(n^{n(V)}\hat{q}^{o(V)}L^{-1}c_V\phi_V(1 + t^{e(V)}))$, then note that $h(t) = \Theta(h(0))$ for $t = O(1)$, and as $o(V) > 0$ there is some $t_0 = O(1)$ such that $h'(t) < 0$ for $t > t_0$. ■

2.5 | Stopping times and the main technical result

In this section we formulate our main result regarding the stopping time I (mentioned above) that provides the lower bound in Theorem 1.1. For convenience in breaking up the proof into sections, we define

$$I = \min\{I_{\text{ext}}, I_{\text{glo}}, I_{\text{con}}, I_{\text{stk}}\}$$

in terms of 4 other stopping times defined below, which are in turn analyzed over the next 4 sections. Each of these stopping times is defined as the first step at which certain good events fail (or ∞ if there is no such step). The stopping time I_{ext} controls various events that we think of as "external" to the main martingale strategy of critical window events in Section 2.1. The other stopping times control critical window events for each of the three ensembles: I_{glo} controls global variables, I_{con} controls variables in the controllable ensemble at times when they are controllable (at other times they are controlled by I_{ext}) and I_{stk} controls stacking variables. We start by defining these critical window stopping times in terms of stopping times J_V and I_V associated to each variable V as follows.

Definition 2.11. Consider any variable V in any ensemble, and write (see Section 2.2.2, all variables can be thus expressed) $V = X_{\phi,J,\Gamma}$ for some extension (A, J, Γ).

We say that V is *bad* (at step i) if there is an edge $e = \phi(x)\phi(y)$ of $G(i)$ with x, y in A and some $w \in V_\Gamma \setminus A$ such that Γ contains xw and yw, and at least one of them is in J.

If V is not bad we call it *good*.

We let J_V be the smallest[11] $i \geq i_V$ such that V is bad (or ∞ if there is no such time).

We let the stopping time I_V be the smallest i with $i_V < i < J_V$ such that $|\mathcal{D}V(i)| > \delta_V(t)v(t)$ and V is controllable if in the controllable ensemble (or ∞ if there is no such time).

We let I_{glo}, I_{con}, and I_{stk} be the respective minima of I_V over all variables V in the global, controllable, and stacking ensembles.

We note that the global variables are always good. We also note if some V is bad then $V = 0$, as a copy of V would require either a triangle in $G(i)$ (which does not exist in the triangle-free process!) or a triangle containing two edges in $G(i)$ and one open pair (which contradicts the definition of "open"). For example, if uv is an edge then $Y_{uv} = 0$. On the other hand, if uv is closed (not an edge or open) then we do track Y_{uv}; this will be important for example, for (67) in the proof of Theorem 1.1.

As indicated earlier, for the actual definition of the variable $\mathcal{Z}V(i)$ appearing in the trend hypothesis of Section 2.1 we "freeze" it at step J_V, as follows:

$$\mathcal{Z}V(i) = \begin{cases} |\mathcal{D}V(i)| - \delta_V(t)v(t) & \text{if } i < J_V \\ \mathcal{Z}V(J_V - 1) & \text{if } i \geq J_V. \end{cases} \tag{13}$$

While the stopping times $I_{\text{glo}}, I_{\text{con}}, I_{\text{stk}}$ are the main subject of the proof, we will also need some additional information about the evolution of the process, which will be captured by the "external" stopping time I_{ext}. This includes properties of $G(i)$ for $i < n^{5/4}$, sharper estimates on Q and Y_{uv} for $i < i_Y$, crude estimates for a broad class of extension variables, and control of vertex degrees (which cannot be treated by the general strategy applied to all other variables).

Definition 2.12. We let the stopping time I_{ext} be the first step i at which $G(i) \notin \mathcal{G}_i$ (or ∞ if there is no such step), where $\mathcal{G}_i$ is the "good event" that the following estimates hold, fixing a large absolute constant C:

[11] See Definition 2.9 for i_V (the "activation step").

(i) $Q(i)/q(t)$, $X_u/x_1(t)$, and $X_{uv}/x(t)$ are $1 \pm Ct^2$ for every vertex u and pair uv, whenever $n^{-0.49} \leq t \leq 0.01$,

(ii) $Y_{uv}(i)/y(t)$ and $Y_u(i)/y_1(t)$ are $1 \pm CL^8t^2 \pm Ct^{-0.4}n^{-0.2}$ for every vertex u and non-edge uv, whenever $n^{-0.49} \leq t \leq 0.01$,

(iii) $Z_{uv}(i) \leq L^4$ for all pairs uv, where the *codegree* $Z_{uv}(i)$ is the number of vertices adjacent to both u and v in $G(i)$,

(iv) for every extension (A, J, Γ) on at most M^3 vertices and all injections $\phi : A \to [n]$ we have $X_{\phi,J,\Gamma}(i) < L^{4|V_\Gamma|} \max_{A\subseteq B\subseteq V_\Gamma} S_B^{V_\Gamma}(i)$,

(v) For every extension (A, J, Γ) on $M^3 + 1$ vertices such that $S_B^{V_\Gamma} \leq y/L^7$ for all $A \subseteq B \subseteq V_\Gamma$ and all injections $\phi : A \to [n]$ we have $X_{\phi,J,\Gamma}(i) < L^{4|V_\Gamma|} \max_{A\subseteq B\subseteq V_\Gamma} S_B^{V_\Gamma}(i)$,

(vi) no good variable $V = X_{\phi,J,\Gamma}$ in the controllable ensemble has $|\mathcal{D}V(i)| > n^{-\delta^2}v(t)$ for any $n^{-1/4} \leq t \leq t_V$ such that $S_A^B > n^{\delta'}$ for all $A \subsetneq B \subseteq V_\Gamma$,

(vii) $Y_u(i) = (1 \pm \delta_{Y_1}(t))y_1(t)$ for every vertex u.

To aid intuition, we make some remarks on the use of the various properties in the definition of $\mathcal{G}_i$. The error terms from Q and Y_{uv} are ubiquitous throughout the calculations, and the tighter control expressed by (i) and (ii) handles some technical difficulties that arise for small t; a similar motivation applies to (vi). We include (vii) in $\mathcal{G}_i$ as the vertex degrees cannot be treated by the same method used for the other variables. Combining (i) and (ii) with the martingale estimates for Q and Y_{uv} after their activation steps, we obtain the following bounds that hold for all $n^{5/4} \leq i < I$. We emphasize that we will often use without further comment the facts that the approximation errors δ_Q^* and δ_Y^* for Q and Y_{uv} have sublogarithmic decay and $\delta_Y^* = O(\delta_Y)$ for all $i \geq n^{5/4}$.

$$\text{For } n^{5/4} \leq i < I \text{ we have } Q(i) = (1 \pm \delta_Q^*)q(t) \text{ and } Y_{uv}(i) = (1 \pm \delta_Y^*)y(t) \text{ if } uv \notin E(i), \quad (14)$$
$$\text{where } \delta_Q^* \leq \delta_Q,\ \delta_Q^* = O(t^2),\ \delta_Y^* \leq 2\delta_Y \text{ for } i \geq i_Y \text{ and } \delta_Y^* = O(L^8t^2) + O(t^{-0.4}n^{-0.2}).$$

The intuition for the codegree variable Z_{uv} in (iii) is that it should scale in expectation like $p^2n = 2t < \sqrt{\log n}$, so whp will be at most polylogarithmic. An important application is that

$$\text{For any two open pairs } e \text{ and } e' \text{ at most } L^4 \text{ open pairs can simultaneously close both.} \quad (15)$$

We think of (15) as "destruction fidelity," as it will allow us to approximate the number of possibilities for a set of destruction events by a sum over each event. To see that (15) follows from (iii), we can assume that e and e' share a vertex (otherwise at most 2 pairs can close both), say $e = xu$ and $e' = xv$, and then the required bound is immediate from $Z_{uv} \leq L^4$. The bound on Z_{uv} is similar to those in (iv) and (v), but we state and prove it separately to emphasize its importance and because its proof is much simpler than those of the general statements.

Conditions (iv) and (v) in $\mathcal{G}_i$ both give the same estimate (under different hypotheses) for general extensions. This estimate is quite crude, in that it exceeds by a polylogarithmic factor $L^{4|V_\Gamma|}$ the "worst-case expectation estimate" $\max_{A\subseteq B\subseteq V_\Gamma} S_B^{V_\Gamma}(i')$ (our union bounds cannot rule out the event that ϕ extends to some embedding of (B, J, Γ), which we would then expect to have $S_B^{V_\Gamma}$ extensions). This polylogarithmic loss makes it ineffective when verifying trend hypotheses, but it is easily absorbed when verifying boundedness hypotheses. This will be crucial for the controllable ensemble, where we recall that we imposed the size restriction $|V_\Gamma| \leq M^3$, so condition (v) enables us to verify the boundary case $|V_\Gamma| = M^3$ (this idea makes our treatment of extensions significantly simpler than that in [13]).

Now we state our main result on the triangle-free process.

Theorem 2.13. *With high probability $I > i_{\max} := t_{\max} n^{3/2}$.*

The lower bound in Theorem 1.1 follows from Theorem 2.13. To see this, we note that if $I > i_{\max}$ then $I_Q > i_{\max}$, so the process persists until time $t_{\max} = \frac{1}{2}\sqrt{(1/2-\varepsilon)\log n}$, and $I_{\text{ext}} > i_{\max}$, so by Definition 2.12.vii all vertex degrees at time $t_{\max}$ are $(1 \pm \delta_{Y_1}(t_{\max}))y_1(t_{\max}) = (1+o(1))2t_{\max}n^{1/2}$.

We will prove Theorem 2.13 over the next four sections, in which we in turn bound the probabilities of the events $\{I = I_{\text{ext}} \le i_{\max}\}$ (Theorem 3.1), $\{I = I_{\text{glo}} \le i_{\max}\}$ (Theorem 4.1), $\{I = I_{\text{con}} \le i_{\max}\}$ (Theorem 5.1), and $\{I = I_{\text{stk}} \le i_{\max}\}$ (Theorem 6.1); in combination these theorems imply Theorem 2.13.

Note that if $I \le i_{\max}$ then either $G(I) \notin \mathcal{G}_I$ or there is some V such that $I = I_V \le i_{\max}$, that is, $|\mathcal{D}V(I)|$ is too large and V is good at step I. We emphasize that, since we can restrict our attention to $i < I$, we may assume $\mathcal{G}_i$ and $|\mathcal{D}V(i)| \le \delta_V(t)v(t)$ for all good variables V when verifying the trend and boundedness hypotheses.

2.6 | Some calculations and further notation

We will employ the following useful lemma extensively to estimate sums of products. The proof given here is due to Patrick Bennett.

Lemma 2.14 (Product lemma)**.** *Suppose $x, y, (x_i)_{i\in I}$ and $(y_i)_{i\in I}$ are real numbers such that $|x_i - x| \le \delta$ and $|y_i - y| < \varepsilon$ for all $i \in I$. Then we have*

$$\left|\sum_{i\in I} x_i y_i - \frac{1}{|I|}\left(\sum_{i\in I} x_i\right)\left(\sum_{i\in I} y_i\right)\right| \le 2|I|\delta\varepsilon.$$

Proof. The triangle inequality gives

$$\left|\sum_{i\in I}(x_i - x)(y_i - y)\right| \le |I|\delta\varepsilon.$$

Rearranging this inequality gives

$$\begin{aligned}\sum_{i\in I} x_i y_i &= x\sum_{i\in I} y_i + y\sum_{i\in I} x_i - |I|xy \pm |I|\delta\varepsilon \\ &= \frac{1}{|I|}\left(\sum_{i\in I} x_i\right)\left(\sum_{i\in I} y_i\right) - |I|\left(\frac{1}{|I|}\sum_{i\in I} x_i - x\right)\left(\frac{1}{|I|}\sum_{i\in I} y_i - y\right) \pm |I|\delta\varepsilon.\end{aligned}$$

∎

The following notation and conventions that are used throughout the paper.

- We use compact notation for one-step differences, writing $\Delta_i(F) = F(i+1) - F(i)$ for any sequence $F(i)$ and $\Delta_i(f) = f(t + n^{-3/2}) - f(t)$ for any function $f(t)$.
- The "O-tilde" notation $f = \widetilde{O}(g)$ means $|f| \le (\log n)^A |g|$ for some absolute constant A.
- "whp" means "with high probability"; all such statements will have subpolynomial failure probability, which will justify us taking a polynomial number of them in union bounds.
- We reiterate that we denote the vertex set by $[n] = \{1, \dots, n\}$.

We conclude this section by estimating the one-step differences for variable scalings v and error terms $v\delta_V$ (recall Definitions 2.7 and 2.8). To interpret the latter formula, note that in the main term we have factored out the scaling $v\delta_V$ and the approximate probability $8tn^{-3/2}$ (see (5)) of closing any given open pair at step t; a crucial feature of the trend hypothesis calculations later will be the self-correction of open pairs in V that cancels the $o(V)$ term. We let P_V denote the power of e in δ_V, that is, P_V equals 2, 1 or δ according as V is in the global, stacking or controllable ensembles.

Lemma 2.15. *For any variable V in any ensemble and $t \geq n^{-1/4}$ we have*

$$\Delta_i(v) = v'n^{-3/2} + O(v)n^{-5/2}, \text{ and}$$

$$\Delta_i(v\delta_V) = \left(\tfrac{e(V)}{8t^2} - o(V)\right)\delta_V v \cdot 8tn^{-3/2} + \delta_V' vn^{-3/2} + O(\delta_V v)n^{-5/2}, \text{ where}$$

$$\delta_V' \geq 4tP_V\delta_V + (\vartheta'/\vartheta - e(V)t^{-1})2g_V.$$

Proof. By Taylor's theorem, for any smooth function $h(t)$ we have

$$\Delta_i(h) = h'(t)n^{-3/2} + O(n^{-3}|h''(t')|), \quad \text{where} \quad t < t' < t + n^{-3/2}.$$

We apply this first with $h = v$, which has the form $v(t) = a(t)e^{b(t)}$, where a and b are polynomials in t and b has degree at most 2, so satisfies $v'/v = O(t + t^{-1}) = O(n^{1/4})$ and $v''/v = O(t^2 + t^{-2}) = O(n^{1/2})$ for $t \geq n^{-1/4}$; this gives the first estimate. For the second, we recall that $v = n^{n(V)}p^{e(V)}\hat{q}^{o(V)}$, so

$$v'/v = e(V)/t - 8to(V).$$

Applying Taylor's theorem to $h = v\delta_V$, as $h'/h = v'/v + \delta_V'/\delta_V$ the main term in the second estimate is equal to $h'(t)n^{-3/2}$, so it remains to show $|h''(t')| = O(n^{1/2})\delta_V v$ for $t \leq t' \leq t + n^{-3/2}$. To see this, we recall that $\delta_V = f_V + 2g_V$, where vf_V and vg_V/ϑ both have the form $a(t)e^{b(t)}$ as above, so $(vf_V)'' = O(t^2 + t^{-2})vf_V = O(n^{1/2})v\delta_V$, $(vg_V/\vartheta)' = O(t + t^{-1})vg_V/\vartheta = O(n^{1/4})v\delta_V$, and $(vg_V/\vartheta)'' = O(t^2 + t^{-2})vg_V/\vartheta = O(n^{1/2})v\delta_V$. Recalling that ϑ' and ϑ'' are bounded (see Definition 2.8) we deduce $(vg_V)'' = (vg_V/\vartheta)''\vartheta + 2(vg_V/\vartheta)'\vartheta' + \vartheta'' = O(n^{1/2})v\delta_V$, as required. The bound on δ_V' follows from $f_V'/f_V = 4tP_V$ and $g_V'/g_V = 4tP_V + \vartheta'/\vartheta - e(V)t^{-1}(1 + t^{e(V)})^{-1}$. ∎

3 | COUPLING AND UNION BOUNDS

In this section we gather two types of estimates that can be made without using dynamic concentration, namely coupling and union bounds. The two key applications of these arguments are (i) showing that whp every variable V in each of three ensembles obeys its required estimates at its activation step i_V (see Lemma 3.9), and (ii) showing that whp the stopping time I_{ext} of Definition 2.12 controlling the good event $\mathcal{G}_i$ does not occur by step $i_{\max}$. We state the latter as the main theorem of this section.

Theorem 3.1. *With high probability we do not have $I = I_{ext} \leq i_{\max}$.*

Theorem 3.1 follows by combining various lemmas proved in this section showing that each of the defining properties of the event $\mathcal{G}_i$ in Definition 2.12 hold whp; specifically, properties (i), (ii), and (vi) are in Lemma 3.9, (iii) in Lemma 3.10, (iv) in Lemma 3.13, (v) in Lemma 3.12, and (vii) in Lemma 3.15.

3.1 | Extension variables in $G_{n,p}$

Our coupling arguments will compare extension variables in the triangle-free process $G(i)$ with extension variables in the Erdős-Rényi random graph $G_{n,p}$. In this subsection we briefly review some well-known theory of the latter. Suppose J is a graph and $A \subseteq V_J$. We refer to (A, J) as an *extension*. Given an injective map $\phi : A \to [n]$, where $[n]$ is the vertex set of $G_{n,p}$, we let $X^{ER}_{\phi,J}$ be the number of injective maps $f : V_J \to [n]$ such that f restricts to ϕ on A and $f(e)$ is an edge of $G_{n,p}$ for every $e \in J \setminus J[A]$. Thus $X^{ER}_{\phi,J}$ is formally defined in the same way as the extension variable $X_{\phi,J,J}$ on $G(i)$ (see Section 2.2.2), but we emphasize that $X^{ER}_{\phi,J}$ is defined on $G_{n,p}$, not on $G(i)$.

The following definition and accompanying lemma describe how a general extension can be naturally decomposed into a series of extensions that are "strictly balanced," in that they do not have any "dense subextension."

Definition 3.2. Given $A \subseteq B \subseteq B' \subseteq V_J$ we define the *scaling*

$$S^{B'}_B = S^{B'}_B(J) = n^{|B'|-|B|} p^{e_{J[B']}-e_{J[B]}}.$$

We say that (A, J) is *strictly balanced* (in $G_{n,p}$) if $S^{V_J}_B < 1$ for all $A \subsetneq B \subsetneq V_J$. The *extension series* (in $G_{n,p}$) for (A, J), denoted $(B_0, \dots, B_d)$, is constructed by the following rule. We let $B_0 = A$. For $i \geq 0$, if (B_i, J) is not strictly balanced then we choose B_{i+1} to be a minimal set C with $B_i \subsetneq C \subsetneq V_J$ that minimizes $S^C_{B_i}$; otherwise we choose $B_{i+1} = V_J$, set $d = i + 1$ and terminate the construction.

Lemma 3.3. *Let (A, J) be an extension and $(B_0, \dots, B_d)$ be its extension series in $G_{n,p}$. Then*

(i) if $A \subseteq B \subseteq B' \subseteq B'' \subseteq V_J$ then $S^{B''}_B = S^{B'}_B S^{B''}_{B'}$,
(ii) if $A \subseteq B \subseteq V_J$ and $C \subseteq V_J \setminus B$ then $S^{B\cup C}_{A\cup C} \leq S^B_A$,
(iii) each extension $(B_i, J[B_{i+1}])$ is strictly balanced,
(iv) $S^{B_{i+1}}_{B_i} \geq 1$ for $i > 0$.

Proof. Statements (i) and (ii) are clear. For (iii), we cannot have $S^{B_{i+1}}_B \geq 1$ for some $B_i \subsetneq B \subsetneq B_{i+1}$, as then $S^B_{B_i} = S^{B_{i+1}}_{B_i} / S^{B_{i+1}}_B \leq S^{B_{i+1}}_{B_i}$ contradicts minimality of B_{i+1}. For (iv), suppose for contradiction that $S^{B_{i+1}}_{B_i} < 1$ for some $i > 0$. If $i + 1 < d$ then $S^{B_{i+1}}_{B_{i-1}} < S^{B_i}_{B_{i-1}}$ contradicts the definition of B_i. On the other hand, if $i + 1 = d$ we will obtain a contradiction by showing that (B_{i-1}, J) is strictly balanced (so the extension series should have terminated with $B_i = V_J$).

To see this, consider any $B_{i-1} \subsetneq B \subsetneq V_J$ and write $B^{\cup} = B \cup B_i$, $B^{\cap} = B \cap B_i$. By strict balance of (B_i, J) we have $S^{V_J}_{B^{\cup}} \leq 1$, with equality only if $B^{\cup} = V_J$ (as $S^{B_{i+1}}_{B_i} < 1$). By (ii) and strict balance of $(B_{i-1}, J[B_i])$ we have $S^{B^{\cup}}_B \leq S^{B_i}_{B^{\cap}} \leq 1$, with equality only if $B^{\cap} = B_i$. At least one of these inequalities is strict, so $S^{V_J}_B = S^{B^{\cup}}_B S^{V_J}_{B^{\cup}} < 1$. This contradiction completes the proof. ■

Next we quote the following general extension estimate of Kim and Vu [18, Theorem 4.2.4] in a weakened form that suffices for our purposes.

Lemma 3.4. *For any $\alpha > 0$ there is $\beta > 0$ so that for any extension (A, J) with $S^B_A > n^{\alpha}$ for all $A \subsetneq B \subseteq V_J$ in $G_{n,p}$ whp $X^{ER}_{\phi,J} = (1 \pm n^{-\beta}) S^{V_J}_A$ for all injections $\phi : A \to [n]$.*

We also require a weaker estimate that can be applied to sparse extensions, as given by the following union bound lemma. We include a brief proof as it illustrates a method we will also use for similar estimates in the triangle-free process. We recall that $L = \sqrt{\log n}$.

Lemma 3.5. *If* (A, J) *is strictly balanced in* $G_{n,p}$ *then whp* $X^{ER}_{\phi,J} < L^{4|V_J\setminus A|}\max\{S^{V_J}_A, 1\}$ *for all injections* $\phi : A \to [n]$.

Proof. First we note that for any fixed $f : V_J \to [n]$ restricting to ϕ on A we have $\mathbb{P}(f \in X^{ER}_{\phi,J}) = p^{e_J - e_{J[A]}}$. Next we estimate the probability that there are s extensions in $X^{ER}_{\phi,J}$ that are disjoint outside of $\phi(A)$. An upper bound is $s!^{-1}(n^{v_J-|A|})^s \cdot (p^{e_J-e_{J[A]}})^s < (3s^{-1}S^{V_J}_A)^s$, which is subpolynomial for $s = L^4\max\{S^{V_J}_A, 1\}$.

Now we show the statement of the lemma by induction on $|V_J\setminus A|$. The base case $|V_J\setminus A| = 1$ holds by the bound on disjoint extensions. Now suppose $|V_J \setminus A| > 1$. We consider a maximal collection C of extensions disjoint outside of A. As shown above, whp $|C| \le s = L^4\max\{S^{V_J}_A, 1\}$. By maximality, any extension $\phi' \in X^{ER}_{\phi,J}$ intersects some extension $\phi^* \in C$ outside of A. By strict balance and the induction hypothesis, for any ϕ^* the number of choices for ϕ' is at most $2^{|V_J|}L^{4(|V_J\setminus A|-1)} < L^{4|V_J\setminus A|-1}$. Therefore $X^{ER}_{\phi,J} < L^{4|V_J\setminus A|-1}|C| < L^{4|V_J\setminus A|}\max\{S^{V_J}_A, 1\}$. ∎

We deduce the following estimate on general extensions.

Lemma 3.6. *For any extension* (A, J) *whp* $X^{ER}_{\phi,J} < L^{4|V_J\setminus A|}\max_{A\subseteq B\subseteq V_J} S^{V_J}_B$ *for all* ϕ.

Proof. Let $(B_0, \dots, B_d)$ be the extension series in $G_{n,p}$ for (A, J). By Lemma 3.3.iii we can apply Lemma 3.5 bound to each step of the extension series, so whp for each $0 \le i < d$ and injection $\phi_i : B_i \to [n]$ we have $X^{ER}_{\phi_i,J[B_{i+1}]} < L^{4|B_{i+1}\setminus B_i|}\max\{S^{B_{i+1}}_{B_i}, 1\}$. Thus for any injection $\phi : A \to [n]$ we have $X^{ER}_{\phi,J} < \prod_{i=0}^{d-1} L^{4|B_{i+1}\setminus B_i|}\max\{S^{B_{i+1}}_{B_i}, 1\}$. By Lemma 3.3.iv we have $S^{B_{i+1}}_{B_i} \ge 1$ for $i \ge 1$, so $X^{ER}_{\phi,J} < L^{4|V_J|}\max\{S^{B_1}_{B_0}, 1\}S^{V_J}_{B_1} = L^{4|V_J|}\max\{S^{V_J}_{B_0}, S^{V_J}_{B_1}\}$.

It remains to show that this bound is identical to that claimed by the lemma. To see this, consider any $A \subseteq B \subseteq V_J$ and write $B^\cup = B \cup B_1$, $B^\cap = B \cap B_1$ and $S^{V_J}_B = S^{B^\cup}_B S^{V_J}_{B^\cup}$. Then $S^{B^\cup}_B \le S^{B_1}_{B^\cap} \le \max\{S^{B_1}_{B_0}, 1\}$ and $S^{V_J}_{B^\cup} \le S^{V_J}_{B_1}$ by Lemma 3.3, as required. ∎

3.2 | Coupling estimates

In this subsection we estimate our variables for small t by coupling the triangle-free process $G(i)$ inside the Erdős-Rényi random graph process $ER(n, j)$, which is defined in the same way as $G(i)$ but without the condition of being triangle-free, that is, we consider a uniform random order of the set of pairs in $[n]$ and let the edge-set of $ER(n, j)$ consist of the first j pairs in this order. The coupling is defined by rejecting any pair in $ER(n, j)$ that is closed, in that it forms a triangle with previous (nonrejected) edges. Thus after j steps the selected edges form the triangle-free process $G(i)$ after i steps, where $j - i$ edges were rejected. The number of rejected edges is bounded by the number of triangles in $ER(n, j)$; call this $T(j)$.

The intuition (made precise in Lemma 3.8) is that for small t few edges are rejected, so variables in $G(i)$ are well-approximated by corresponding variables in $ER(n, j)$. This allows us to side-step technical difficulties that arise for small t when implementing the main martingale strategy of Section 2.1 (i.e., that powers of t in the error functions blow up for small t, and in any case we have to exclude very small t to obtain concentration). We will see in the calculations below that the coupling gives us the required bounds up to $t = n^{-1/4}$ (and beyond in some cases), which explains our choice of activation step i_V in Definition 2.9.

A well-known paradigm of random graphs is that the random graph $ER(n,j)$ of fixed size is very similar to the usual binomial model G_{n,p_j} where edges are chosen independently with probability $p_j = j/\binom{n}{2}$; the following lemma makes this statement precise.

Lemma 3.7 (Lemma 1.2 in [15]). *Let $\mathcal{P}$ be any graph property and $p_j = j/\binom{n}{2}$ where $j = j(n) \to \infty$ and $\binom{n}{2} - j \to \infty$. Then for n sufficiently large*

$$\mathbb{P}(ER(n,j) \in \mathcal{P}) \leq 10j^{1/2}\mathbb{P}(G_{n,p_j} \in \mathcal{P}).$$

We will view $j = j(i)$ as a random variable on the probability space of the coupling of $G(i)$ and $ER(n,j)$, which is equal to the number of steps of the Erdős-Rényi process $ER(n,j)$ that are revealed in order to obtain i edges in the coupled triangle-free process $G(i)$. We can approximate $j(i)$ and so estimate variables in $G(i)$ by those in $G_{n,p}$ as follows.

Lemma 3.8. *If $i = tn^{3/2}$ with $t \in (n^{-0.49}, 0.01)$ then whp $i \leq j(i) < (1 + O(t^2))i$. Thus for any extension (A, J, Γ) and injection $\phi : A \to [n]$ whp $X_{\phi,J,\Gamma} \leq X^{ER}_{\phi,J}$ in $G_{n,p'}$ with $p' = (1 + O(t^2))p$.*

Proof. By definition of the coupling we have $0 \leq j - i \leq T(j)$, where $T(j)$ is the number of triangles in $ER(n,j)$. As $t > n^{-0.49}$, by Lemmas 3.4 and 3.7 whp $T(j) < 2p_j^3n^3 < 20(j/n)^3$. We deduce $j < 2i$, as at step $2i$ we have seen at least $2i - 20(2i/n)^3 = (1 - 80t^2)2i > i$ edges of the triangle-free process (using $t < 0.01$). Thus $T(j) = O(t^2)i$, which gives the first statement.

To see the second, note that $X_{\phi,J,\Gamma}$ is bounded deterministically (via the coupling) by $X^{ER}_{\phi,J}$ in $ER(n, j(i))$, and by Chernoff bounds on the number of edges in $G_{n,p'}$ we can include $G_{n,p'}$ in the coupling ("tripling"?) so that whp $ER(n, j(i)) \subseteq G_{n,p'}$. ■

Having established the coupling, we now turn to its application, which is to show that any good variable V is not in or beyond its critical interval at its activation step i_V when we begin its martingale analysis; this is the final statement of the next lemma (we also include some stronger bounds required for the event $\mathcal{G}_i$ in Definition 2.12, and a stronger statement for stacking variables). We require these bounds as earlier steps are not covered by the martingale analysis: we recall from Definition 2.11 that the stopping time I_V is the smallest i with $i_V < i < J_V$ such that $|\mathcal{D}V(i)| > \delta_V(t)v(t)$ (or ∞ if there is no such time). We can assume V is good by definition of J_V (also in Definition 2.11). For convenience, we recall the estimates on $t_V = i_V n^{-3/2}$ given in Lemma 2.10: if $e(V) = 0$ or $V = S$ then $t_V = n^{-1/4}$, otherwise $t_V = \widetilde{\Theta}(n^{-1/4e(V)})$ if V is a stacking variable or $t_V = \widetilde{\Theta}(n^{-\delta/4e(V)})$ if V is in the controllable ensemble.

Lemma 3.9. *With high probability*

- *(i) $V(i) = (1 \pm O(t^2))v(t)$ for any good variable V with $e(V) = 0$ and $n^{-0.49} \leq t \leq 0.01$,*
- *(ii) $Y_{uv}(i)/y(t)$ and $Y_u(i)/y_1(t)$ are $1 \pm O(L^8t^2) \pm O(t^{-0.4}n^{-0.2})$ for every vertex u, non-edge uv, and $n^{-0.49} \leq t \leq 0.01$,*
- *(iii) no good variable $V = X_{\phi,J,\Gamma}$ in the controllable ensemble has $|\mathcal{D}V(i)| > n^{-\delta^2}v(t)$ for any $n^{-1/4} \leq t \leq t_V$ such that $S^B_A > n^{\delta'}$ for all $A \subsetneq B \subseteq V_\Gamma$,*
- *(iv) no good stacking variable V has $|\mathcal{D}V(i)| > (f_V(t) + g_V(t))v(t)$ for any $n^{-1/4} \leq t \leq t_V$,*
- *(v) no good variable V has $|\mathcal{D}V(i_V)| > (f_V(t_V) + g_V(t_V))v(t_V)$.*

Proof. For (i), we first estimate the maximum degree $\Delta(i)$ of $G(i)$. By Lemma 3.8, we can bound $\Delta(i)$ whp by the maximum degree in $G_{n,p'}$ with $p' = (1 + O(t^2))p = O(p)$, so whp $\Delta = O(pn) = O(tn^{1/2})$.

Thus any vertex is incident to $O(pn)$ edges and $O(pn)^2 = O(t^2)n$ closed pairs. Now consider any variable V with $e(V) = 0$, and recall that $v(t) = n^{n(V)}\hat{q}(t)^{o(V)}$, where $\hat{q}(t) = e^{-4t^2} = 1 - O(t^2)$. We have $n^{n(V)} \geq V(0) \geq V(i) \geq v(t) - O(t^2)n \cdot n^{n(V)-1}$, so $V(i) = (1 + O(t^2))v(t)$, as required. This also proves (v) for such variables; indeed, we have $t_V = n^{-1/4}$, so $\mathcal{D}V(i_V) = O(n^{-1/2})v(t)$ and $f_V(n^{-1/4}) + g_V(n^{-1/4}) \geq f_R(n^{-1/4}) + g_R(n^{-1/4}) = \Theta(L^{40}n^{-1/2}) \gg \mathcal{D}V(i_V)/v(t)$.

For (ii), consider any non-edge uv and $n^{-0.49} < t < 0.01$. By Lemma 3.8 we can bound Y_{uv} whp above by the degree $d(v)$ of v in $G_{n,p'}$ with $p' = (1 + O(t^2))p$. By Chernoff bounds whp $d(v) = (1 + O(t^2))pn \pm (pn)^{0.6}$, where $d(v)/y(t) = 1 + O(t^2) + O(tn^{1/2})^{-0.4}$ as $y(t) = (1 + O(t^2))pn$ and $pn = 2tn^{1/2}$. We can bound Y_{uv} below whp by $d^{ER}(v) - T(v) - P_3(uv)$, where $d^{ER}(v)$ is the degree of v in $ER(n, j(i))$, and $T(v)$, $P_2(v)$ are the numbers of triangles containing v and paths of length 3 from u to v, both in $G_{n,p'}$ (a bound on the same quantities in $ER(n, j(i))$). By Lemma 3.6, noting that $pn > n^{0.01}$, we can bound $T(v)$ and $P_2(v)$ by $L^8 \max\{p^3n^2, 1\} = O(L^8t^2)y$, which gives the stated estimate for Y_{uv}. The argument for Y_u is the same, except that there is no $P_3(uv)$ term.

For (iii), we have already shown the required bounds when $e(V) = 0$, so we can assume $e(V) > 0$. By Lemma 3.8 (which applies as $n^{-1/4} \leq t \leq t_V = \widetilde{\Theta}(n^{-\delta/4e(V)}) < 0.01$) we can bound $V(i)$ whp above by $X_{\phi,J}$ in $G_{n,p'}$ with $p' = (1 + O(t^2))p$. As $S_A^B > n^{\delta'}$ for all $A \subsetneq B \subseteq V_\Gamma$ by Lemma 3.4 we have $X_{\phi,J} = (1 \pm n^{-2\delta^2})v(t)$, say, as $\delta \ll \delta' \ll \varepsilon$ and $e(V) < M^2 = 9\varepsilon^{-2}$. For a lower bound on $V(i)$, we consider for each pair xy in V_J not contained in A how it can prevent extensions in $X_{\phi,J}$ from being counted in V (we do not need to consider $xy \subseteq A$, as such edges either make V bad or have no effect on V). We let $J + xy$ be obtained from J by adding xy as an edge and $J * xy$ be obtained from J by adding a new vertex z adjacent to both x and y. Then we can bound $V(i)$ whp below by $X_{\phi,J} - \sum_{xy} X_{\phi,J+xy} - \sum_{xy} X_{\phi,J*xy}$.

We will bound both $X_{\phi,J+xy}$ and $X_{\phi,J*xy}$ by $n^{-2\delta^2}v$. To see this bound for $X_{\phi,J+xy}$, note that $S_A^{V_J}(J + xy) = pv$ and for any $A \subsetneq B \subseteq V_J$ that $S_B^{V_J}(J + xy) \leq S_B^{V_J}(J) = v/S_A^B < n^{-\delta'}v$, so $X_{\phi,J+xy} < n^{-2\delta^2}v$ by Lemma 3.6. A similar argument applies to $X_{\phi,J*xy}$ (also using $t = \widetilde{O}(n^{-\delta/4e(V)})$), so $V(i) = (1 \pm 4n^{-2\delta^2})v(t)$. As $\mathcal{T}V(i) = v(t)(Q/q)^{o(V)} = (1 + O(t^2))v(t) = (1 \pm n^{-2\delta^2})v(t)$, this gives (iii). As $g_V(t_V) = L^{-1}$ by definition, this also proves (v) for controllable variables.

For (iv), we may assume $e(V) > 0$. As $\delta_V(t) = \Omega(\delta_V(t_V))$ for $n^{-1/4} \leq t \leq t_V$ it suffices to show $|\mathcal{D}V(i)|/v(t) = o(\delta_V(t_V))$. Applying (i) and (ii) to each step in the stacking order of V, noting that only $O(1)$ choices are forbidden at each step due to using a vertex already used by a previous step, we obtain $V(i)/v(t) = 1 \pm O(L^8t^2) \pm O(t^{-0.4}n^{-0.2})$. Similarly, the tracking variable $\mathcal{T}V$ satisfies the same estimate for $\mathcal{T}V(i)/v(t)$, so $|\mathcal{D}V(i)|/v(t) < O(L^8t^2) + O(t^{-0.4}n^{-0.2})$. This satisfies the desired bound, as if $e(V) \neq 1$ we have $\delta_V(t_V) = \widetilde{\Theta}(1)$ and $t_V = \widetilde{\Theta}(n^{-1/4e(V)})$, so $|\mathcal{D}V(i)|/v(t) = \widetilde{O}(n^{-1/2e(V)} + n^{-0.1})$ or if $e(V) = 1$ (see Definition 2.9) we have $\delta_V(t_V) = \widetilde{\Theta}(n^{-0.05})$ and $t_V = n^{-0.24}$, so $|\mathcal{D}V(i)|/v(t) = O(n^{-0.1})$. This proves (iv) and (v) for stacking variables.

For (v), the only remaining case is $V = S$, for which we recall $t_S = n^{-1/4}$. We have $S(n^{5/4}) \leq 2n^{5/4} \cdot n = 2n^{9/4}$, as each triple counted by S determines an ordered edge and a vertex. We do not count such triples if the other pairs are closed or edges, so $S(n^{5/4}) \geq 2n^{9/4} - 2P_2 - 2P_3$, where P_ℓ is the number of paths of length ℓ in $G_{n,p'}$ with $p' = O(n^{-3/4})$ (using Lemma 3.8). As $G_{n,p'}$ whp has degrees $O(n^{1/4})$ we have $S(n^{5/4}) = 2n^{9/4} \pm O(n^{7/4})$, which is well within the desired bound $(f_S(n^{-1/4}) + g_S(n^{-1/4}))s(n^{-1/4}) = \widetilde{\Theta}(n^2)$. ■

3.3 | Union bounds

In this subsection we adapt the argument of Lemmas 3.5 and 3.6 to give a crude bound on general extension variables that holds throughout the triangle-free process. Along the way, we prove Theorem 1.4, assuming Theorem 2.13. We start with the simplest instance of this argument, which is bounding the codegree $Z_{uv}(i)$ of any two vertices u and v in $G(i)$.

Lemma 3.10. *Whp for every pair uv, if* $i' - 1 < I$ *then* $Z_{uv}(i') \leq L^4$.

Proof. We may assume uv is a non-edge, otherwise $Z_{uv} = 0$. At any step $i \leq i'$ the edge added at step i completes a path of length two between u and v with probability $(Y_{uv} + Y_{vu})Q^{-1}$. We can bound this probability by $O(y/q) = O(Ln^{-3/2})$ for $t \geq 1$ or by $O(y(1)/q(0)) = O(n^{-3/2})$ for $t \leq 1$. Taking a union bound over all subsets of L^4 steps at which we might increment Z_{uv}, the probability that Z_{uv} reaches L^4 by step i' is at most $\binom{i_{\max}}{L^4} O\left(Ln^{-3/2}\right)^{L^4} = O\left(L^{-2}\right)^{L^4}$. ∎

We need some further notation and terminology for general extensions in the triangle-free process, which mirrors that used previously for extensions in the Erdős-Rényi process. We say that (A, J, Γ) is *strictly balanced at time t* if $S_B^{V_\Gamma} < 1$ for all $A \subsetneq B \subsetneq V_\Gamma$. The *extension series at time t* for (A, J, Γ), denoted $(B_0, \dots, B_d)$, is constructed by the following rule. We let $B_0 = A$. For $i \geq 0$, if (B_i, J, Γ) is not strictly balanced then we choose B_{i+1} to be a minimal set C with $B_i \subsetneq C \subsetneq V_\Gamma$ that minimizes $S_{B_i}^C$; otherwise we choose $B_{i+1} = V_\Gamma$, set $d = i + 1$ and terminate the construction.

In Lemma 3.13 we will give a general estimate for extension variables in the triangle-free process. First we illustrate the argument in the following lemma, which shows that sparse graph pairs do not appear; this is the main tool needed for the proof of Theorem 1.4. Here we take $A = \emptyset$, write $V_{J,\Gamma} = X_{\phi,J,\Gamma}$, where ϕ is the unique map from $\emptyset$ to $[n]$, and $v_{J,\Gamma} = S_\emptyset^{V_\Gamma}(J, \Gamma)$.

Lemma 3.11. *Suppose* $v_{J,\Gamma}(t') < n^{-c}$ *for some* $c > 0$ *and time* t'. *Then the probability that* $\mathcal{G}_{i'}$ *holds,* $i' - 1 < I$ *and* $V_{J,\Gamma}(i') > 0$ *is at most* $2n^{-c}$.

Proof. For $t' \leq L^{-1}$ we appeal to the coupling with the Erdős-Rényi random graph process. By Lemma 3.8 it suffices to estimate the probability that J appears in $G_{n,j}$, where $j = (1 + o(1))i'$. The expected number of copies of J is at most $2n^{-c}$, so the required bound follows from Markov's inequality. Thus it suffices to consider $t' \geq L^{-1}$.

To estimate $\mathbb{P}(V_{J,\Gamma}(i') > 0)$, we take a union bound of events, where we specify the injection $f : V_\Gamma \to [n]$, and for $e \in J$ we specify the *selection step* i_e at which the process selects the edge $f(e)$. Fix some choice and let $\mathcal{E}$ be the specified event.

For each $i \leq i'$ we estimate the probability that the selected edge is compatible with $\mathcal{E}$. At a selection step $i = i_e$ the selected edge is specified, so the probability is $2/Q(i_e) = (1 + o(1))2q(t_e)^{-1}$, where $t_e = n^{-3/2} i_e$ (the approximation of Q by q holds on $\mathcal{G}_{i'}$ and $i' - 1 < I$).

For other i, the required probability is $1 - N_i/Q$, where N_i is the number of ordered open pairs that cannot be selected at step i on $\mathcal{E}$. If i is a selection step we write $N_i = 0$. Therefore

$$\mathbb{P}(\mathcal{E} \wedge \mathcal{G}_{i'}) \leq \prod_{e \in J} (1 + o(1))2q(t_e)^{-1} \cdot \prod_{i=1}^{i'} (1 - N_i/Q). \tag{16}$$

Now we estimate N_i when i is not a selection step. For $i < L^{-1}n^{3/2}$ we use the trivial estimate $N_i \geq 0$, so suppose $i \geq L^{-1}n^{3/2}$. Suppose there are k_i choices of $e \in J$ with $i_e > i$. Then there are $|\Gamma \setminus J| + k_i$ open pairs that must not become closed, namely the open pairs of $f(\Gamma \setminus J)$ and the k_i pairs of $f(J)$ that have yet to be selected as edges. We recall from (15) that by property (iii) of $\mathcal{G}_{i'}$ only $O(L^4) = o(y)$ choices of e_i can close more than one such open pair.

As $\mathcal{G}_{i'}$ holds and $i'-1 < I$, by (14) all Y-variables are $(1+o(1))y$, so we obtain $N_i = (1+o(1))(|\Gamma \setminus J| + k_i) \cdot 4y$. Thus for $i \geq L^{-1}n^{3/2}$ we can write $1 - N_i/Q \leq 1 - (1+o(1))(A_i + B_i)$, where

$$A_i = |\Gamma \setminus J| \cdot 8tn^{-3/2} = |\Gamma \setminus J| \cdot 8in^{-3} \text{ and } B_i = k_i \cdot 8in^{-3}.$$

This holds for all i if we set $A_i = B_i = 0$ for $i < L^{-1}n^{3/2}$.

We estimate each factor by $1-(1+o(1))(A_i+B_i) \leq \exp\{-(1+o(1))A_i\}\exp\{-(1+o(1))B_i\}$ and bound separately the contributions from all A_i and from all B_i. The contribution from all A_i is

$$\begin{aligned}\exp\left\{-\sum_{i=1}^{i'}(1+o(1))A_i\right\} &= \exp\left\{-(1+o(1))|\Gamma \setminus J| \sum_{i=L^{-1}n^{3/2}}^{i'} 8in^{-3}\right\} \\ &= (1+o(1))\exp\left\{-|\Gamma \setminus J| \cdot 4(i')^2 n^{-3}\right\} \\ &= (1+o(1))e^{-4(t')^2|\Gamma \setminus J|} = (1+o(1))\hat{q}(t')^{|\Gamma \setminus J|},\end{aligned}$$

since $\sum_{i=1}^{L^{-1}n^{3/2}} in^{-3} < L^{-2} = o(1)$. The contribution from all B_i is

$$\begin{aligned}\exp\left\{-\sum_{i=1}^{i'}(1+o(1))B_i\right\} &= \exp\left\{-(1+o(1))\sum_{e\in J}\sum_{i=L^{-1}n^{3/2}}^{i_e} 8in^{-3}\right\} \\ &= \prod_{e\in J}(1+o(1))\hat{q}(t_e).\end{aligned}$$

Substituting in (16) we obtain

$$\mathbb{P}(\mathcal{E} \wedge \mathcal{G}_{i'}) \leq (1+o(1))\hat{q}(t')^{|\Gamma\setminus J|}\prod_{e\in J} 2\hat{q}(t_e)/q(t_e) = (1+o(1))\hat{q}(t')^{|\Gamma\setminus J|}(2n^{-2})^{|J|}.$$

Summing over at most $n^{|V_\Gamma|}$ choices for f and $(i')^{|J|}$ choices for the selection steps, we estimate $\mathbb{P}\left(\left\{V_{J,\Gamma}(i') > 0\right\} \wedge \mathcal{G}_{i'}\right) < (1+o(1))v(t') < 2n^{-c}$. ■

Proof of Theorem 1.4. Statement (i) is immediate from [8, Theorem 1.6(iii)]. For (ii), fix $H' \subseteq H$ with $d(H') > 2$. By choosing the global parameter $\epsilon > 0$ sufficiently small we can assume $|E_{H'}|(1/2-\epsilon) > |V_{H'}| + \epsilon$. Note that if $H \subseteq G$ then $V_{J,H'}(i_{\max}) > 0$ for some spanning subgraph J of H', that is, there is some potential embedding ϕ of H' that survives until step $i_{\max}$, in that some subgraph $\phi(J)$ is selected by the triangle-free process, and the remaining subgraph $\phi(H' \setminus J)$ remains open, so that it is available for the remainder of the process (which we do not analyze). We have

$$v_{J,H'}(t_{\max}) = n^{|V_{H'}|}p^{|E_J|}\hat{q}(t_{\max})^{|E_{H'}|-|E_J|} = n^{|V_{H'}|-|E_J|/2-(1/2-\epsilon)(|E_{H'}|-|E_J|)} < n^{-\epsilon}.$$

Thus the result follows from Theorem 2.13 and Lemma 3.11. ■

We now turn to a key lemma which includes the union bound arguments that are most significant for the whole proof: it implies property (v) of Definition 2.12 and will also be used in the proof of Lemma 3.13, which implies property (iv) of Definition 2.12.

Lemma 3.12. *For any extension (A, J, Γ) with $|V_\Gamma| = O(1)$, if $S_B^{V_\Gamma} < y/L^7$ for all $A \subseteq B \subseteq V_\Gamma$ at step i' then whp we do not have $I = I_{ext} = i'$ due to some ϕ with $X_{\phi,J,\Gamma}(i') \geq L^{4|V_\Gamma \setminus A|}\max_{A\subseteq B\subseteq V_\Gamma} S_B^{V_\Gamma}$.*

Proof. As in the proof of Lemma 3.11, it suffices to consider $t' \geq L^{-1/2}$, as for smaller t' we can simply appeal to the coupling with the Erdős-Rényi random graph process (Lemma 3.8) and apply the bound from Lemma 3.6. Furthermore, the general case of the lemma follows from the case that (A, J, Γ) is strictly balanced, by applying it to each step of the extension series (in the same way that Lemma 3.6 followed from Lemma 3.5). We will therefore only consider the case that (A, J, Γ) is strictly balanced.

We argue by induction on $|V_\Gamma \setminus A|$. Similarly to the proof of Lemma 3.5, we first estimate the probability that there is a set of s extensions $\{f_1, \dots, f_s\}$ in $V(i') := X_{\phi,J,\Gamma}(i')$ that are disjoint outside of $\phi(A)$, where $s = \max\{L^4, 6 \max_{A \subseteq B \subseteq V_\Gamma} S_B^{V_\Gamma}\}$.

Our method for estimating this probability is similar to the argument of Lemma 3.11, but now we consider s embeddings simultaneously. We take a union bound of events in which we specify $f_1, \dots, f_s$, and for each $1 \leq j \leq s$ and $e \in J \setminus J[A]$ we specify the *selection step* $i_{j,e}$ at which the process selects the edge $f_j(e)$. Fix some choice and let $\mathcal{E}$ be the specified event.

For each $i \leq i'$ we estimate the probability that the selected edge e_i is compatible with $\mathcal{E}$. At a selection step $i = i_{j,e}$ the selected edge is specified, so the probability is $2/Q(i_{j,e}) = (1+o(1))2q(t_{j,e})^{-1}$, where $t_{j,e} = n^{-3/2} i_{j,e}$. For other i, the required probability is $1 - N_i/Q$, where N_i is the number of ordered open pairs that cannot be selected at step i on $\mathcal{E}$. If i is a selection step we write $N_i = 0$. Then we estimate

$$\mathbb{P}(\mathcal{E} \wedge \mathcal{G}_{i'}) \leq \prod_{j=1}^{s} \prod_{e \in J \setminus J[A]} (1 + o(1)) 2q(t_{j,e})^{-1} \cdot \prod_{i=1}^{i'} (1 - N_i/Q).$$

Now we estimate N_i when i is not a selection step, assuming that we are in the event $\mathcal{G}_i$ and $i < I$. For $i < L^{-1/2} n^{3/2}$ we use the trivial estimate $N_i \geq 0$, so suppose $i \geq L^{-1/2} n^{3/2}$. Suppose there are k_i choices of (j, e) with $i_{j,e} > i$. Then there are $o(V)s + k_i$ open pairs that must not become closed, namely the $o(V)$ open pairs specified by each $f_1, \dots, f_s$ and the k_i pairs that have yet to be selected as edges (these pairs are distinct by disjointness of $f_1, \dots, f_s$ outside $\phi(A)$). By (15) the number of choices of the selected edge e_i that close more than one such open pair is $O(s^2 L^4) = o(syL^{-2})$, as by assumption on $\max_B S_B^{V_\Gamma}$ and choice of s we have $s < y(t')L^{-7} < y(t)L^{-6.5}$.

As $i < I$, by (14) all Y-variables are $(1 + o(1))y$, so $N_i = (1 + o(1))(o(V)s + k_i) \cdot 4y$. Similarly to the proof of Lemma 3.11, we write

$$1 - N_i/Q \leq 1 - (1 + o(1))(A_i + B_i) \leq \exp\{-(1 + o(1))A_i\} \exp\{-(1 + o(1))B_i\},$$

where $A_i = B_i = 0$ for $i < L^{-1/2} n^{3/2}$ and otherwise $A_i = o(V)s \cdot 8tn^{-3/2}$ and $B_i = k_i \cdot 8tn^{-3/2}$. As before, we estimate separately all A_i terms and all B_i terms to obtain

$$\exp\left\{-\sum (1 + o(1))A_i\right\} \leq \left[(1 + o(1))\hat{q}(t')^{o(V)}\right]^s, \text{ and}$$

$$\exp\left\{-\sum (1 + o(1))B_i\right\} \leq \prod_{j=1}^{s} \prod_{e \in J \setminus J[A]} (1 + o(1))\hat{q}(t_{j,e}), \text{ so}$$

$$\mathbb{P}(\mathcal{E} \wedge \mathcal{G}_{i'}) \leq \hat{q}(t')^{o(V)s} \prod_{j=1}^{s} \left((1 + o(1)) \prod_{e \in J \setminus J[A]} 2n^{-2} \right).$$

Summing over at most $s!^{-1}n^{n(V)s}$ choices for $f_1, \dots, f_s$ and $(i')^{e(V)s}$ choices for the selection steps, the probability that such $f_1, \dots, f_s$ exist is at most $s!^{-1}[(1+o(1))v(t')]^s < (3s^{-1}v(t'))^s$, which is subpolynomial.

The required bound on $X_{\phi,J,\Gamma}(i')$ follows from this estimate by induction as in the proof of Lemma 3.5. (The base case $|V_J \setminus A| = 1$ holds by the bound on disjoint extensions, and for $|V_J \setminus A| > 1$ the bound follows by considering a maximal collection C of extensions disjoint outside of A—we have just shown whp $|C| \le s$—noting by strict balance and the induction hypothesis that at most $L^{4|V_J \setminus A|-1}$ embeddings intersect some embedding in C outside of $\phi(A)$.) This completes the proof when (A, J, Γ) is strictly balanced, and as noted above, the general case follows by applying this to each step of the extension series. ∎

Lemma 3.13. *For any extension (A, J, Γ) with $|V_\Gamma| \le M^3$, whp we do not have $I = I_{ext} = i$ due to some ϕ with $X_{\phi,J,\Gamma} \ge L^{4|V_\Gamma \setminus A|} \max_{A \subseteq B \subseteq V_\Gamma} S_B^{V_\Gamma}$.*

Proof. By bounding each step of the extension series we can assume that (A, J, Γ) is strictly balanced. If $S_A^{V_\Gamma}(t) < n^{\delta'}$ then the required bound follows from Lemma 3.12. On the other hand, if $S_A^{V_\Gamma}(t) \ge n^{\delta'}$ then $X_{\phi,J,\Gamma}$ is controllable at time t, so the required bound follows from $i < I_{\text{con}}$. ∎

Remark 3.14. We emphasize here a subtlety in the analysis of the controllable ensemble in the previous two lemmas that may not be immediately apparent. Consider for example the variable V in the controllable ensemble that for three fixed vertices a, b, c counts vertices v with av and bv open and cv an edge. Before time L^{-1} we control V by coupling. Throughout the interval of times t from $t_V = \widetilde{\Theta}(n^{-\delta/4})$ to $t' \approx \frac{1}{4}\sqrt{\log n}$ where $\hat{q}(t')^2 p(t') n = n^{\delta'}$ we control V via the stopping time I_{con}. At later times t^*, we bound V via I_{ext}, using the union bound argument in Lemma 3.12, which operates throughout the interval from L^{-1} to t^*, including the times t in the interval when V is controllable and is much larger. The point is that this union bound argument only considers sets of s extensions where $s \ll y(t)$ for all t in the interval and shows that whp no such set survives until time t^*.

3.4 | Vertex degrees

Recall that we cannot apply our general strategy to vertex degree variables, as $g_{Y_1}(t) y_1(t)$ is not approximately nonincreasing. We conclude this section with a separate (much simpler) argument for these variables, which establishes property (vii) of $\mathcal{G}_i$ in Definition 2.12.

Lemma 3.15. *whp we do not have $I_{ext} = i'$ due to some uv with $|Y_u(i') - y_1(t')| \ge \delta_{Y_1}(t') y_1(t')$.*

Proof. For each $1 \le i \le i'$, the probability that we choose an edge incident to u is

$$\frac{2X_u(i)}{Q(i)} = \frac{(1 \pm \delta_{X_1})2x_1}{(1 \pm \delta_Q)q} = \left(1 \pm (1+o(1))\delta_{X_1}\right)\frac{2}{n}.$$

By coupling, we can bound $Y_u(i')$ by sums $\Sigma^\pm$ of independent Bernoulli random variables with probabilities $(1 \pm 2\delta_{X_1})2/n$. Now we recall from Definition 2.8 that $c_{Y_1} = 2.2c_{X_1}$, and note that $f_{Y_1} = 2.2f_{X_1}$ and $g_{Y_1}/g_{X_1} = 2.2(1+t^{-1})/2 > 1.1$, so $\delta_{Y_1} > 1.1\delta_{X_1}$. Thus on the event $|Y_u(i') - y_1| \ge \delta_{Y_1} y_1$ one of $\Sigma^\pm$ deviates from its mean $(1+o(1))2tn^{1/2}$ by more than $\delta_{Y_1} y_1/100 > L^{13} n^{1/4}$. By Chernoff bounds, whp this does not occur for any vertex u. ∎

4 | GLOBAL ENSEMBLE

In this section we prove that the global variables have the desired concentration, assuming that this is the case for all ensembles at earlier times. Recall that the global variables are the number $Q(i)$ of ordered open pairs in $G(i)$, the number of ordered triples $R(i)$ where all the pairs within the triple are open, and the number $S(i)$ of ordered triples abc such that ab is an edge while bc and ac are open pairs. The global variables have scalings $q = \hat{q}n^2$, $r = \hat{q}^3n^3$, and $s = 2t\hat{q}^2n^{5/2}$. Recall that we track each variable V relative to a *tracking random variable* $\mathcal{T}V$ to isolate variations in V from variations in other variables that might have an impact on V. We use the tracking variables

$$\mathcal{T}Q = q, \quad \mathcal{T}R = Q^3n^{-3}, \quad \text{and} \quad \mathcal{T}S = 2tn^{-3/2}Q^2.$$

(Note that the tracking variable for Q is a deterministic function.)

We show that the *difference random variables*

$$\mathcal{D}V = V - \mathcal{T}V$$

for $V \in \{Q, R, S\}$ are all small throughout the process. Recall that I_{glo} is the minimum of the stopping times I_V over all variables V in the global ensemble, that is, the first time at which some global variable V (is good and) fails to satisfy $|\mathcal{D}V| \leq \delta_V v$. (Global variables are automatically good, so we can ignore that part of the definition.) The following theorem bounds the probability that we reach the universal stopping time I before step $i_{\max}$ because a global variable V fails to satisfy the required bounds $|\mathcal{D}V| \leq \delta_V v$.

Theorem 4.1. *With high probability we do not have $I = I_{glo} \leq i_{\max}$.*

We prove Theorem 4.1 using the strategy described in Section 2.1. We divide the argument into three parts, in which we respectively bound the one-step expected changes in the difference variables, determine variation equations that suffice to establish the trend hypothesis, and verify the boundedness hypothesis.

4.1 | One-step changes in the difference variables

In this subsection, for each variable V in the global ensemble, we give an upper bound on the one-step expected change in the difference variable, conditional on the history of the process, that is,

$$\mathbb{E}[\Delta_i\mathcal{D}V \mid \mathcal{F}_i] = \mathbb{E}[\mathcal{D}V(i+1) - \mathcal{D}V(i) \mid \mathcal{F}_i],$$

under the assumption that V is in its upper critical window, that is,

$$(f_V + g_V)v < \mathcal{D}V < (f_V + 2g_V)v.$$

Recall that we can assume $n^{5/4} \leq i < I$, so we can apply the estimates from $\mathcal{G}_i$ in Definition 2.12 and the bounds $V = \mathcal{T}V \pm \delta_V v$ for any variable V if $i_V \leq i \leq J_V$ (i.e., if V is good and activated). To illustrate later calculations, which are often more complicated than those in this section, as we proceed we will indicate how certain specific calculation are instances of a more general framework.

We will consider the effect of each open pair and edge in the structure counted by V separately; the final expression is then obtained by linearity of expectation. When an open pair in a copy of the structure counted by V is chosen or closed, we say that the copy is *destroyed*. We balance the change in V due to destructions with the change in $\mathcal{T}V$ due to the change in Q. (The case $V = Q$ is handled differently as Q is tracked relative to the deterministic function q.) Adding the edge e_{i+1} can also create new copies of the structure counted by V in which e_{i+1} plays the role of one of the edges in the structure; then we say that a copy of V is *created* (for global variables this only applies to $V = S$). The change in V that comes from creations is balanced with the change in t in $\mathcal{T}V$.

We begin with destructions. The main point to note in these calculations is that the assumption that V is in its critical window gives a self-correction term of $-8tf_V vn^{-3/2}$ for each open pair, which will cancel with a corresponding $8tf_V vn^{-3/2}$ term from the change in $\delta_V v$; this arises from the critical window excess of $f_V v$ in V relative to $\mathcal{T}V$, recalling from (4) that in each such "excess copy" of V the corresponding open pair becomes closed with probability about $8tn^{-3/2}$.

4.1.1 | Q: simple destructions

We will show the following estimate for the expected one-step change in Q.

Lemma 4.2. *If $n^{5/4} \le i < I$ and $Q \ge (1 + f_Q + g_Q)q$ then*

$$\mathbb{E}[\Delta_i(\mathcal{D}Q) \mid \mathcal{F}_i] \le -(f_Q + g_Q - (1 + o(1))\delta_S)8tqn^{-3/2}.$$

For the variable Q there is another variable S in our ensemble that counts situations when some open pair counted by Q is closed. We call destructions of this form *simple destructions*. (We will see examples of this type again in Section 6 where we treat the stacking variables.)

Proof. Each triple in S contains 4 ordered open pairs, each of which would decrease Q by 2 ordered pairs if selected as the edge at step i, and by symmetry in S we count each of these possibilities twice. The selected edge itself also removes 2 ordered open pairs, so

$$\mathbb{E}[\Delta_i Q \mid \mathcal{F}_i] = -2 - 4S/Q. \tag{17}$$

Recalling from Lemma 2.15 that $\Delta_i(q) = -8tqn^{-3/2} + O(qn^{-5/2})$, we calculate

$$\begin{aligned}\mathbb{E}[\Delta_i(\mathcal{D}Q) \mid \mathcal{F}_i] &= \mathbb{E}[\Delta_i(Q) - \Delta_i(q) \mid \mathcal{F}_i] \\ &= -(2 + 4S/Q) + 8tqn^{-3/2} \pm O(qn^{-5/2}) \\ &= -8tQn^{-3/2} \pm (8 + O(\delta_Q))\delta_S tn^{1/2}\hat{q} + 8tqn^{-3/2} \pm O(1) \\ &\le -(f_Q + g_Q - (1 + o(1))\delta_S)8tqn^{-3/2}.\end{aligned}$$

In the third estimate we used $S = (1 \pm \delta_S)\mathcal{T}S = (1 \pm \delta_S)2tn^{-3/2}Q^2$ and $Q = (1 \pm \delta_Q)q$, which are valid as $n^{5/4} = i_S = i_Q \le i < I$. In the last line we used $\mathcal{D}Q = Q - q \ge (f_Q + g_Q)q$ when Q is in its upper critical window, and $\delta_S \ge g_S \ge c_S L^{-1}e^2t^{-1}$, where $c_S = 2L^{40}$ (see Definition 2.8), so $t\delta_S qn^{-3/2} \ge L^{-1}c_S e^2 qn^{-3/2} = L^{-1}c_S \gg 1$. ∎

4.1.2 | R: product destructions

We will show the following estimate for the expected one-step change in R.

Lemma 4.3. *If $n^{5/4} \le i < I$ and $R \ge (1 + f_R + g_R)\mathcal{T}R$ then*

$$\mathbb{E}[\Delta_i(\mathcal{D}R) \mid \mathcal{F}_i] \le \left[-(3 + o(1))(f_R + g_R) + O(\delta_Y\delta_X) + O(t^{-1}e^2) + O(L^{16}t^2n^{-1/2})\right] 8trn^{-3/2}.$$

The destructions for R are not simple destructions, as no variable in our ensembles counts ways in which triples counted by R are destroyed. Instead, we will apply the product lemma (Lemma 2.14). For clarity we will write out the calculation separately for R and S (in later sections we will be more efficient by introducing extra notation that unifies all cases).

Proof. To estimate the expected change, we first recall from (4) that any pair $\alpha\beta \in Q(i)$ becomes closed with probability $2(1 + Y_{\alpha\beta} + Y_{\beta\alpha})/Q$. Noting that closing $\alpha\beta$ reduces R by $3X_{\alpha\beta}$, we write

$$\mathbb{E}[\Delta_i(R) \mid \mathcal{F}_i] = -\sum_{\alpha\beta\in Q} 2Q^{-1}(1 + Y_{\alpha\beta} + Y_{\beta\alpha}) \cdot 3X_{\alpha\beta} + \mathbb{E}[F_i(R) \mid \mathcal{F}_i],$$

where $F_i(R)$ is a "destruction fidelity" correction term to remove overcounting of triples in R for which the selected edge closes two open pairs in the triple. Thus $\mathbb{E}[F_i(R) \mid \mathcal{F}_i] = F^*/Q$, where F^* is the number of ordered quadruples where two adjacent pairs are edges and the other four pairs are open. As $i < I_{\text{ext}}$, by property (iv) of $\mathcal{G}_i$ in Definition 2.12 we have $F^* < L^{16}n^4p^2\hat{q}^4 = 4L^{16}t^2\hat{q}r$, so

$$\mathbb{E}[F_i(R) \mid \mathcal{F}_i] = O(L^{16}t^2r/n^2). \tag{18}$$

Next, noting that

$$\sum_{\alpha\beta\in Q} Y_{\alpha\beta} = S \quad \text{and} \quad \sum_{\alpha\beta\in Q} X_{\alpha\beta} = R,$$

we estimate the main term using the product lemma as

$$-\sum_{\alpha\beta\in Q} 6Q^{-1}(1 + Y_{\alpha\beta} + Y_{\beta\alpha})X_{\alpha\beta} = -12SRQ^{-2} \pm O(\delta_Y y\delta_X x) \pm O(x), \tag{19}$$

whereas $n^{5/4} \le i < I$ we have the estimates $X = (1 + O(\delta_X))x$ and all Y-variables are $(1 \pm \delta_Y^*)y = (1 + O(\delta_Y))y$ from (14). The important point to observe regarding the product error term is that δ_X and $t\delta_Y$ are $\widetilde{O}(e)$, whereas δ_R is $\widetilde{O}(e^2)$, so the error term is negligible for appropriate choices of the polylogarithmic constants c_X, c_Y and c_R (see Definition 2.8).

Next we consider the expected change in the tracking variable $\mathcal{T}R = Q^3n^{-3}$. We have

$$\Delta_i(\mathcal{T}R) = Q(i+1)^3n^{-3} - Q(i)^3n^{-3} = 3\Delta_i(Q)Q^2n^{-3} + H_i(R), \tag{20}$$

where $H_i(R)$ is a "higher order" term correcting for the linear approximation of the difference in Q^3, and as $\Delta_i(Q) = O(y)$ we have $H_i(R) = (3\Delta_i(Q)^2Q + \Delta_i(Q)^3)n^{-3} = O(t^2rn^{-3})$. By (17) we have

$$\mathbb{E}[\Delta_i(\mathcal{T}R) \mid \mathcal{F}_i] = -12SQ^{-2}\mathcal{T}R + O(t^2rn^{-3}). \tag{21}$$

Combining (18), (19), and (21) gives

$$\begin{aligned}\mathbb{E}[\Delta_i(\mathcal{D}R) \mid \mathcal{F}_i] &= \mathbb{E}\left[\Delta_i(R) - \Delta_i(\mathcal{T}R) \mid \mathcal{F}_i\right] \\ &= -\sum_{\alpha\beta\in Q} 6Q^{-1}(1 + Y_{\alpha\beta} + Y_{\beta\alpha})X_{\alpha\beta} - 3Q^2n^{-3}\mathbb{E}[\Delta_i(Q) \mid \mathcal{F}_i] + \mathbb{E}[F_i(R) - H_i(R) \mid \mathcal{F}_i]\end{aligned}$$

$$= -12SQ^{-2}R \pm O(\delta_Y y \delta_X x) + O(x) + 12SQ^{-2}\mathcal{T}R + O(L^{16}t^2 r/n^2)$$
$$= -(1 \pm (3+o(1))\delta_S) 8tn^{-3/2}\mathcal{D}R \pm O(\delta_Y \delta_X) trn^{-3/2} \pm O(r/q) + O(L^{16}t^2 r/n^2)$$
$$\leq \left[-(3+o(1))(f_R + g_R) + O(\delta_Y \delta_X) + O(t^{-1}e^2) + O(L^{16}t^2 n^{-1/2})\right] 8trn^{-3/2}.$$

■

An important point to note in the above calculation is that the same factor $12SQ^{-2}$ appears with R and $\mathcal{T}R$, and that we approximate S by $(1 \pm \delta_S)\mathcal{T}S$ only after using the critical window bound $\mathcal{D}R = R - \mathcal{T}R \geq (f_R + g_R)r$; thus the fact that our approximation of S is weaker than that of R does not cause any difficulty in this calculation for R.

4.1.3 | S: product destructions and creations

For S we have both creations and destructions, so we will now elaborate on how we group the calculations for each edge of a structure (we could gloss over this for R, as it has 3 indistinguishable edges, but it will be important for most other variables, including S). Recall that S is the number of ordered triples abc where ab is an edge and ac, bc are open pairs. We write

$$\Delta_i(S) = \Delta_i(S^{12}) + \Delta_i(S^{13}) + \Delta_i(S^{23}),$$

where we think of 123 as labeling each such abc, and each $\Delta_i(S^e)$ is the change in S due to e, that is, $\Delta_i(S^{12})$ is the number of triples abc in S created due to ab being the edge selected at step i, $-\Delta_i(S^{13})$ is the number of triples abc in S destroyed due to ac being selected or closed at step i, and similarly for $-\Delta_i(S^{23})$.

Usually, we would also include a "fidelity" term $F_i(S)$ in this decomposition of changes by edges, reflecting the fact that the selected edge might affect more than one pair in a triple counted by S, but in fact this is not possible, so we can set $F_i(S) = 0$. Indeed, if selecting the edge ab creates a triple abc in S then by definition of S it does not close ac or bc, and a triple abc cannot be destroyed by some edge e_i that simultaneously closes ac and bc, as this would require $e_i = cd$ such that ad and bd are edges, but then abd would be a triangle, which is impossible.

We also decompose the change in the tracking variable $\mathcal{T}S$ into terms that we assign to the different parts of the calculation corresponding to each of the edges in S. Recalling that $\mathcal{T}S = 2tn^{-3/2}Q^2$, we have $\Delta_i(\mathcal{T}S) = 2(t + n^{-3/2})n^{-3/2}(Q + \Delta_i(Q))^2 - 2tn^{-3/2}Q^2$, which we write as

$$\Delta_i(\mathcal{T}S) = \Delta_i(\mathcal{T}S^{12}) + \Delta_i(\mathcal{T}S^{13}) + \Delta_i(\mathcal{T}S^{23}) + H_i(S),$$

where $\Delta_i(\mathcal{T}S^{12}) = 2n^{-3}Q^2 = \mathcal{T}S/tn^{3/2}$ and $\Delta_i(\mathcal{T}S^{13}) = \Delta_i(\mathcal{T}S^{23}) = 2tn^{-3/2}\Delta_i(Q)Q = \frac{\Delta_i Q}{Q}\mathcal{T}S$, with the higher-order correction term $H_i(S) = 2n^{-3}(2\Delta_i(Q)Q + \Delta_i(Q)^2) + 2tn^{-3/2}\Delta_i(Q)^2 = O(yqn^{-3}) + O(tn^{-3/2}y^2) = O(sn^{-3}) + O(t^2 sn^{-3})$.

Now we show the calculations for the change $\Delta_i(\mathcal{D}S^{13}) := \Delta_i(S^{13}) - \Delta_i(\mathcal{T}S^{13})$ (the one with 23 instead of 13 is the same); these are product destructions very similar to those for R.

Lemma 4.4. *If $n^{5/4} \leq i < I$ and $S \geq \mathcal{T}S + (f_S + g_S)s$ then*

$$\mathbb{E}[\Delta_i(\mathcal{D}S^{13}) \mid \mathcal{F}_i] \leq \left[-(1+o(1))(f_S + g_S) + O(\delta_Y^2) + O(t^{-1}e^2)\right] 8tsn^{-3/2}.$$

Proof. Similarly to the proof of Lemma 4.3, we calculate

$$\mathbb{E}[\Delta_i(\mathcal{D}S^{13}) \mid \mathcal{F}_i] = \mathbb{E}\left[\Delta_i(S^{13}) - \frac{\Delta_i(Q)}{Q}\mathcal{T}S \mid \mathcal{F}_i\right]$$

$$
\begin{aligned}
&= -\sum_{\alpha\beta\in Q} 2Q^{-1}(1 + Y_{\alpha\beta} + Y_{\beta\alpha})Y_{\alpha\beta} + (2 + 4SQ^{-1})Q^{-1}\mathcal{T}S \\
&= -4SQ^{-2}S \pm O(\delta_Y y)^2 + O(y) + 4SQ^{-2}\mathcal{T}S \pm O(s/q) \\
&= (1 \pm (1 + o(1))\delta_S) 8tn^{-3/2}\mathcal{D}S \pm O(\delta_Y^2)tsn^{-3/2} \pm O(s/q) \\
&\le \left[-(1 + o(1))(f_S + g_S) + O(\delta_Y^2) + O(t^{-1}e^2)\right] 8tsn^{-3/2}.
\end{aligned}
$$

■

Finally, we turn to creations, which among the global variables occur only for S.

Lemma 4.5. *If $n^{5/4} \le i < I$ then $\mathbb{E}[\Delta_i(\mathcal{D}S^{12}) \mid \mathcal{F}_i] \le (1 + o(1))\frac{\delta_R}{8t^2} 8tsn^{-3/2}$.*

Proof. We have $\mathbb{E}[\Delta_i(S^{12}) \mid \mathcal{F}_i] = 2R/Q$, as for each triple abc in R, with probability $2/Q$ the edge e_{i+1} selected at step $i + 1$ falls in position ab and turns abc into a triple in S. Thus

$$
\begin{aligned}
\mathbb{E}[\Delta_i(\mathcal{D}S^{12}) \mid \mathcal{F}_i] &= \mathbb{E}[\Delta_i(S^{12}) - t^{-1}n^{-3/2}\mathcal{T}S \mid \mathcal{F}_i] \\
&= 2Q^{-1}(\mathcal{T}R \pm \delta_R r) - t^{-1}n^{-3/2}\mathcal{T}S \\
&= \pm 2\delta_R r Q^{-1} = \pm(1 + o(1))t^{-1}\delta_R s n^{-3/2}.
\end{aligned}
$$

■

Note that there is no self-correction in creation, but this term will be negligible as our approximation of R is better than that of S.

4.2 | Trend hypothesis and variation equations

For each variable V in the global ensemble we consider the sequence of random variables

$$\mathcal{Z}V(i) = \mathcal{D}V - v\delta_V.$$

The following lemma establishes the trend hypothesis, that is, that this sequence is a supermartingale when V is in its upper critical window. During the proof we will derive the variation equations, which give conditions on the constants c_V under which the trend hypothesis holds; we will see that these conditions are satisfied by the choices in Definition 2.8.

Lemma 4.6. *For each $V \in \{Q, R, S\}$, if $n^{5/4} \le i < I$ and $\mathcal{D}V > (f_V + g_V)v$ then $\mathbb{E}[\Delta_i\mathcal{Z}V \mid \mathcal{F}_i] \le 0$.*

Proof. We begin by gathering together the relevant creation and destruction calculations from the previous subsections; these are obtained by combining Lemmas 4.2, 4.3, 4.4, and 4.5.

$$
\begin{aligned}
\mathbb{E}[\Delta_i\mathcal{D}Q \mid \mathcal{F}_i] &\le -\left(f_Q + g_Q - (1 + o(1))\delta_S\right) 8tqn^{-3/2}, \\
\mathbb{E}[\Delta_i\mathcal{D}R \mid \mathcal{F}_i] &\le -(1 + o(1))\left[3(f_R + g_R) - O(\delta_Y\delta_X) - O(t^{-1}e^2)\right] 8trn^{-3/2}, \\
\mathbb{E}[\Delta_i\mathcal{D}S \mid \mathcal{F}_i] &\le -(1 + o(1))\left[2(f_S + g_S) - \frac{\delta_R}{8t^2} - O(\delta_Y^2) - O(t^{-1}e^2)\right] 8tsn^{-3/2}.
\end{aligned}
$$

For R we have omitted the fidelity term in (18); this is valid as $F_i(R) = O(L^{16}t^2rn^{-2}) = o(g_R)trn^{-3/2}$, where we recall from Definition 2.8 that

$$c_R = L^{40} \gg L^{20} \text{ (say).} \tag{22}$$

Next we consider the change in $v\delta_V$. From Lemma 2.15 we have

$$\Delta_i(v\delta_V) = \left(\tfrac{e(V)}{8t^2} - o(V)\right)\delta_V v \cdot 8tn^{-3/2} + \delta'_V vn^{-3/2} + O(\delta_V v)n^{-5/2}.$$

Recalling that $\delta_V = f_V + 2g_V$, we see that we can cancel the $8to(V)f_V vn^{-3/2}$ term that occurs both in $\Delta_i(\delta_V v)$ and in $\mathbb{E}[\Delta_i \mathcal{D}V \mid \mathcal{F}_i]$; this is the self-correction that is fundamental to the analysis.

Thus we obtain

$$\begin{aligned}
\mathbb{E}[\Delta_i \mathcal{Z}Q \mid \mathcal{F}_i] &\le -\left(\frac{\delta'_Q}{8t} + o(f_Q) - (1+o(1))(g_Q + \delta_S)\right) 8tqn^{-3/2},\\
\mathbb{E}[\Delta_i \mathcal{Z}R \mid \mathcal{F}_i] &\le -\left(\frac{\delta'_R}{8t} + o(f_R) - (1+o(1))3g_R + O(\delta_Y\delta_X) - O(t^{-1}e^2)\right) 8trn^{-3/2},\\
\mathbb{E}[\Delta_i \mathcal{Z}S \mid \mathcal{F}_i] &\le -\left(\frac{\delta'_S}{8t} + \frac{\delta_S}{8t^2} + o(f_S) - (1+o(1))(2g_S + \tfrac{\delta_R}{8t^2}) + O(\delta_Y^2) - O(t^{-1}e^2)\right) 8tsn^{-3/2}.
\end{aligned}$$

Recall that our error functions have the form $\delta_V = f_V + 2g_V$, where

$$f_V = c_V e^2 \quad \text{and} \quad g_V = c_V \vartheta L^{-1}(1 + t^{-e(V)})e^2 \quad \text{if} \quad V \in \{Q, R, S\}.$$

We now show that these error functions grow quickly enough for each of these sequences to be supermartingales (i.e., the δ'_V term will be dominant in each case). We stress that the $t \ll 1$ regime behaves a bit differently from the rest of the process in the estimates that follow. For each global variable in turn we apply the bound on δ'_V from Lemma 2.15, that is,

$$\delta'_V \ge 8t\delta_V + (\vartheta'/\vartheta - e(V)t^{-1})2g_V.$$

For Q we have

$$\begin{aligned}
\mathbb{E}[\Delta_i \mathcal{Z}Q \mid \mathcal{F}_i] &\le -(1+o(1))\left[(f_Q + (\tfrac{2\vartheta'}{8t\vartheta} + 2)g_Q) - (g_Q + \delta_S)\right] 8tqn^{-3/2}\\
&\le -(1+o(1))\left[(f_Q - f_S) + (\tfrac{\vartheta'}{4t\vartheta} g_Q + g_Q - 2g_S)\right] 8tqn^{-3/2}.
\end{aligned}$$

Then the sequence $\mathcal{Z}Q$ forms a supermartingale provided

$$c_Q \ge 2c_S. \tag{23}$$

Indeed, then the dominant terms are $-f_Q$ for $t \ge 1$ and/or $-\frac{\vartheta'}{4t\vartheta} g_Q$ for $t \le 1$ (for $t \le 1$ we recall that $\vartheta'/\vartheta = (3/\varepsilon)^6$ and note that the t^{-1} in $\frac{\vartheta'}{4t\vartheta} g_Q$ matches the t^{-1} in g_S).

Next consider R, where we have

$$\begin{aligned}
\mathbb{E}[\Delta_i \mathcal{Z}R \mid \mathcal{F}_i] &\le -(1+o(1))\left[f_R + (\tfrac{2\vartheta'}{8t\vartheta} + 2)\cdot g_R - 3g_R - O(\delta_Y\delta_X) - O(t^{-1}e^2)\right] 8trn^{-3/2}\\
&\le -(1+o(1))\left[f_R + (\tfrac{\vartheta'}{4t\vartheta} - 1)g_R - O(f_Y f_X) - O(g_Y f_X)\right] 8trn^{-3/2},
\end{aligned}$$

as $t^{-1}e^2 \ll t^{-1}g_R$. Then $\mathcal{Z}R$ forms a supermartingale provided

$$c_R \ge Lc_Y c_X, \tag{24}$$

for this implies that the $g_R\vartheta'/(4t\vartheta)$ term dominates for $f_R < g_R/t$ and that the f_R term dominates otherwise. As noted earlier, we chose powers of e in the error functions so that δ_R and the product error $t\delta_Y\delta_X$ are comparable up to log factors (i.e., e in δ_X and δ_Y and e^2 in δ_R); then the choice of polylogarithmic constants c_V in Definition 2.8 was such that (24) holds.

The final global variable is S, where we have

$$\begin{aligned}\mathbb{E}[\Delta_i\mathcal{Z}S \mid \mathcal{F}_i] &\leq -(1+o(1))\left[(1+\tfrac{1}{8t^2})f_S + (\tfrac{2\vartheta'}{8t\vartheta}+2)\cdot g_S\right]8tsn^{-3/2}\\ &\quad + (1+o(1))\left[2g_S + \tfrac{\delta_R}{8t^2} + O(\delta_Y\delta_Y) + O(t^{-1}e^2)\right]8tsn^{-3/2}\\ &\leq -(1+o(1))\left[f_S + \tfrac{f_S-f_R}{8t^2} + \tfrac{\vartheta' tg_S/\vartheta - g_R}{4t^2} - O(f_Y^2) - O(g_Y^2) + o(g_S)\right]8tsn^{-3/2}.\end{aligned}$$

Then $\mathcal{Z}S$ forms a supermartingale provided

$$c_S \geq 2c_R \quad \text{and} \quad c_S \geq Lc_Y^2, \tag{25}$$

for this implies that the f_S/t^2 term dominates for $t \ll 1$ and the f_S term dominates otherwise. ∎

4.3 | Boundedness hypothesis

For the boundedness hypothesis, for each V in the global ensemble we estimate $\mathrm{Var}_V = \mathrm{Var}(\mathcal{Z}V(i) \mid \mathcal{F}_{i-1})$ and $N_V = |\Delta_i\mathcal{Z}V|$. Recall that it suffices to establish (2) and (3); that is, it suffices to show the following lemma.

Lemma 4.7. *For each $V \in \{Q,R,S\}$, if $n^{5/4} \leq i < I$ then* $\mathrm{Var}_V = o\left(\frac{(g_Vv)^2}{L^3n^{3/2}}\right)$ *and* $N_V = o\left(\frac{g_Vv}{L^2}\right)$.

Proof. For convenience we replace $\mathcal{Z}V$ by $\mathcal{D}V$ in our calculations, as this does not change Var_V and only changes N_V by an additive term which we can bound by $O(n^{-5/4}v\delta_V)$.

For one-step variances we use the simple estimate $\mathrm{Var}_V \leq N_V^2$ (so for the global variables we do not need the full power of Freedman's inequality: it suffices to apply the Hoeffding-Azuma inequality).

For Q we have $g_Qq \geq c_QL^{-1}n^{3/2}$, so it suffices to show $\mathrm{Var}_Q = o(c_Q^2L^{-5}n^{3/2})$ and $N_Q = o(c_QL^{-3}n^{3/2})$. The change in $\mathcal{D}Q$ when the process chooses the edge $e_{i+1} = uv$ is

$$\Delta_i\mathcal{D}Q = 2(Y_{uv} + Y_{vu} + 1) - \Delta_i(q) = 4(y \pm y\delta_Y) - 4y + O(1) = O(y\delta_Y) = O(c_YLn^{1/4}).$$

Then $N_Q = \widetilde{O}(n^{1/4})$, $Var_Q = \widetilde{O}(n^{1/2})$, and the required bounds hold easily.

For R we have $g_Rr \geq c_RL^{-1}\hat{q}^2n^{5/2}$, so it suffices to show $\mathrm{Var}_R = o\left(c_R^2L^{-5}\hat{q}^4n^{7/2}\right)$ and $N_R = o\left(c_RL^{-3}\hat{q}^2n^{5/2}\right)$. Recall from (20) that $\Delta_i(\mathcal{T}R) = 3\Delta_i(Q)Q^2n^{-3} + H_i(R)$, where $H_i(R) = O(t^2rn^{-3}) = \widetilde{O}(1)$. On choosing $e_{i+1} = uv$ we have

$$\Delta_iR = F_i(R) - \sum_{ab\in Y_{uv}\cup Y_{vu}\cup\{uv\}} 6X_{ab},$$

where, as in the proof of Lemma 4.3, $F_i(R)$ is a "destruction fidelity" correction term to remove overcounting of triples in R for which the selected edge closes two open pairs in the triple. We can bound $F_i(R)$ by the number of triples uab counted by R such that va and vb are edges (and similarly interchanging u and v). As $n^{5/4} \leq i < I$, by property (iv) of $\mathcal{G}_i$ in Definition 2.12 we have $F_i(R) = \widetilde{O}(1 + n\hat{q}^3)$. Combining these estimates gives

$$\begin{aligned}
\Delta_i \mathcal{D}R &= \Delta_i R - 3\frac{\Delta_i(Q)}{Q}\mathcal{T}R + \widetilde{O}(1) \\
&= -\sum_{ab \in Y_{uv} \cup Y_{vu} \cup \{uv\}} 6X_{ab} - 2(Y_{uv} + Y_{vu} + 1) \cdot 3Q^2 n^{-3} + \widetilde{O}(1 + n\hat{q}^3) \\
&= -6\left[\sum_{ab \in Y_{uv} \cup Y_{vu} \cup \{uv\}} (X_{ab} - Q^2 n^{-3})\right] + \widetilde{O}(y + n\hat{q}^3) \\
&= O(yx\delta_X) + \widetilde{O}(y + n\hat{q}^3) = \widetilde{O}(\hat{q}^{5/2} n^{5/4}).
\end{aligned}$$

Then $N_R = \widetilde{O}(\hat{q}^{5/2} n^{5/4})$, $Var_R = \widetilde{O}(\hat{q}^5 n^{5/2})$, and the required bounds hold easily.

For S we have $g_S s \geq c_S L^{-1} \hat{q} n^2$, so it suffices to show $\mathrm{Var}_S = o(c_S^2 L^{-5} \hat{q}^2 n^{5/2})$ and $N_S = o(c_S L^{-3} \hat{q} n^2)$. We bound the impact of creations and destructions separately, recalling the decompositions of the change in S as $\Delta_i(S) = \Delta_i(S^{12}) + \Delta_i(S^{13}) + \Delta_i(S^{23})$, where $\Delta_i(S^{12})$ counts creations and $\Delta_i(S^{13})$, $\Delta_i(S^{13})$ count destructions. We also recall the corresponding decomposition of the change in the tracking variable as $\Delta_i(\mathcal{T}S) = \Delta_i(\mathcal{T}S^{12}) + \Delta_i(\mathcal{T}S^{13}) + \Delta_i(\mathcal{T}S^{23}) + H_i(S)$, where $\Delta_i(\mathcal{T}S^{12}) = 2n^{-3}Q^2 = \mathcal{T}S/tn^{3/2}$, $\Delta_i(\mathcal{T}S^{13}) = \Delta_i(\mathcal{T}S^{23}) = 2tn^{-3/2}\Delta_i(Q)Q = \frac{\Delta_i Q}{Q}\mathcal{T}S$, and $H_i(S) = O((1 + t^2)sn^{-3})$.

On choosing $e_{i+1} = uv$, we estimate the destruction terms (e.g., that for S^{13}) by

$$\begin{aligned}
\Delta_i \mathcal{D}S^{13} &= \Delta_i S^{13} - 2\frac{\Delta_i(Q)}{Q}\mathcal{T}S + O((1 + t^2)sn^{-3}) \\
&= -\sum_{ab \in Y_{uv} \cup Y_{vu} \cup \{uv\}} (Y_{ab} + Y_{ba} - 2 \cdot 2tQn^{-3/2}) + O((1 + t^2)sn^{-3}) \\
&= O(y \cdot \delta_Y y) = \widetilde{O}(\hat{q}^{3/2} n^{3/4}).
\end{aligned}$$

For the creation term we have

$$\Delta_i \mathcal{D}S^{12} = \Delta_i S^{12} - \mathcal{T}S/(tn^{3/2}) = 2X_{uv} - 2Q^2 n^{-3} = O(\delta_X x) = \widetilde{O}(\hat{q}^{3/2} n^{3/4}).$$

The required bounds on N_S and Var_S hold easily. ∎

Having verified the trend and boundedness hypotheses in Lemmas 4.6 and 4.7, Theorem 4.1 now follows from Lemmas 2.2 and 3.9.

5 | THE CONTROLLABLE ENSEMBLE

In this section we prove that all variables $V = X_{\phi,J,\Gamma}$ in the controllable ensemble have the desired concentration, assuming that all variables in all ensembles are well-behaved at earlier times. Recall that I_{con} is the minimum of the stopping times I_V over all variables V in the controllable ensemble. The following theorem bounds the probability that we reach the universal stopping time I before step $i_{\max}$ because some controllable variable V is good (see Definition 2.11) but fails to satisfy the required bound $|\mathcal{D}V| \leq \delta_V v$.

Theorem 5.1. *With high probability we do not have $I = I_{con} \leq i_{\max}$.*

As with the other ensembles, we prove this theorem by applying the strategy set forth in Section 2.1. In particular, we apply Lemma 2.2.

Recall our convention that the phrase "controllable variable" is in reference to controllability at a particular time t. Thus, we reach the stopping time I_{con} at time $t \geq 1$ only if we violate the bound $|\mathcal{D}V| \leq \delta_V v$ for some variable V in the controllable ensemble that is controllable at time t. Throughout this section, and without further comment, we restrict our attention to times t such that $t \leq 1$ or $t \geq 1$ and the controllable variable V is controllable at time t.

5.1 | Preliminaries

We start by recalling the definition of the ensemble. We say $V = X_{\phi,J,\Gamma}$ is *controllable at time* t' if $o(V) > 0$ and for any $1 \leq t \leq t'$ we have

$$S_A^B(J,\Gamma) \geq n^{\delta'} \quad \text{for all} \quad A \subsetneq B \subseteq V_\Gamma. \tag{26}$$

The *controllable ensemble* consists of all such V with $|V_\Gamma| \leq M^3$ that are controllable at time 1.

Next we record some preliminary observations.

Lemma 5.2. *Let V be controllable at time t'. Then $v(t) \geq n^{\delta'}$ for $1 \leq t \leq t'$ and $S_A^B(J,\Gamma) \geq n^{\delta'}$ for all $A \subsetneq B \subseteq V_\Gamma$ and $t_V \leq t \leq t'$. Furthermore, if V^+ is obtained from V by changing some edge to an open pair then V^+ is controllable at time t'.*

Proof. The first inequality is immediate from the definition with $B = V_\Gamma$. The final statement holds as for any $A \subsetneq B \subseteq V_\Gamma$ we have $S_A^B(V^+) = S_A^B(V)$ or $S_A^B(V^+) = \hat{q}p^{-1}S_A^B(V) \geq S_A^B(V)$, using (12). For the remaining inequality, consider any $A \subsetneq B \subseteq V_\Gamma$. By (26) at $t = 1$ we have

$$n^{|B|-|A|}(2n^{-1/2})^{|J[B]|-|J[A]|} \geq n^{\delta'},$$

so $|J[B]| - |J[A]| < 2(|B| - |A|)$. This gives the much stronger bound

$$n^{|B|-|A|}(n^{-1/2})^{|J[B]|-|J[A]|} \geq n^{1/2},$$

so for $t_V \leq t \leq 1$, recalling from Lemma 2.10 that $t_V = \widetilde{\Theta}(n^{-\delta/4e(V)})$, we have $S_A^B(J,\Gamma) = \Omega\left(n^{|B|-|A|}(t_V n^{-1/2})^{|J[B]|-|J[A]|}\right) = \widetilde{\Omega}\left(n^{1/2-\delta/4}\right) > n^{1/4}$. ∎

It will be convenient to approximation V by the following modified variable V^* which has better behavior for the martingale arguments.

Definition 5.3. Consider $V = X_{\phi,J,\Gamma}$ in the controllable ensemble. Given an injective map $f : V_\Gamma \to [n]$, we say that a pair ab in $f(V_\Gamma)$ is *f-open* if there is no vertex c such that ac, bc are edges and $c \notin f(V_\Gamma)$; note that it is the last condition that distinguishes the definition from that of "open.' Let $V^* = X^*_{\phi,J,\Gamma}(i)$ be defined in the same way as $X_{\phi,J,\Gamma}(i)$, except that pairs that are required to be open in $X_{\phi,J,\Gamma}(i)$ are only required to be f-open in $X^*_{\phi,J,\Gamma}(i)$.

We will apply our usual martingale strategy to show whp $V^* = (1 \pm \delta_{V^*})v$ for $i_V \leq i < I$, where $\delta_{V^*} = \delta_V - g_V/2 = f_V + 3g_V/2$; we recall

$$e = \hat{q}^{-1/2}n^{-1/4}, \quad f_V = e^\delta \quad \text{and} \quad g_V = \vartheta L^{-1}(1 + t^{-e(V)})e^\delta.$$

This will suffice in combination with the following straightforward approximation of V by V^*.

Lemma 5.4. *If $i_V \leq i < I$ then $V = V^* \pm g_V v/2$.*

Proof. Fix $e \in \binom{V_\Gamma}{2} \setminus \Gamma$ with e not contained in the base A. Let $J^e = J \cup \{e\}$ and $\Gamma^e = \Gamma \cup \{e\}$. We bound $|V - V^*|$ by the sum over all such e of X_{ϕ,J^e,Γ^e}. As $i < I$, by property (iv) of Definition 2.12 we have $X_{\phi,J^e,\Gamma^e} \leq L^{4|V_\Gamma|} S_A^{V_\Gamma}(J^e,\Gamma^e)/S_A^B(J^e,\Gamma^e)$, where B is chosen to minimize $S_A^B(J^e,\Gamma^e)$. For any $A \subseteq B \subseteq V_\Gamma$, if $B = A$ then $S_A^B(J^e,\Gamma^e) = 1$; otherwise, by controllability $S_A^B(J^e,\Gamma^e) \geq pS_A^B(J,\Gamma) \geq pn^{\delta'}$. As $S_A^{V_\Gamma}(J^e,\Gamma^e) = pv$, it follows that $X_{\phi,J^e,\Gamma^e} \leq L^{4|V_\Gamma|} vn^{-\delta} \ll g_V v$, as $e^\delta > n^{-\delta/4}$. ∎

5.2 | Decomposition by pairs

We decompose the one-step change in V^* as

$$\Delta_i(V^*) = \sum_{e\in\Gamma\setminus\Gamma[A]} \Delta_i(V^e) \pm F_i(V^*),$$

where each $\Delta_i(V^e)$ accounts for the change in V due to e, as follows. If $e \in J$ then, letting V^+ be obtained from V by changing e from an edge to an open pair, $\Delta_i(V^e)$ is the number of embeddings $f \in (V^+)^*$ such that $f(e)$ is the edge e_{i+1} selected at step $i+1$. If $e \in \Gamma \setminus J$ then $-\Delta_i(V^e)$ is the number of embeddings $f \in V^*$ which are destroyed at step $i+1$ by $f(e)$ not remaining f-open. The fidelity term $F_i(V^*)$ is to correct for embeddings $f \in V^*$ where $f(e)$ is affected for more than one e simultaneously. Note that by definition of "f-open" this cannot occur for creation, that is, if $f(e) = e_{i+1}$ for some $e \in J \setminus J[A]$; thus $F_i(V^*)$ accounts for embeddings $f \in V^*$ where $f(e)$ becomes not f-open for more than one $e \in \Gamma \setminus \Gamma[A]$. This requires the selected edge e_{i+1} to be xy for some $x \in f(V_\Gamma)$ such that y is a common neighbor of some pair u,v in $f(V_\Gamma)$. As $i < I$, by property (iii) of Definition 2.12 all codegrees are $O(L^4)$, so

$$\mathbb{E}[F_i(V^*) \mid \mathcal{F}_i] = O(L^4)v/q. \tag{27}$$

We also decompose the one step change in the tracking variable as

$$\Delta_i(\mathcal{T}V^*) = \sum_{e\in\Gamma\setminus\Gamma[A]} \Delta_i(\mathcal{T}V^e) \pm H_i(V^*),$$

where $\Delta_i(\mathcal{T}V^e)$ is $\mathcal{T}V/(tn^{3/2})$ if e is an edge or $-\frac{\Delta_i Q}{Q}\mathcal{T}V$ if $e \in \Gamma \setminus J$ if e is open, and the higher-order correction term is

$$H_i(V^*) = O((tn^{3/2})^{-1} + Q^{-1}\Delta_i Q)^2 \mathcal{T}V = O(t^2 + t^{-2})n^{-3}v. \tag{28}$$

Our calculations for the trend and boundedness hypotheses will consider separately each $\Delta_i(\mathcal{D}V^e) := \Delta_i(V^e) - \Delta_i(\mathcal{T}V^e)$.

5.3 | One-step expected changes

Here we estimate the one-step expected change in V^* when it is in its upper critical window.

Lemma 5.5. *If $i_V \leq i < I$ and $\mathcal{D}V^* > (f_V + g_V)v$ then*

$$\mathbb{E}[\Delta_i(V^*) \mid \mathcal{F}_i] \leq (1 + o(1)) \left[\frac{e(V)}{8t^2} \delta_{(V^+)^*} - o(V)(f_V + g_V - 2g_Y) \right] 8tvn^{-3/2}.$$

Proof. We estimate the one-step expected changes $\mathbb{E}[\Delta_i(V^e) \mid \mathcal{F}_i]$ for each $e \in \Gamma \setminus \Gamma[A]$.

We start with creation, that is, the case that $e \in J$ is an edge. As for the global variables, we do not use the critical window assumption or obtain any self-correction term in this calculation. Writing $V^+ = X^*_{\phi,J\setminus e,\Gamma}$, we have

$$\begin{aligned}\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] &= \mathbb{E}[\Delta_i(V^e) - \Delta_i(\mathcal{T}V^e) \mid \mathcal{F}_i]\\ &= 2Q^{-1}(V^+)^* - \mathcal{T}V/(tn^{3/2})\\ &= 2Q^{-1}\mathcal{D}(V^+)^*\\ &\le (1+o(1))t^{-1}\delta_{(V^+)^*}vn^{-3/2}.\end{aligned}$$

In the third equality we used $\mathcal{T}V/(2tn^{3/2}) = \mathcal{T}V^+/Q$ and in the last inequality we estimated $\mathcal{D}(V^+)^*$ using $i_V \le i < I$ and $i_{V^+} \le i_V$ (see Lemma 2.10).

Now we consider destruction, that is, the case that $ab = e \in \Gamma \setminus J$ is open. We have $\mathbb{E}[\Delta_i(V^e) \mid \mathcal{F}_i] = 2Q^{-1}\sum_{f\in V^*}(Y_{f(a)f(b)} + Y_{f(b)f(a)} \pm O(1))$, where the $O(1)$ term corrects for the difference between "open" and "f-open" and also for the possibility that $f(ab)$ may become selected rather than closed. Then, recalling (17), we have

$$\begin{aligned}\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] &= -2Q^{-1}\sum_{f\in V}(1 + Y_{f(a)f(b)} + Y_{f(b)f(a)} \pm O(1)) - \mathbb{E}\left[\frac{\Delta_i(Q)\mathcal{T}V}{Q} \mid \mathcal{F}_i\right]\\ &= -4Q^{-1}V(\mathcal{T}Y \pm \delta_Y y) \pm O(q^{-1}v) + Q^{-1}\mathcal{T}V(2+4S/Q)\\ &= -(1 \pm (1+o(1))\delta_Y)8tn^{-3/2}V \pm O(q^{-1}v) + (1 \pm O(\delta_S))8tn^{-3/2}\mathcal{T}V\\ &= -8tn^{-3/2}\mathcal{D}V \pm (1+o(1))8tn^{-3/2}\delta_Y V \pm O(\delta_S 8tn^{-3/2}v) \pm O(q^{-1}v)\\ &\le -\left[(1+o(1))(f_V+g_V) - 2g_Y\right]8tvn^{-3/2}.\end{aligned}$$

In the above calculation we note that we can afford to approximate the multipliers of V and $\mathcal{T}V$ independently as our approximations for controllable variables are weaker than those in the other ensembles. The approximations of Y and S hold for all $n^{5/4} \le i < I$; we also used $f_Y + f_S = o(f_V)$ and $g_S = \widetilde{O}(1+t^{-1})e^2 = o(g_V)$, which holds as

$$(1+t^{-1})e^2 = O(e) \text{ for } t \ge n^{-1/4}. \tag{29}$$

The lemma follows by summing the creation estimate over $e(V)$ edges and the destruction estimate over $o(V)$ open pairs. The $o(1)$ terms absorb the corrections of $O(L^4)v/q$ for fidelity (see (27)) and $O(t^2+t^{-2})n^{-3}v$ for higher-order terms (see (28)), ∎

5.4 | Trend hypothesis and variation equation

The following lemma establishes the trend hypothesis, that is, that $\mathcal{Z}V^* = \mathcal{D}V^* - (f_V + 3g_V/2)v$ is a supermartingale when V^* is in its upper critical window; we will see that this is valid under the choice $c_V = 1$ made in Definition 2.8.

Lemma 5.6. *If $i_V \le i < I$ and $\mathcal{D}V^* > (f_V + g_V)v$ then $\mathbb{E}[\Delta_i\mathcal{Z}V^* \mid \mathcal{F}_i] \le 0$.*

Proof. By Lemma 2.15 (replacing $2g_V$ by $\frac{3}{2}g_V$ to adjust for V^*) we have

$$\Delta_i(v\delta_{V^*}) = \left(\tfrac{e(V)}{8t^2} - o(V)\right)\delta_{V^*}v \cdot 8tn^{-3/2} + \delta'_{V^*}vn^{-3/2} + O(\delta_{V^*}v)n^{-5/2}, \text{ where}$$

$$\delta'_{V^*} \ge 4\delta t\delta_{V^*} + (\vartheta'/\vartheta - e(V)t^{-1})\tfrac{3}{2}g_V.$$

Since

$$c_V = 1 \tag{30}$$

for all V in the controllable ensemble, we have $f_{V^{+*}} = f_{V^*}$, so $\delta_{V^{+*}} - \delta_{V^*} = (3/2)(g_{V^+} - g_V)$.

There is no V^+ term if $e(V) = 0$, and otherwise $\frac{g_{V^+}}{g_V} = \frac{t^{e(V)}+t}{t^{e(V)}+1} < 2t$, so by Lemma 5.5

$$\begin{aligned}
\frac{\mathbb{E}[\Delta_i(\mathcal{Z}V) \mid \mathcal{F}_i]}{8tvn^{-3/2}} &\leq (1+o(1))\frac{e(V)}{8t^2}\delta_{V^{+*}} - (1+o(1))o(V)(f_V + g_V - 2g_Y)\\
&\quad - \left[\left(\frac{e(V)}{8t^2} - o(V)\right)\delta_{V^*} + \frac{\delta'_{V^*}}{8t}\right] + \widetilde{O}(\delta_{V^*}t^{-1}n^{-1})\\
&\leq \frac{e(V)}{8t^2}\cdot\frac{3}{2}(g_{V^+} - g_V) + \frac{o(V)}{2}g_V - \frac{1}{2}\delta f_V\\
&\quad - \left(\frac{\vartheta'}{8t\vartheta} - \frac{e(V)}{8t^2} + \frac{\delta}{2}\right)\cdot\frac{3g_V}{2} + O(g_Y) + o(\delta_{V^+}/t^2) + o(\delta_V)\\
&\leq \frac{g_V}{2}\left(o(V) - \frac{3\delta}{2} + \frac{3e(V)}{4t} - \frac{3\vartheta'}{8t\vartheta} + o(1)\right) - \frac{1}{4}\delta f_V + O(g_Y) + o(\delta_{V^+}/t^2).
\end{aligned}$$

For the last inequality, we have cancelation of two terms $\frac{e(V)g_V}{8t^2}$ with opposite signs, and we used $g_{V^+} \leq 2tg_V$. Finally, $\mathbb{E}[\Delta_i(\mathcal{Z}V) \mid \mathcal{F}_i] \leq 0$, as the dominant terms are $-\frac{3\vartheta'}{16t\vartheta}g_V$ and/or $-\frac{1}{4}\delta f_V$. ∎

5.5 | Boundedness hypothesis

For the boundedness hypothesis, we fix any controllable $V = X_{\phi,J,\Gamma}$ and estimate $\text{Var}_{V^*} = \text{Var}(\mathcal{Z}V^*(i) \mid \mathcal{F}_{i-1})$ and $N_{V^*} = |\Delta_i \mathcal{Z}V^*|$. Recall that it suffices to establish (2) and (3), as in the following lemma. We remark that the proof of the "boundary case" $|V_\Gamma| = M^3$ is quite delicate, and it is here that the details of property (v) in Definition 2.12 are important.

Lemma 5.7. *If $i_V \leq i < I$ and V is good and controllable (at time t) then* $\text{Var}_{V^*} = o\left(\frac{(t^{-e(V)}e^{\delta}v)^2}{L^3n^{3/2}}\right)$ *and* $N_{V^*} = o\left(\frac{t^{-e(V)}e^{\delta}v}{L^2}\right)$.

Proof. Recalling that we restrict our attention to $t \geq t_V$, we can bound the one-step change in $\mathcal{T}V^* + (f_V + 3g_V/2)v$ by $O((t + t^{-1})vn^{-3/2}) = \widetilde{O}(vn^{-5/4})$, which is negligible in comparison with the required estimates. It therefore suffices to consider changes in V^* rather than $\mathcal{Z}V^*$. As in the trend hypothesis, we can obtain these estimates as a sum over all $e \in \Gamma \setminus \Gamma[A]$. (Here we use $|V_\Gamma| \leq M^3 = O(1)$ and the simple observation that if random variables A and B each have variance at most σ^2 then $A + B$ has variance at most $4\sigma^2$.)

Thus for each $e = \alpha\beta \in \Gamma \setminus \Gamma[A]$ we estimate $N_e = |\Delta_i V^e|$ and $\text{Var}_e = \text{Var}(\Delta_i V^e \mid \mathcal{F}_{i-1})$. (We drop the stars in the superscripts in an attempt to reduce notational clutter.)

We start with the creation calculation, that is, the case $e \in J$. All scalings here will be with respect to the extension $(\phi, J \setminus e, \Gamma)$ obtained by changing e to an open pair: for example, $S_A^{V_\Gamma} = v\hat{q}p^{-1}$. Let $A' = A \cup \{\alpha, \beta\}$, where $A \subseteq V_\Gamma$ is the base of the extension. We note that if $\Delta_i V^e \neq 0$ then for any B with $A' \subseteq B \subseteq V_\Gamma$ the edge e_{i+1} selected at step $i+1$ must fall in some extension in $X_{\phi,(J\setminus e)[B],\Gamma[B]}$. We consider the "hardest" such extension: let $S_{\text{m}} = \min_{A' \subseteq B \subseteq V_\Gamma} S_A^B$. Let B_{m} be some set B achieving the minimum in this definition. We note that

(B1) $\max_{A' \subseteq C \subseteq B_{\text{m}}} S_C^{B_{\text{m}}} = 1$,
(B2) $\max_{B_{\text{m}} \subseteq C \subseteq V_\Gamma} S_C^{V_\Gamma} = S_A^{V_\Gamma}/S_{\text{m}}$,

(B3) $S_m \geq n^{\delta'}(\hat{q}/p)$,
(B4) $\max_{A\subseteq C\subseteq B_m} S_C^{B_m} = S_m$.

Indeed, (B1) and (B2) follow from the definition of B_m, and (B3) and (B4) from controllability of V. By property (iv) of Definition 2.12 applied to the extension from A to B_m (and (B4)) we estimate

$$p_e := \mathbb{P}[\Delta_i V^e \neq 0] < L^{4|V_\Gamma|} S_m/q.$$

Also, applying property (iv) of Definition 2.12 to the extensions from A' to B_m (using (B1)) and from B_m to V_Γ (using (B2)), we estimate

$$N_e < L^{4|V_\Gamma|} \cdot L^{4|V_\Gamma|} S_A^{V_\Gamma}/S_m \leq L^{8|V_\Gamma|} n^{-\delta'} v,$$

using (B3) for the second inequality. Then

$$\mathrm{Var}_e < p_e N_e^2 < L^{20|V_\Gamma|}(S_m/q)(S_A^{V_\Gamma}/S_m)^2 = L^{20|V_\Gamma|}(\hat{q}/p)^2 v^2/(qS_m) < L^{20|V_\Gamma|}(2tn^{3/2})^{-1} n^{-\delta'} v^2.$$

Noting that creation only occurs when $e(V) \geq 1$, these estimates are well within the required bounds, as $e^{\delta} > n^{-\delta/4}$ and $\delta \ll \delta'$.

It remains to consider destruction, that is, the case $e = \alpha\beta \in \Gamma \setminus J$. Let (A', J', Γ') be obtained from (A, J, Γ) by "gluing a Y-variable on $\alpha\beta$" as follows. Let γ be a new vertex, $V' = V_\Gamma \cup \{\gamma\}$, $A' = A \cup \{\alpha, \gamma\}$, $J' = J \cup \{\beta\gamma\}$ and $\Gamma' = \Gamma \cup \{\alpha\gamma, \beta\gamma\}$ (so this definition depends on the order of α and β). To analyze destruction of extensions $f \in V^*$ due to closures of e by selecting the edge corresponding to $\alpha\gamma$, we consider extensions in $X_{\phi', J', \Gamma}$ where $\phi' : A' \to [n]$ restricts to ϕ on A and $\phi'(\alpha\gamma)$ is the edge e_{i+1} added at step $i+1$. In only considering the case that γ is a new vertex we make crucial use of the distinction between V^* and V.

As in the creation calculation, we consider the "hardest" extension that includes the $\alpha\gamma$. We set $S_m = S_A^{B_m} = \min_{A' \subseteq B \subseteq V'} S_A^B$, where all scalings are with respect to the pair (J', Γ'), and note that conditions (B1) and (B4) hold. Also note that in place of the condition (B2) we have

(B2′) $vy/S_m = S_A^{V'}/S_m = \max_{B_m \subseteq C \subseteq V'} S_C^{V'}$.

If $|B_m| = M^3 + 1$ then we have the simple bound $p_e := \mathbb{P}[\Delta_i V^e \neq 0] \leq 2yv/q$. If $|B_m| \leq M^3$, applying property (iv) of Definition 2.12 and (B4) we have $p_e < L^{4|V_\Gamma|} S_m/q$. Thus, in either case we have

$$p_e := \mathbb{P}[\Delta_i V^e \neq 0] < L^{4|V_\Gamma|} S_m/q.$$

We claim that

$$S_m \geq yn^{\delta'}. \tag{31}$$

To see this, note that if $B_m = A \cup \{\gamma\}$ then $S_m = \hat{q}n \geq yn^{\delta'}$. Otherwise, we write $S_A^{B_m} = S_{B_m\setminus\gamma}^{B_m} S_A^{B_m\setminus\gamma}$. We have $S_{B_m\setminus\gamma}^{B_m} \geq y$ by construction of (J', Γ') and $S_A^{B_m\setminus\gamma} \geq n^{\delta'}$, since V is controllable. This proves the claim.

Now we claim that the magnitude of the change due to e is bounded as

$$N_e < 2L^{12|V'|+7} vy/S_m. \tag{32}$$

The lemma follows from this bound; indeed, substituting (31) gives $N_e = \widetilde{O}(n^{-\delta'} v)$ and

$$\mathrm{Var}_e < p_e N_e^2 = \widetilde{O}(S_{\mathrm{m}}/q)(yv/S_{\mathrm{m}})^2 = \widetilde{O}(y^2 v^2/qS_{\mathrm{m}}) = \widetilde{O}(n^{-\delta'} v^2 n^{-3/2}).$$

Thus it remains to prove (32).

First we note that the argument we used for creation establishes (32) if we are not at the boundary of the ensemble, that is, if $|V_\Gamma| < M^3$, so $|V'| \le M^3$. Indeed, applying property (iv) of Definition 2.12 to the extensions from A' to B_{m} (using (B1)) and from B_{m} to V' (using (B2′)), we estimate

$$N_e < L^{4|V'|} \cdot L^{4|V'|} S_A^{V'}/S_{\mathrm{m}} = L^{8|V'|} yv/S_{\mathrm{m}},$$

as desired.

It remains to establish (32) in the boundary case $|V_\Gamma| = M^3$. Note that we still have at most $L^{4|V'|}$ extensions from A' to B_{m}, using property (v) of Definition 2.12 if $B_{\mathrm{m}} = V'$. As this observation establishes (32) in the case $B_{\mathrm{m}} = V'$, we henceforth assume $B_{\mathrm{m}} \subsetneq V'$. Next we consider the extension series from B_{m} to V' and let $C \subsetneq V'$ be the set preceding V'. We claim that if $\beta \in C$ then we can still implement the above bound using extensions on at most M^3 vertices, so that property (iv) of Definition 2.12 still applies. Indeed, writing $C^- = C \setminus \{\gamma\}$ we have

$$yv/S_{\mathrm{m}} = S_A^{V'}/S_{\mathrm{m}} = S_{B_{\mathrm{m}}}^{V'} = S_{B_{\mathrm{m}}}^{C} S_C^{V'} = S_{B_{\mathrm{m}}}^{C} S_{C^-}^{V_\Gamma},$$

so considering extensions from C^- to V_Γ we still have at most $L^{4|V'|} S_A^{V'}/S_{\mathrm{m}}$ extensions from B_{m} to V', and (32) follows.

Now we may assume $\beta \notin C$. We can also assume $S_C^{V'} \ge y/L^7$, otherwise we can still implement the previous calculation using property (v) of Definition 2.12. On the other hand, by definition of the extension series we have $S_C^{V'} \le S_C^{C\cup\beta} \le y$, as the extension from C to $C \cup \beta$ contains the edge $\beta\gamma$ and the open pair $\alpha\beta$. Thus $S_C^{V'} \le S_C^{C\cup\beta} \le L^7 S_C^{V'}$ and we give up a factor of at most L^7 if we can count extensions from C to V' by first bounding the number of extensions from C to $C \cup \beta$ with a Y variable and then counting extensions from $C \cup \beta$ to V'. Furthermore, $S_{C\cup\beta}^{V'} \le 1$ by definition of C, and we can estimate extensions from $C \cup \beta$ to V' using extensions from C to V_Γ, since $S_{C\cup\beta}^{V'} = S_C^{V_\Gamma}$. This gives

$$N_e < L^{4|B_{\mathrm{m}}|} \cdot L^{4|C|} S_{B_{\mathrm{m}}}^{C} \cdot 2y \cdot L^{4|V_\Gamma|} < 2L^{12|V'|+7} S_A^{V'}/S_{\mathrm{m}},$$

which completes the proof of the claim (32), and so of the lemma. ∎

Now that we have verified the trend and boundedness hypotheses for V^*—these are Lemmas 5.6 and 5.7, respectively—we can apply Lemma 2.2 and show whp $V^* = (1 \pm \delta_{V^*})v$ for $i_V \le i < I$. (Recall that Lemma 3.9 establishes this condition at step i_V.) In combination with Lemma 5.4 this proves Theorem 5.1.

6 | STACKING ENSEMBLE

In this section we prove that all variables in the stacking ensemble have the desired concentration, assuming that all variables in all ensembles are well-behaved at earlier times. Recall that I_{stk} is the minimum of the stopping times I_V over all variables V in the stacking ensemble. The following theorem bounds the probability that we reach the universal stopping time I before step $i_{\max}$ because some stacking variable V is good (see Definition 2.11) but fails to satisfy the required bound $|\mathcal{D}V| \le \delta_V v$.

Theorem 6.1. *With high probability we do not have $I = I_{stk} \leq i_{\max}$.*

As for the other ensembles, we will prove this theorem by verifying the trend and boundedness hypotheses and applying Lemma 2.2.

Throughout the section we consider some stacking variable $V = S^{\pi}_{uv} = X_{\phi,J,\Gamma}$, for some non-edge uv, where we recall that $V(\Gamma) = V(S^{\pi}_{uv}) = \{\alpha_u, \alpha_v, \alpha_1, \dots, \alpha_{|\pi|}\}$, $A = \{\alpha_u, \alpha_v\}$, $\phi(\alpha_u) = u$, $\phi(\alpha_v) = v$ and (J, Γ) is defined so that edges specified by the extension are mapped to edges of $G(i)$, and likewise for open pairs. Recalling that we gave a separate argument for vertex degree variables in Lemma 3.15, we can assume V is not such a variable. Similarly to the analysis of controllable variables (except that here we do not approximate V by V^*), we decompose the one-step change in V as

$$\Delta_i(V) = \sum_{e \in \Gamma \setminus \Gamma[A]} \Delta_i(V^e) \pm F_i(V),$$

where each $\Delta_i(V^e)$ accounts for the change in V due to e, as follows. If $e \in J$ then, letting V^+ be obtained from V by changing e from an edge to an open pair, $\Delta_i(V^e)$ is the number of embeddings $f \in V^+$ such that $f(e)$ is the edge e_{i+1} selected at step $i + 1$. If $e \in \Gamma \setminus J$ then $-\Delta_i(V^e)$ is the number of embeddings $f \in V$ which are destroyed at step $i + 1$ by $f(e)$ being selected or closed. The fidelity term $F_i(V)$ corrects for embeddings $f \in V$ where $f(e)$ is affected for more than one e simultaneously (see Section 6.4).

6.1 | Subextensions of stacking variables

This subsection concerns certain subextensions of stacking variables that will be particularly important throughout this section. For the following two special structures we will appeal to the controllable ensemble for our estimates, and so we need to show that these extensions are indeed controllable at time $t_{\max}$ (so that we can apply our bounds on controllable variables throughout the process).

- Let (uv, J, Γ) be the extension corresponding to some stacking sequence $\pi \in \mathcal{S}_M$ at the boundary of the ensemble, that is, with $w(\pi) = 2M$. Let $y = |\pi|$ and let $\alpha_x\alpha_y \in \Gamma \setminus J$ (so $\alpha_x\alpha_y$ is an open pair that contains the final vertex in stacking order). The *backward extension* B_π is the extension (A', J', Γ') with $A' = \{\alpha_u, \alpha_v, \alpha_x, \alpha_y\}$, $J' = J$ and $\Gamma' = \Gamma \setminus \alpha_x\alpha_y$.
- An *h-fan* at the triple $A = abc$ is any extension of the form (A, J, Γ), where the base is $A = abc$, there are h additional vertices $v_1, \dots, v_h$ in V_Γ, the sequence $bv_1 \dots v_hc$ is a path of length $h + 1$ in Γ, and $av_i \in \Gamma \setminus J$ is open for $i \in [h]$. We emphasize that the pairs in the path $bv_1 \dots v_hc$ can be either edges or open pairs.

Both of these extensions arise from the boundary conditions in our choice to restrict the stacking ensemble to M-bounded variables. Recalling Definition 2.5, we need to consider backward extensions due to condition (i) that $w(\pi) \leq 2M$ and fans due to condition (ii) forbidding a subsequence of length M using only $\{X^I, Y^I\}$: in both cases there is at least one direction in which we cannot stack Y on the last rung.

Now we show that these two extensions are controllable. We recall that $M = 3/\varepsilon$ and $\hat{q}(t_{\max}) = n^{-1/2+\varepsilon}$.

Lemma 6.2. *All M-fans and backward extension variables are controllable at time $t_{\max}$.*

Proof. We start by considering an M-fan (A, J, Γ). Among all such extensions, the minimum scaling is $(\hat{q}n)^M p^{M+1} > n^{\varepsilon M - 1/2} = n^{5/2}$, which is achieved when the path $bv_1 \dots v_Mc$ belongs entirely to J. Fix

B with $A \subsetneq B \subseteq V$ that minimizes $S_A^B = S_A^B(J,\Gamma)$. We need to show that $S_A^B \geq n^{\delta'}$. As $S_A^{V_\Gamma} > n^{5/2}$ we can assume that $B \neq V_\Gamma$, so we can find v_i in B such that not both v_{i-1} and v_{i+1} are in B. (Here $v_0 = c$ and $v_{M+1} = b$.) Now removing v_i from B reduces the scaling by at least $y > \hat{q}n^{1/2} = n^{\varepsilon}$, so by minimality we have $|B| = |A| + 1$, so $S_A^B \geq y > n^{\varepsilon} > n^{\delta'}$ (recalling (8)).

Now consider (with notation as above) a backward extension $B_\pi = (A', J', \Gamma')$ with $w(\pi) = 2M$. We fix B with $A' \subsetneq B \subseteq V$ and estimate $S_{A'}^B$ as a sequence of single-vertex extensions. First we consider the case that there is some $T \subseteq V$ disjoint from B such that some component C of $\Gamma' \setminus T$ contains $\{\alpha_x, \alpha_y\}$, but not α_u or α_v. Then we consider vertices of $B \setminus C$ in stacking order and vertices of $B \cap C$ in reverse stacking order. Each step contributes a factor of at least $y > n^{\varepsilon}$ to the scaling, so $S_{A'}^B > n^{\varepsilon} > n^{\delta'}$.

Now we can assume there is no such T, which implies that B intersects every rung and contains all α_i such that $\pi(i+1) = O$. We claim that $|B| \geq M + 2$. We note that this will imply the lemma, as estimating S_{uv}^B by a sequence of single-vertex extensions gives

$$S_{A'}^B = S_{uv}^B/(n^2\hat{q}) > (n^{\varepsilon})^{|B|-2}/n^2 \geq (n^{\varepsilon})^{|M|}/n^2 = n > n^{\delta'}.$$

It remains to show the claim. We bound the intersection of B with the set of $2M$ vertices that contribute to $w(\pi)$. Suppose π has i occurrences of the symbol O in the sequence $\pi(2), \dots, \pi(|\pi| - 1)$ and j occurrences of O or E in $\{\pi(1), \pi(|\pi|)\}$. Then there are at most $i + 1$ triangular ladders and π has $2M - i - j$ turning points (recall that the positions with the symbols X^O or Y^O give turning points), of which at most $2 - j$ are in A' (namely α_u and $\alpha_{|\pi|-1}$). Let T be the set of turning points not in A', so that $|T| \geq 2M - i - 2$. For each triangular ladder there is a path of rungs spanned by $T \cap L$, so we must have $|B \cap T \cap L| \geq \lfloor |T \cap L|/2 \rfloor$. We deduce $|B \setminus A'| \geq i + \frac{2M-2-i}{2} - \frac{i+1}{2} \geq M - 2$, which proves the claim, and so the lemma. ∎

Remark 6.3. The proof of Lemma 6.2 shows moreover that to show that a fan of any size is controllable at any time t it suffices to show that it has scaling at least $n^{\delta'}$, as then this will also be true for all subextensions needed for controllability.

6.2 | Boundedness hypothesis

Here we verify the boundedness hypothesis, for which the arguments are somewhat similar to those given above for the controllable ensemble, and are relatively short (the bulk of the section will then be occupied with verifying the trend hypothesis). Recalling (2) and (3), and that $c_V \geq L^{15}$ for all V in the stacking ensemble (see Definition 2.8), it suffices to prove the following lemma.

Lemma 6.4. *If $i_V \leq i < \min\{I, J_V\}$ then*

$$N_V = O\left((1 + t^{-e(V)})ev\right) \quad \text{and} \quad \mathrm{Var}_V = O\left(n^{-3/2}((1 + t^{-e(V)})ev)^2\right).$$

Proof. As in the proof of Lemma 5.7, it suffices to establish the stated bounds for each $e \in \Gamma \setminus \Gamma[A]$ on $N_e = |\Delta_i V^e|$ and $\mathrm{Var}_e = \mathrm{Var}(\Delta_i V^e \mid \mathcal{F}_{i-1})$ (we do not need to take advantage of better bounds available on the change in the difference between these variables and their tracking variables). There are two cases, according to whether e is an open pair or an edge.

We start by considering the case that $e \in J$ is an edge. Let $e = \alpha_x\alpha_y$ where $x < y$. Let $A' = A \cup \{\alpha_x, \alpha_y\}$ and $S_{\mathrm{m}} = \min_{A' \subseteq B \subseteq V} S_A^B = S_A^{B_{\mathrm{m}}}$, where all scalings are with respect to $(J \setminus e, \Gamma)$. As in the proof of Lemma 5.7, the extension to B_{m} is the "hardest" extension that contains the pair $\alpha_x\alpha_y$. We have

(C1) $\max_{A' \subseteq C \subseteq B_m} S_C^{B_m} = 1$,
(C2) $\max_{B_m \subseteq C \subseteq V_\Gamma} S_C^{V_\Gamma} = S_A^{V_\Gamma}/S_m$,
(C3) $\max_{A \subseteq C \subseteq B_m} S_C^{B_m} = S_m$.

Applying part (iv) of Definition 2.12, noting that $S_A^{V_\Gamma} = \hat{q}p^{-1}v$, we have

$$p_e := \mathbb{P}[\Delta_i V^e \neq 0] < L^{4|V_\Gamma|} S_m/q \quad \text{and} \quad N_e < L^{8|V_\Gamma|}\hat{q}p^{-1}v/S_m,$$

Note that we use (C3) to establish the bound on p_e, and we use (C1) and (C2) to establish the bound on N_e. We have

$$\text{Var}_e < p_e N_e^2 < L^{20|V_\Gamma|}(\hat{q}p^{-1}v)^2/(qS_m).$$

We calculate the scaling S_m one vertex at a time, proceeding in stacking order. Each vertex contributes a factor of at least $p\hat{q}n = y$, and α_y contributes at least $\hat{q}^2 n = x$, since the edge $\alpha_x\alpha_y$ was switched to an open pair in $(J \setminus e, \Gamma)$. If $|B_m \setminus A| \geq 2$ we have $S_m \geq xy$, so

$$N_e < L^{8|V_\Gamma|}vp^{-1}\hat{q}/(xy) = t^{-1}ev \cdot L^{8|V_\Gamma|}(4t)^{-1}e^3 \ll t^{-1}ev \text{ and}$$
$$\text{Var}_e < L^{20|V_\Gamma|}(\hat{q}p^{-1}v)^2/(qxy) = n^{-3/2}((2t)^{-1}ev)^2 \cdot y^{-1}L^{20|V_\Gamma|} \ll n^{-3/2}(t^{-1}ev)^2,$$

which are sufficient, as $e \in J$ implies $e(V) \geq 1$.

On the other hand, if $|B_m \setminus A| = 1$, then $B_m = A'$, and this corresponds to the edge $e_{i+1} = u'v'$ added at step $i+1$ playing the role of an edge that intersects $A = \alpha_u\alpha_v$. There are two possibilities for such an edge. If $\pi(1) = Y^I$ or $\pi(1) = Y^O$ then this edge could create this first Y-extension of π or this edge could be the edge of $\pi(j) = Y^O$ where $j > 1$ and $\pi(1), \dots, \pi(j-1) \in \{Y^I, X^I\}$. In the first case, writing π' for the stacking sequence obtained from π by removing $\pi(1)$, and $V' = S_{u'v'}^{\pi'}$ for the corresponding stacking variable based at $u'v'$ (which is open before we add e_{i+1}), we can improve the above bounds to $p_e \leq 2x/q$ and $N_e \leq V' \leq 2v/y$, so $\text{Var}_e \leq 8q^{-1}(t^{-1}v)^2$, which again suffices. In the second case, we have $p_e \leq 2x/q = 2/n$ and we claim that $N_e \leq L^{4M}v/(np)$, which will also be sufficient. We obtain the claimed bound for N_e as the product of bounds for the fan extension from $\{\alpha_u, \alpha_v, \alpha_j\}$ to $\{\alpha_1, \dots, \alpha_{j-1}\}$ and the forward extension from $u'v'$ to the remainder of the stacking variable. The fan extension has scaling at least 1 by definition of B_m, and we include the logarithmic factor in the bound on N_e because we need to apply (iv) of Definition 2.12.iv if the extension to the fan is not controllable. The claim follows.

It remains to consider the changes due to closing some open pair $\alpha_x\alpha_y = e \in \Gamma \setminus J$ (which may be a rung or a stringer). This is described by a structure where for some vertex γ we already have the edge $\alpha_y\gamma$ and then we add the edge $\alpha_x\gamma$. There are two subcases according to whether γ belongs to V_Γ or is a new vertex. In both subcases, we consider $J' = J \cup \{\alpha_y\gamma\}$ and $\Gamma' = \Gamma \cup \{\alpha_x\gamma, \alpha_y\gamma\}$ on the vertex set $V' = V_\Gamma \cup \{\gamma\}$ (which is V_Γ if $\gamma \in V_\Gamma$), we let $A = \{\alpha_u, \alpha_v\}$, $A' = A \cup \{\alpha_x, \gamma\}$ and $S_m = \min_{A' \subseteq B \subseteq V'} S_A^B = S_A^{B_m}$, where all scalings are with respect to (J', Γ').

We begin with the subcase $\gamma \notin V_\Gamma$. Then Γ' is obtained from Γ by adding a Y extension on $\alpha_x\alpha_y$, so $S_A^{V'} = vy$. The analogues of (C1)-(C3) hold, so, as above, we have

$$p_e := \mathbb{P}[\Delta_i V^e \neq 0] < L^{4|V'|} S_m/q \quad \text{and} \quad N_e < L^{8|V'|} S_A^{V'}/S_m,$$

$$\text{so} \quad \text{Var}_e < p_e N_e^2 < L^{20|V'|}(S_A^{V'})^2/(qS_m).$$

If $S_{\rm m} > L^{40|V_\Gamma|}y^2$ (say) then these bounds are easily sufficient. Estimating $S_{\rm m}$ vertex by vertex in the stacking order we see that this holds if $|B_{\rm m} \setminus A| \geq 3$ (when $S_{\rm m} \geq y^3 \gg y^2$) or if $|B_{\rm m} \setminus A| = 2$ and not both steps from A to $B_{\rm m}$ are Y extensions (this gives $S_{\rm m} \geq xy \gg y^2$).

The remaining possibilities in the subcase $\gamma \notin V_\Gamma$ need more precise estimates on N_e and Var_e that avoid the polylogarithmic loss in these crude estimates. Consider the case that $|B_{\rm m} \setminus A| = 2$ and $B_{\rm m}$ is obtained by two Y extensions, so $S_{\rm m} = y^2$. Here we can use stacking variables to estimate p_e and N_e, as $(A, B_{\rm m})$ induces the extension $S_{uv}^{\pi(1)Y'}$, and $N_e \leq S_{\alpha_y\alpha_x}^{\pi'}$, where $\pi = \pi(1)\pi'$. We have the better bounds $p_e < 2S_{\rm m}/q = 2y^2/q$ and $N_e < 2S_A^{V'}/S_{\rm m} = 2v/y$, so $\mathrm{Var}_e < 8v^2/q = 8n^{-3/2}(ev)^2$, which suffices.

Now consider $|B_{\rm m} \setminus A| = 1$, so $\alpha_x \in \{\alpha_u, \alpha_v\}$ and $B_{\rm m} = \{\alpha_u, \alpha_v, \gamma\}$. The extension from A to $B_{\rm m}$ is an open degree, with scaling $S_{\rm m} = x_1 = \hat{q}n$, so we estimate $p_e \leq 2x_1/q = 2/n$. To estimate N_e we consider the extension (A', J', Γ') in two steps, where in the first step we add all vertices in the stacking order up to α_y, and in the second step we add the remaining vertices. Thus we bound $N_e \leq \sum_{f\in V^1} V_f^2$, where V^1 is a fan extension with base A', and V_f^2 is a stacking variable with base $f(\alpha_x\alpha_y)$. The scalings v_1 and v_2 satisfy $v_1v_2 = S_{A'}^{V'} = vy/x_1$. If V^1 is controllable at time t we obtain the required bounds from $N_e < 2v_1 \cdot 2v_2 = 4vy/x_1 = 8tn^{-1/2}v$ and $\mathrm{Var}_e < 2n^{-1}(4vy/x_1)^2 = 32t^2n^{-2}v^2$. Now suppose V^1 is not controllable at time t, so $v_1 < n^{\delta'}$ by Remark 6.3. Note further that $(4t)^2\hat{q} < 1$ for all t. This implies that the fan extension in question is not comprised of a single vertex with two edges and a single open pair. The condition $v_1 < n^{\delta'}$ then implies $\hat{q} < t^{-3/2}n^{\delta'/2-1/4}$, so $S_{\rm m} = \hat{q}n > L^{40|V_\Gamma|}y^2$, and we have already completed the proof when this holds.

Now consider the final subcase, namely $\gamma \in V_\Gamma$. Here we write $S_A^{V'} = vp^\alpha\hat{q}^\beta$ where $\alpha, \beta \in \{0, 1\}$. Note that the analogues of (C1) and (C2) hold, but the analogue of (C3) does not necessarily hold. As $S_{\rm m} \geq yp^\alpha\hat{q}^\beta$ we deduce

$$N_e < L^{8|V_\Gamma|}vp^\alpha\hat{q}^\beta/S_m \leq evt^{-1} \cdot L^{8|V_\Gamma|}e/2,$$

which is sufficient, noting that we can improve the bound on S_m to $S_m \geq xp^\alpha\hat{q}^\beta$ in the case $e(V) = 0$.

It remains to bound Var_e in the case $\gamma \in V_\Gamma$. If we have $S_A^C \geq 1$ for all $A \subseteq C \subseteq B_{\rm m}$ then the analogue of (C3) holds, so $p_e < L^{4|V_\Gamma|}S_m/q$ and

$$\mathrm{Var}_e < L^{20|V_\Gamma|}v^2/qy = n^{-3/2}(ev)^2t^{-1} \cdot L^{20|V_\Gamma|}e^2/2.$$

This bound suffices (again appealing to the improvement in the bound on S_m if $e(V) = 0$). If $S_{\rm m} < 1$ and $S_A^C \geq S_{\rm m}$ for all proper subsets C of $B_{\rm m}$ we have $p_e < L^{4|V_\Gamma|}/q$ and $\mathrm{Var}_e < L^{20|V_\Gamma|}v^2/(qy^2)$, which is sufficient. Finally, suppose there is a proper subset C of $B_{\rm m}$ such that $S_A^C < 1, S_{\rm m}$. Note that we have $\alpha = 1$, corresponding to an extra edge in $J' \setminus J$, and C must contain this extra edge. Writing $S'_{\rm m} = \max_{A\subsetneq C\subsetneq B_{\rm m}} S_C^{B_{\rm m}}$, we have $p_e < L^{8|V_\Gamma|}S'_{\rm m}/q$ and

$$\mathrm{Var}_e < L^{8|V_\Gamma|}S'_{\rm m}/q \cdot L^{16|V_\Gamma|}(S_A^{V'}/S_{\rm m})^2 < L^{|24|V_\Gamma|}S'_{\rm m}/q \cdot (S_A^{V'}/(ypS'_{\rm m}))^2 < L^{|24|V_\Gamma}v^2/(y^2q),$$

which again suffices. ■

6.3 | Tracking variables

Here we will recall and explain in more detail the definition of the tracking variables $\mathcal{T}V$ in Section 2.3. We also describe the pair decomposition of their one step changes. There will be two cases for $V = S_{uv}^\pi$ depending on the form of π.

6.3.1 | Standard tracking variables

The first case, which we call *standard*, is that $\pi(|\pi|-1) \neq O$ or $\pi(|\pi|) \in \{O, E\}$. We write $\pi = \pi^- \circ U$, where U is the last element of π, and let

$$\mathcal{T}V = V^-\mathcal{T}U, \text{ where } V^- = S_{uv}^{\pi^-}.$$

Note that this choice of $\mathcal{T}V$ isolates variations that are not caused by variations in V^-.

We say that a pair e is *terminal* if it belongs to U, that is, it contains the final vertex of V; otherwise we say that e is *internal*. We write

$$\Delta_i(\mathcal{T}V) = \Delta_i(V^-)\mathcal{T}U + V^-\Delta_i(\mathcal{T}U) = \sum_{e \in \Gamma \setminus \Gamma[A]} \Delta_i(\mathcal{T}V^e) + H_i(V), \tag{33}$$

where similarly to (28) the higher-order correction term is

$$H_i(V) = O(t^2 + t^{-2})n^{-3}v, \tag{34}$$

and $\Delta_i(\mathcal{T}V^e)$ is defined as follows.

(i) If e is a terminal edge then $\Delta_i(\mathcal{T}V^e) = \frac{\mathcal{T}V}{tn^{3/2}}$,
(ii) If e is a terminal open pair then $\Delta_i(\mathcal{T}V^e) = \frac{\Delta_i(Q)}{Q}\mathcal{T}V$,
(iii) If e is internal then $\Delta_i(\mathcal{T}V^e) = \Delta_i((V^-)^e)\mathcal{T}U$.

Note that (iii) uses the definition of $\Delta_i(V^e)$ above with V^- in place of V.

6.3.2 | Partner tracking variables

The other case, which we call *partner*, is that $\pi(|\pi|-1) = O$ and $\pi(|\pi|) \notin \{O, E\}$. We must have $|\pi| \geq 2$, and the vertices $\{\alpha_{|\pi|-2}, \alpha_{|\pi|-1}, \alpha_{|\pi|}\}$ form a triangle in $V = S_{uv}^{\pi}$, in which at most one pair is an edge and the other pairs are open. We say that the open pair $\alpha_{|\pi|-2}\alpha_{|\pi|-1}$ and the pair $\alpha_{|\pi|-2}\alpha_{|\pi|}$ (which can be an edge or an open pair) are *partner pairs*; it is natural to treat them together because of the "symmetry" interchanging $\alpha_{|\pi|-1}$ and $\alpha_{|\pi|}$ (although it can be that one is an edge and the other is open). The pair $\alpha_{|\pi|-1}\alpha_{|\pi|}$ is still called terminal; its treatment is exactly as in (i) and (ii) above.

We emphasize that we do not consider partner pairs to be terminal, even though one of them uses the last vertex of V. We also do not consider partner pairs to be internal.

We write $\pi = \pi^- OU$, $V^- = S_{uv}^{\pi^-}$, $\beta = \alpha_{|\pi|-2}$ and let $\mathcal{T}V = \sum_{f \in V^-} X_{f(\beta)}\hat{U}_f$, where

$$\hat{U}_f = \begin{cases} X_{f(\beta)} \cdot Qn^{-2} & \text{if } U \in \{X^I, X^O\} \\ X_{f(\beta)} \cdot 2tn^{-1/2} & \text{if } U = Y^I \\ Y_{f(\beta)} \cdot Qn^{-2} & \text{if } U = Y^O. \end{cases}$$

To interpret this formula, note that for each $f \in V^-$ we are approximating the number of choices for the three remaining edges as if they were independent events: for the partner pairs we include a degree or open degree factor $Y_{f(\beta)}$ for an edge or $X_{f(\beta)}$ for an open pair, and for the terminal pair we include a probability factor of Qn^{-2} for an open pair or $2tn^{-1/2}$ for an edge.

We unify the two definitions of $\mathcal{T}V$ by writing

$$\mathcal{T}V = \sum_{f \in V^-} \mathcal{T}_f V, \text{ where } \mathcal{T}_f V = \mathcal{T}U \text{ if } \pi = \pi^- U \text{ or } \mathcal{T}_f V = X_{f(\beta)}\hat{U}_f \text{ if } \pi = \pi^- OU. \tag{35}$$

We keep the same definition as in points (i) and (ii) above of $\Delta_i(\mathcal{T}V^e)$ for terminal pairs, and extend it to internal pairs (consistently with (iii) above) and partner pairs as follows.

(iii) If e is an internal edge then $\Delta_i(\mathcal{T}V^e) = \sum_{f\in V^{-+}} I_f^e \mathcal{T}_f V$, where V^{-+} is obtained from V^- by changing e to an open pair and I_f^e is the indicator of the event that $e_{i+1} = f(e)$.
If e is an internal open pair then $\Delta_i(\mathcal{T}V^e) = \sum_{f\in V^-} I_f^e \mathcal{T}_f V$, where I_f^e is the indicator of the event that e_{i+1} closes $f(e)$.

(iv) If e is a partner edge then $\Delta_i(\mathcal{T}V^e) = \sum_{f\in V^-} \Delta_i(Y_{f(\beta)}) \cdot X_{f(\beta)} \cdot Qn^{-2}$.
If e is a partner open pair then $\Delta_i(\mathcal{T}V^e) = \sum_{f\in V^-} \Delta_i(X_{f(\beta)})\hat{U}_f$.

6.3.3 | Classification of pairs

As in the controllable ensemble, we will verify the trend and boundedness hypotheses by considering separately $\Delta_i(\mathcal{D}V^e) := \Delta_i(V^e) - \Delta_i(\mathcal{T}V^e)$ for each $e \in \Gamma \setminus \Gamma[A]$. We will organize the trend hypothesis by grouping together terms that use the same method of calculation, so here we introduce some terminology to classify these terms. We have met special cases of some of these terms earlier when we considered the global variables: again "simple" terms are those described by another variable in our ensemble, and the "product" terms in the global variables are analogous to the "internal" terms here. We use the following notation:

- For any $y \le |\pi|$ we let $\pi|_y$ denote the prefix of π of length y.
- If the final symbol $\pi(|\pi|) \in \{X^I, X^O, Y^I, Y^O\}$ we let π^o (the "opposite" variable) be obtained from π by interchanging superscripts I and O in $\pi(|\pi|)$.

For our classification we use the same terms internal, terminal and partner as above, but we must pay special attention to the terminal open pairs, which we divide into the following three subtypes (recall that if a pair is not a rung we call it a stringer):

(a) If e is a rung and πY^I and πY^O are both M-bounded we say that e is *simple*.
If e is a stringer and $\pi^o Y^I$ and $\pi^o Y^O$ are both M-bounded we say that e is *simple*.

(b) If $w(\pi) = 2M$ and e is the terminal rung we say that e is *outer*.
If $w(\pi) = 2M - 1$, $\pi(|\pi|) = X^I$ and e is the terminal stringer then we say that e is *outer*.

(c) If e is not simple or outer we say that e is a *fan end* pair.

To explain this classification, we note the following:

- Outer pairs are not simple, as adding Y^O to any π' with $w(\pi') = 2M$ gives a variable not in $\mathcal{S}_M$ (consider $\pi' = \pi$ if e is the terminal rung or $\pi' = \pi^o$ if e is the terminal stringer).
- Fan end pairs are aptly named, as if there is a fan end pair it follows from the definition of the M-bounded stacking ensemble $\mathcal{S}_M$ (see Definition 2.5) that π must end with an $(M-1)$-fan.

6.4 | Correction terms

Before starting on the main calculations for the trend hypothesis, here we will summarize various correction terms which are negligible by comparison with the terms appearing in the variation equations. Besides the higher-order corrections (34) to changes in the tracking variable mentioned above, we also have the following "injectivity" and "fidelity" corrections.

Lemma 6.5 (Injectivity). *Suppose $i < I$ and $V = X_{\phi,J,\Gamma}$ is a stacking variable or fan extension with $v \ge y$. Then for any vertex $x \notin A$ (the base) there are $\tilde{O}(t^{-1}e^2)v$ choices of $f \in V$ with $x \in Im(f)$.*

Proof. Fix $a \in V_\Gamma \setminus A$, let $A' = A \cup \{a\}$ and extend ϕ to ϕ' on A by $\phi'(a) = x$. It suffices to show that the stated bound holds for $X_{\phi',J,\Gamma}$. Fix $A' \subseteq B \subseteq V_\Gamma$ minimizing S^B_A. If V is a stacking variable, then considering vertices one by one in the stacking order we have $S^B_A \geq y$. If V is a fan then either $B = V_\Gamma$, when $S^B_A = v \geq y$, or $B = A'$ (as in the proof of Lemma 6.2), so again $S^B_A \geq y$. As $i < I$, by property (iv) of Definition 2.12 the number of choices for f is at most $L^{4|V_\Gamma|}v/S^B_A$. The lemma follows as $y = 2te^{-2}$. ∎

Lemma 6.6 (Fidelity). *Suppose $i < I$ and $V = S^\pi_{uv}$ is good.*

(i) There are $O(L^4 v)$ pairs (f, xy) where $f \in V$ such that if xy were the edge e_{i+1} selected at step $i+1$ then at least two open pairs in f would become closed.

(ii) Let V^+ be a stacking variable obtained from V by changing some edge e to an open pair. There are $\widetilde{O}(e^2 v^+)$ choices of $f \in V^+$ such that if $f(e)$ were the edge e_{i+1} selected at step $i+1$ then some open pair in f would become closed.

Proof. Let (uv, J, Γ) be the extension corresponding to V.

For (i), we first note that for each $f \in V$ there are only $O(1)$ choices of $xy \subseteq \text{Im}(f)$. Any other xy with the stated property must have one of its vertices in $\text{Im}(f)$, say y, and the open pairs in f closed by xy are of the form ya, yb with a, b in $\text{Im}(f)$ where xa, xb are edges. As $i < I$, by property (iii) of Definition 2.12 the number of choices for x given f is at most $Z_{ab} < L^4$. This proves (i).

For (ii), note first that for such a configuration to exist we must have $|\pi| \geq 2$, so V has scaling $v \geq y^2$. We consider the extension $(uv, J \setminus e, \Gamma)$ corresponding to V^+ and any variable V^* corresponding to an extension (uv, J^*, Γ) with $J^* = (J \setminus e) \cup e'$ for some $e' \in \binom{V_\Gamma}{2} \setminus \Gamma$. It suffices to show $V^* = \widetilde{O}(e^2 v^+)$.

Note that $v^+ = \hat{q}p^{-1}v = (2te^2)^{-1}v$, so $e^2 v^+ = (2t)^{-1}v > 1$, and $v^* = pv^+ = \hat{q}v$. Fix $uv \subseteq B \subseteq V_\Gamma$ minimizing S^B_{uv}, taking scalings with respect to (uv, J^*, Γ). If $e' \not\subseteq B$ or $B = uv$ then $S^B_{uv} \geq 1$, as the scaling is the same as in V^+, so by property (iv) of Definition 2.12 we have $V^* = \widetilde{O}(v^*) = \widetilde{O}(e^2 v^+)$. If $|B| \geq 4$ we have $S^B_{uv} \geq y^2$, so $V^* = \widetilde{O}(v^+/y^2) = \widetilde{O}(e^2 v^+)$.

The remaining case is that $|B| = 3$ and $B = uv \cup e'$. Write $B = \{u, v, \alpha_j\}$. We cannot have $j = 1$, as $e' \notin \Gamma$ would then imply $\pi(1) \in \{O, E\}$, so the assumption of the lemma could not hold: selecting $f(e')$ as an edge for such e' cannot close any other pair in f. Thus α_j is adjacent in Γ to at most one of u, v, so $S^B_{uv} \geq pn$, giving $V^* = \widetilde{O}(v^*/pn) = \widetilde{O}(e^2 v^+)$. ∎

6.5 | Creation

Now we will estimate the one-step expected changes $\mathbb{E}[\Delta_i(V^e) \mid \mathcal{F}_i]$ for each $e \in J \setminus J[A]$, according to the classification of pairs described above. As for the other ensembles, the error terms for creation are not as significant as those for destruction, and the calculations do not require self-correction or use the fact that V is in its critical window. We do use $i_V \leq i < I$. Note that we do not include in these calculations the fidelity corrections (see Lemma 6.6.ii).

6.5.1 | Terminal creation

Suppose that e is the terminal edge of π. Then $\pi(|\pi|)$ is E, Y^I, or Y^O, and if $\pi(|\pi|) = Y^O$ then $\pi(|\pi| - 1) \neq O$ (otherwise e would be partner). Let V^+ be the variable obtained by changing e to an open pair, that is, replacing Y by X in $U = \pi(|\pi|)$. Then $\mathbb{E}[\Delta_i(V^e) \mid \mathcal{F}_i] = 2Q^{-1}V^+$. For the tracking variable, we note that $\Delta_i(\mathcal{T}V^e) = \frac{\mathcal{T}V}{tn^{3/2}} = 2Q^{-1}\mathcal{T}V^+$ (whether V is standard or partner). As $v^+ = v \cdot \frac{\hat{q}n^{1/2}}{2t}$ and $Q = (1 + o(e))q$ for $i_V \leq i < I$ we have

$$\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] = \mathbb{E}[\Delta_i(V^e) - \Delta_i(\mathcal{T}V^e) \mid \mathcal{F}_i] = 2Q^{-1}\mathcal{D}V^+ = \pm(1 + o(e))t^{-1}\delta_{V^+}vn^{-3/2}.$$

6.5.2 | Partner creation

Suppose that $e = \alpha_x\alpha_y$ with $x < y = |\pi|$ is the partner edge of π. We must have $x = |\pi|-2, y = |\pi|-1$ and $\pi = \pi^- OY^O$. In this case, we recall that the tracking variable is $\mathcal{T}V = \sum_{f\in V^-} X_{f(\alpha_x)}Y_{f(\alpha_x)}Qn^{-2}$, where $V^- = S^{\pi^-}_{uv}$. We let V^+ be obtained from V by changing e to an open pair. Then $\mathbb{E}[\Delta_i(V^e) \mid \mathcal{F}_i] = 2Q^{-1}V^+$. We also recall that $\mathcal{T}V^+ = \sum_{f\in V^-} X^2_{f(\alpha_x)}Qn^{-2}$ and $\Delta_i(\mathcal{T}V^e) = \sum_{f\in V^-} \Delta_i(Y_{f(\alpha_x)}) \cdot X_{f(\alpha_x)} \cdot Qn^{-2}$, so $\mathbb{E}[\Delta_i(\mathcal{T}V^e) \mid \mathcal{F}_i] = 2Q^{-1}\mathcal{T}V^+$. Thus we obtain the same estimate as in terminal creation for $\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i]$.

Note that the definition of the tracking variables isolates variations in V from those in V^-, which is crucial in this calculation: we cannot afford the larger error term δ_{V^-}.

6.5.3 | Internal creation

Suppose that $e = \alpha_x\alpha_y$ with $x < y < |\pi|$ is an internal edge of π (which must be a stringer). Let V^+ be obtained from V by changing e to an open pair. Then $\mathbb{E}[\Delta_i(V^e) \mid \mathcal{F}_i] = 2Q^{-1}V^+$. For the tracking variable, we recall from (35) that $\mathcal{T}V = \sum_{f\in V^-} \mathcal{T}_f V$ and $\Delta_i(\mathcal{T}V^e) = \sum_{f\in V^{-+}} I^e_f \mathcal{T}_f V$, where V^{-+} is obtained from V^- by changing e to an open pair and I^e_f is the indicator of the event that $e_{i+1} = f(e)$. Thus $\mathbb{E}[\Delta_i(\mathcal{T}V^e) \mid \mathcal{F}_i] = \sum_{f\in V^{-+}} \mathcal{T}_f V = 2Q^{-1}\mathcal{T}V^+$, so we obtain the same estimate for $\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i]$ as in terminal and partner creation.

As for partner creation, it is crucial that $\mathcal{T}V$ isolates variations in V^- from this calculation.

6.6 | Destruction

Now we will estimate the one-step expected changes $\mathbb{E}[\Delta_i(V^e) \mid \mathcal{F}_i]$ for each $e \in \Gamma \setminus \Gamma[A]$, according to the classification of pairs described above, assuming that $V = S^\pi_{uv}$ is in its upper critical window, so that $\mathcal{D}V > (f_V + g_V)v$. As usual, the key point is that every open pair yields a self-correcting term of the form $(f_V + g_V)8tvn^{-3/2}$. We remark that the calculations for terminal open pairs will be the source of the most significant error terms in the variation equations.

6.6.1 | Simple destruction

Let $e = \alpha_x\alpha_y$ be a simple rung, that is, the last rung of π such that πY^I and πY^O both belong to $\mathcal{S}_M$. Write $V^I = S^{\pi Y^I}_{uv}$ and $V^O = S^{\pi Y^O}_{uv}$. We have

$$\mathbb{E}[\Delta_i(V^e) \mid \mathcal{F}_i] = 2Q^{-1}\sum_{f\in V}(Y_{f(\alpha_x\alpha_y)} + Y_{f(\alpha_y\alpha_x)} \pm O(1)) = 2Q^{-1}(V^I + V^O \pm O(v)).$$

Note that $\mathcal{T}V^I = \mathcal{T}V^O = 2tQn^{-3/2}V$ and $v^I = v^O = 2t\hat{q}n^{1/2}v$. Since $\Delta_i(\mathcal{T}V^e) = \frac{\Delta_i(Q)}{Q}\mathcal{T}V$, recalling (17) we have

$$\begin{aligned}
\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] &= \mathbb{E}\left[\Delta_i(V^e) - \Delta_i(\mathcal{T}V^e) \mid \mathcal{F}_i\right] \\
&= -2Q^{-1}(V^I + V^O \pm O(V)) + (2 + 4SQ^{-1})Q^{-1}\mathcal{T}V \\
&= -2Q^{-1}(\mathcal{T}V^I + \mathcal{T}V^O \pm v^I\delta_{V^I} \pm v^O\delta_{V^O}) + (8tn^{-3/2} \pm 4\delta_S sq^{-2})\mathcal{T}V \pm O(v/q) \\
&= -8tn^{-3/2}\mathcal{D}V \pm (1 + o(1))8t(\delta_{V^I}/2 + \delta_{V^O}/2 + \delta_S)vn^{-3/2} \pm O(v/q) \\
&\le -(1 + o(1))(f_V + g_V - \delta_{V^I}/2 - \delta_{V^O}/2 - \delta_S - O(t^{-1}e^2))8tvn^{-3/2}.
\end{aligned}$$

The same calculation applies if e is a simple stringer (using π^o in place of π). Note that the estimates for V^I and V^O are valid even before their activation steps by Lemma 3.9.iv. The appearance of their approximation errors δ_{V^l} and δ_{V^o} in this calculation indicates why we need these errors to decrease as we increase the length of the stacking extensions (see Definition 2.8).

6.6.2 | Internal destruction

Suppose that $e = \alpha_x\alpha_y$ with $x < y < |\pi|$ is an internal open pair (note that we do not include partners here). We let $W = S^{\pi'}_{uv}$, where $\pi' = \pi|_y$ if e is a rung or $\pi' = \pi|^o_y$ if e is a stringer.

For each $f \in W$ let $F_{f,\pi}$ count *forward extensions* from f to copies of V, that is, $F_{f,\pi} = X_{f,J,\Gamma}$ with $f : A \to [n]$, where $A = \{\alpha_u, \alpha_v, \dots, \alpha_y\}$.

We note that $F_{f,\pi}$ is closely approximated, up to the injectivity correction from Lemma 6.5, by another variable $V_1^f = S^{\pi_1}_{f(e')}$ in the stacking variable, where e' is the active rung at step y and $\pi|_y \circ \pi_1 = \pi$: we have $F_{f,\pi} = V_1^f + \widetilde{O}(t^{-1}e^2)v_1$, so

$$V = \sum_{f\in W} F_{f,\pi} = \sum_{f\in W}\left(V_1^f + \widetilde{O}(t^{-1}e^2)v_1\right).$$

For the tracking variable, we recall from (35) that $\mathcal{T}V = \sum_{f'\in V^-}\mathcal{T}_{f'}V$. Similarly to above, we define the forward extension F_{f,π^-} from $f \in W$ to copies of V^- and approximate it by $F_{f,\pi^-} = V_2^f + \widetilde{O}(t^{-1}e^2)v_2$, where $V_2^f = S^{\pi_2}_{f(e')}$ and $\pi|_y \circ \pi_2 = \pi^-$. Then

$$\mathcal{T}V = \sum_{f\in W}\sum_{f'\in F_{f,\pi^-}}\mathcal{T}_{f'}V = \sum_{f\in W}\left(\mathcal{T}V_1^f + \widetilde{O}(t^{-1}e^2)v_1\right), \quad \text{so}$$
$$\mathcal{D}V = V - \mathcal{T}V = \sum_{f\in W}\left(\mathcal{D}V_1^f + \widetilde{O}(t^{-1}e^2)v_1\right). \tag{36}$$

Similarly, writing I^e_f for the indicator of the event that e_{i+1} closes $f(e)$, noting that $\Delta_i(V^e) = \sum_{f'\in V} I^e_{f'} = \sum_{f\in W} I^e_f F_{f,\pi}$ and $\Delta_i(\mathcal{T}V^e) = \sum_{f'\in V'}\mathcal{T}_{f'}VI^e_{f'} = \sum_{f\in W} I^e_f \sum_{f'\in F_{f,\pi^-}}\mathcal{T}_{f'}V$, we have

$$\Delta_i(\mathcal{D}V^e) = \Delta_i(V^e) - \Delta_i(\mathcal{T}V^e) = -\sum_{f\in W}\left(\mathcal{D}V_1^f + \widetilde{O}(t^{-1}e^2)v_1\right)I^e_f. \tag{37}$$

We also note from $i_V \le i < I$ and (14) that

$$W^* := \sum_{f\in W}(Y_{f(xy)} + Y_{f(yx)}) = (1 \pm \delta_Y)2Wy. \tag{38}$$

Taking expectations of (37) and applying Lemma 2.14 (the product lemma) we have

$$\begin{aligned}
\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] &= -2Q^{-1}\sum_{f\in W}(Y_{f(xy)} + Y_{f(yx)} \pm O(1))(\mathcal{D}V_1^f + \widetilde{O}(t^{-1}e^2)v_1)\\
&= -\frac{2W^*\mathcal{D}V}{QW} \pm O(Q^{-1}W\cdot y\delta_Y\cdot v_1\delta_{V_1}) \pm \widetilde{O}(t^{-1}e^2)vtn^{-3/2}\\
&\le -(1+o(1))(f_V + g_V - O(\delta_{V_1}\delta_Y) - \widetilde{O}(t^{-1}e^2))8tvn^{-3/2}\\
&\le -(1+o(1))(f_V + g_V - \widetilde{O}(t^{-1}e^2))8tvn^{-3/2}.
\end{aligned}$$

We used the scaling identities $v = wv_1$ and $v^- = wv_2$. In the application of the product lemma on the third line we used (36) and (38).

The last line exhibits the same crucial feature that we saw earlier in product destruction for global variables: the $O(\delta_{V_1}\delta_Y)$ term is negligible, as for small t the $t^{-e(V)}$ factor in g_V dominates the $t^{-e(V_1)}$ factor in δ_{V_1}, and the δ_Y factor compensates for the larger polylogarithmic factor in δ_{V_1}.

6.6.3 | Partner destruction

Here we consider a partner open pair $e = \alpha_x\alpha_y$ with $x < y$. Recall that this means $\pi(|\pi| - 1) = O$, $\pi(|\pi|) \notin \{O, E\}$, $x = |\pi| - 2$, and $y \in \{|\pi| - 1, |\pi|\}$. Let $\pi = \pi^- OU$ and $V^- = S^{\pi^-}_{uv}$. Recall from Section 6.3.2 that $\mathcal{T}V = \sum_{f\in V^-} X_{f(\alpha_x)}\hat{U}_f$, where

$$\hat{U}_f = \begin{cases} X_{f(\beta)} \cdot Qn^{-2} & \text{if } U \in \{X^I, X^O\} \\ X_{f(\beta)} \cdot 2tn^{-1/2} & \text{if } U = Y^I \\ Y_{f(\beta)} \cdot Qn^{-2} & \text{if } U = Y^O. \end{cases}$$

Note that if both partner pairs are open then the definitions of V and $\mathcal{T}V$ are symmetric under swapping the labels of $\alpha_{|\pi|-1}$ and $\alpha_{|\pi|}$, so we can assume $y = |\pi| - 1$. This would not have been true with our usual practice of using the tracking variable $\mathcal{T}U$ instead of $\hat{U}_f$; the point is that we want the self-correction in this section to apply to both partner pairs. (This property of $\mathcal{T}V$ for partners is also essential for our treatment of fan extensions in Section 6.6.5.) On the other hand, we can think of $\hat{U}_f$ as a proxy for $\mathcal{T}U$ as it is a reasonable approximation to U: as $i < I$ we have

$$\hat{\mathcal{D}}U_{f(\alpha_x)z} := U_{f(\alpha_x)z} - \hat{U}_f = O((\delta_U + \delta_{\hat{U}})u),$$

where $\delta_{\hat{U}} = \delta_{Y_1}$ if $U = Y^O$, otherwise $\delta_{\hat{U}} = \delta_{X_1}$. Writing u for the scaling of U, we have

$$V = \sum_{f\in V^-} \sum_{z\in X_{f(\alpha_x)}\setminus \mathrm{Im}(f)} (U_{f(\alpha_x)z} + O(1)), \quad \text{so}$$
$$\mathcal{D}V = V - \mathcal{T}V = \sum_{f\in V^-} \Big(O(u + x_1) + \sum_{z\in X_{f(\alpha_x)}} \hat{\mathcal{D}}U_{f(\alpha_x)z} \Big). \tag{39}$$

Recalling $\Delta_i(\mathcal{T}V^e) = \sum_{f\in V^-} \Delta_i(X_{f(\alpha_x)})\hat{U}_f$ and writing I_{fz} for the indicator of the event that e_{i+1} closes $f(\alpha_x)z$, we have

$$\Delta_i(\mathcal{D}V^e) = -\sum_{f\in V^-} \Big(\sum_{z\in X_{f(\alpha_x)}} (\hat{\mathcal{D}}U_{f(\alpha_x)z} \pm O(1))I_{fz} - \sum_{z\in \mathrm{Im(f)}} O(u)I_{fz} \Big). \tag{40}$$

Also, writing $W = \sum_{f\in V^-} X_{f(\alpha_x)}$, from $i < I$ and (14) we have

$$W^* := \sum_{f\in W} (Y_{f(xy)} + Y_{f(yx)}) = (1 \pm \delta_Y)2Wy. \tag{41}$$

Taking expectations of (40) and applying Lemma 2.14 (the product lemma) we have

$$\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] = \mathbb{E}[\Delta_i(V^e) - \Delta_i(\mathcal{T}V^e) \mid \mathcal{F}_i]$$
$$= -2Q^{-1}\Big[\sum_{f\in V^-} \sum_{z\in X_{f(\alpha_x)}} (Y_{f(\alpha_x)z} + Y_{zf(\alpha_x)} \pm O(1))\hat{\mathcal{D}}U_{f(\alpha_x)z} \Big] \pm O(x_1 + u)v^- y/q$$

$$= -\frac{2W^*\mathcal{D}V}{QW} \pm O(wq^{-1} \cdot y\delta_Y \cdot u(\delta_U + \delta_{\hat{U}})) \pm \widetilde{O}(e^2)vn^{-3/2}$$
$$= -(1+o(1))\mathcal{D}V \cdot 8tn^{-3/2} \pm (O(\delta_Y\delta_U) + O(\delta_Y\delta_{\hat{U}}) + \widetilde{O}(t^{-1}e^2))tvn^{-3/2}$$
$$\le -(1+o(1))(f_V + g_V - \widetilde{O}(t^{-1}e^2))8tvn^{-3/2}.$$

In the application of the product lemma on the third line we used (39) and (41). The last line is valid because the product errors $\delta_Y\delta_U$ and $\delta_Y\delta_{\hat{U}}$ are $o(\delta_V)$; this holds as δ_Y has sublogarithmic decay and the power of t^{-1} in g_V is at least those in each of g_U and $g_{\hat{U}}$.

6.6.4 | Outer destruction

Let $e = \alpha_x\alpha_y$ be an outer rung, that is, e is terminal and $w(\pi) = 2M$. We cannot apply the same analysis as for simple destructions, as $\pi Y^O \notin \mathcal{S}_M$, so instead we use backward extensions, which are controllable by Lemma 6.2.

We let Q' be the set of $ab \in Q$ such that $\{a,b\} \cap \{u,v\} = \emptyset$, and for each $ab \in Q'$ let B_{uvab} count backward extensions that map the last rung of S^{π}_{uv} to the open pair ab; thus $V = \sum_{ab\in Q'} B_{uvab}$.

Let b and δ_B be the scaling and error function for the backward extension. Then $b = v/q$ and $\delta_B = O(1 + t^{-e(V)})e^{\delta}$. Note also that $Q - Q' = O(x_1)$ and $S = \sum_{ab\in Q} Y_{ab} = O(x_1 y) + \sum_{ab\in Q'} Y_{ab}$. Recalling (17) and $\Delta_i(\mathcal{T}V^e) = \frac{\Delta_i Q}{Q}\mathcal{T}V$, by the product lemma (Lemma 2.14) we have

$$\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] = \mathbb{E}[\Delta_i(V^e) - \Delta_i(\mathcal{T}V^e) \mid \mathcal{F}_i]$$
$$= -2Q^{-1}\sum_{ab\in Q'} X_{uvab}(Y_{ab} + Y_{ba} \pm O(1)) + \frac{4S+2Q}{Q^2}\mathcal{T}V$$
$$= -\frac{2V}{QQ'}\Big(2S - O(yx_1)\Big) \pm \frac{4Q'}{Q} \cdot y\delta_Y \cdot b\delta_B + \frac{4S+2Q}{Q^2}\mathcal{T}V \pm O(v/q)$$
$$= -\frac{4S}{Q^2}\mathcal{D}V \pm O(\delta_Y\delta_B + t^{-1}e^2)tvn^{-3/2}$$
$$\le -(1+o(1))(f_V + g_V - O(\mathbf{1}_{e(V)=0}\delta_Y e^{\delta}) - O(t^{-1}e^2))8tvn^{-3/2}.$$

The last line used $\delta_Y\delta_B = o(g_V)$ when $e(V) \ge 1$, which holds as δ_B has sublogarithmic decay (using $i < I$) and the power of t^{-1} in V is at least that in Y. Thus this term is negligible unless $e(V) = 0$, in which case we can substitute $\delta_B = O(e^{\delta})$.

Note that the same estimate applies if e is an outer stringer (using π^o in place of π).

6.6.5 | Fan end destruction

For destruction, it remains to consider the case when $e = \alpha_x\alpha_y$ is a fan end, that is, π ends with an $(M-1)$-fan and e is the terminal rung. We cannot apply the analysis from simple destructions, as $\pi Y^I \notin \mathcal{S}_M$, so instead we use controllability of fan extensions (see Lemma 6.2).

Let $V^* = S^{\pi^*}_{uv}$, where $\pi^* = \pi|_x O$, that is, V^* is obtained from V by deleting all of the fan except its first pair $\alpha_{x-1}\alpha_x$ and last pair $e = \alpha_x\alpha_y$. Then $V = \sum_{f\in V^*} F_{f,\pi}$, where $F_{f,\pi}$ denotes the forward extension, which is closely approximated by the $(M-1)$-fan extension V_1 from $f(\alpha_{x-1}\alpha_x\alpha_y)$; by Lemma 6.5 we have $F_{f,\pi} = V_1 + O(t^{-1}e^2)v_1$. We recall that V_1 is controllable by Lemma 6.2.

In the calculation below for $\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i]$ we require the following estimate for the expected closures of the terminal open pair $\alpha_x\alpha_y$ in copies of V^*, which are described by

$$V^*_{\text{close}} := 2Q^{-1}\sum_{f\in V^*}(Y_{f(\alpha_x\alpha_y)} + Y_{f(\alpha_x\alpha_y)}).$$

Lemma 6.7. *Let $\pi^\bullet = \pi|_x E$, $V^x = S_{uv}^{\pi|_x}$, $V^\bullet = S_{uv}^{\pi^\bullet}$, $V^{*I} = S_{uv}^{\pi^* Y^I}$ and $V^{*O} = S_{uv}^{\pi^* Y^O}$. Then*

$$V^*_{close} = 8tn^{-3/2}\Big[V^* \pm (1+o(1))\big(\delta_{V^*} + \delta_{V^\bullet} + \delta_{V^{*I}} + \delta_{V^{*O}} + O(\delta_{X_1} + \delta_{Y_1})\delta_{X_1}\big)v^*/2\Big].$$

Proof. First we emphasize that all variables defined in the statement of the lemma are in the stacking ensemble, and this fact makes crucial use of Definitions 2.4 and 2.5. The point is that as non-terminal OX^I and OY^I are forbidden, the fan must start with $\pi(x+1) \in \{X^O, Y^O\}$, and also $w(\pi) \le 2M-1$ as we do not allow a strict subsequence of weight $2M$, so $w(\pi|_x) \le w(\pi) - 1 \le 2M-2$. Now

$$\begin{aligned}\sum_{f\in V^*}(Y_{f(\alpha_x\alpha_y)} + Y_{f(\alpha_x\alpha_y)} \pm O(1)) &= V^{*I} + V^{*O} \pm O(v^*)\\ &= \mathcal{T}V^{*I} + \mathcal{T}V^{*O} \pm (\delta_{V^{I*}}y + \delta_{V^{*I}}y + O(1))v^*,\end{aligned}$$

where, as V^{*I} and V^{*O} are both partner variables, by Lemma 2.14 we have

$$\begin{aligned}\mathcal{T}V^{*I} &= \sum_{f\in V^x} X^2_{f(\alpha_x)} \cdot 2tn^{-1/2} = 2tn^{-1/2}\cdot V^*V^*/V^x \pm O\big(tn^{-1/2}v^x(x_1\delta_{X_1})^2\big) \text{ and}\\ \mathcal{T}V^{*O} &= \sum_{f\in V^x} X_{f(\alpha_x)}Y_{f(\alpha_x)}\cdot Qn^{-2} = Qn^{-2}\cdot V^*V^\bullet/V^x \pm O\big(\hat{q}v^x(x_1\delta_{X_1})(y_1\delta_{Y_1})\big).\end{aligned}$$

The lemma now follows from $V^* = \mathcal{T}V^* \pm \delta_{V^*}v^*$ and $V^\bullet = \mathcal{T}V^\bullet \pm \delta_{V^\bullet}v^\bullet$, where $\mathcal{T}V^* = Qn^{-1}V^x$ and $\mathcal{T}V^\bullet = 2tn^{1/2}V^x$, so V^x cancels (this is crucial to avoid a larger δ_{V^x} error term). ■

Now recalling $F_{f,\pi} = V_1 + O(t^{-1}e^2)v_1$, using $\Delta_i(\mathcal{T}V^e) = \frac{\Delta_i Q}{Q}\mathcal{T}V$ and (17), by Lemma 2.14

$$\begin{aligned}&\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] = -2Q^{-1}\sum_{f\in V^*}(Y_{f(\alpha_x\alpha_y)} + Y_{f(\alpha_x\alpha_y)} \pm O(1))F_{f,\pi} + \frac{4S+2Q}{Q^2}\mathcal{T}V\\ &= -V^*_{\text{close}}V/V^* \pm O(t^{-1}e^2)v^*v_1y/q \pm O(v^*q^{-1}\cdot y\delta_Y\cdot v_1\delta_{V_1}) + \frac{4S}{Q^2}\mathcal{T}V \pm O(v/q)\\ &= -8tn^{-3/2}V + (1+(1+o(1))\delta_S)8tn^{-3/2}\mathcal{T}V\\ &\qquad \pm (1+o(1))\tfrac{1}{2}\big(\delta_{V^*} + \delta_{V^\bullet} + \delta_{V^{*I}} + \delta_{V^{*O}} + O(\delta_Y\delta_{V_1} + t^{-1}e^2 + (\delta_{X_1} + \delta_{Y_1})\delta_{X_1})\big)8tvn^{-3/2}\\ &\le -(1+o(1))\Big(f_V + g_V - \tfrac{1}{2}(\delta_{V^*} + \delta_{V^\bullet} + \delta_{V^{*I}} + \delta_{V^{*O}}) - \delta_S - O(t^{-1}e^2) - O(\delta_Y e^\delta)\Big)8tvn^{-3/2}.\end{aligned}$$

In the third line we applied Lemma 6.7. In the last line, similarly to the case of outer destruction, we note that $(\delta_{X_1} + \delta_{Y_1})\delta_{X_1} = o(\delta_V)$, as $\delta_{X_1} = \widetilde{O}(\delta_V)$ and $\delta_{X_1} + \delta_{Y_1}$ has sublogarithmic decay. Similarly, if $e(V) \ge 1$ then $\delta_Y = \widetilde{O}(\delta_V)$ and δ_{V_1} has sublogarithmic decay, so $\delta_Y\delta_{V_1} = o(\delta_V)$. Thus the only product error is $O(\delta_Y\delta_{V_1}) = O(\delta_Y e^\delta)$ when $e(V) = 0$.

6.7 | Trend hypothesis and variation equations

Now we combine all the estimates in this section to verify the trend hypothesis, that is, that if V is in its upper critical window then $\mathcal{Z}V = \mathcal{D}V - \delta_V v$ forms a supermartingale, given the choice of constants c_V made in Definition 2.8.

Lemma 6.8. *If $i_V \le i < I$ and $\mathcal{D}V > (f_V + g_V)v$ then $\mathbb{E}[\Delta_i\mathcal{Z}V \mid \mathcal{F}_i] \le 0$.*

Proof. Throughout the proof we will measure expected changes using the "yard stick" $8tvn^{-3/2}$, which is an approximation for the expected change in V due to destruction by some fixed open pair. Recall that we decompose the one-step change in $V = X_{\phi,J,\Gamma}$ by its pairs e as

$$\Delta_i(V) = \sum_{e\in\Gamma\setminus\Gamma[A]} \Delta_i(V^e) \pm F_i(V),$$

where $F_i(V)$ is a fidelity correction, which by Lemma 6.6 satisfies

$$\mathbb{E}[F_i(V) \mid \mathcal{F}_i] = O(L^4 v/q) + \widetilde{O}(e^2 v^+/q) = (t^{-1} + t^{-2} 1_{e(V)>0})\widetilde{O}(e^2) \cdot tvn^{-3/2}.$$

Recall also that we decompose the one-step change in the tracking variable as

$$\Delta_i(\mathcal{T}V) = \sum_{e\in\Gamma\setminus\Gamma[A]} \Delta_i(\mathcal{T}V^e) + H_i(V),$$

where the higher-order correction term is

$$H_i(V) = O(t^2 + t^{-2})n^{-3}v = O(n^{-5/4}) \cdot tvn^{-3/2} \text{ for } n^{-1/4} \leq t = O(L).$$

Besides the fidelity and higher-order terms, the remaining contributions to $\mathbb{E}[\Delta_i(\mathcal{D}V) \mid \mathcal{F}_i] = \mathbb{E}[\Delta_i(V) - \Delta_i(\mathcal{T}V) \mid \mathcal{F}_i]$ are obtained by summing $\mathbb{E}[\Delta_i(\mathcal{D}V^e) \mid \mathcal{F}_i] = \mathbb{E}[\Delta_i(V^e) - \Delta_i(\mathcal{T}V^e) \mid \mathcal{F}_i]$ over all $e \in \Gamma \setminus \Gamma[A]$.

There are $e(V)$ edges each giving a creation term of

$$\pm(1 + o(e))t^{-1}\delta_{V^+} vn^{-3/2} = (1 + o(e))\tfrac{\delta_{V^+}}{8t^2} \cdot 8tvn^{-3/2}.$$

There are $o(V)$ open pairs each giving a destruction term in which the main term is a self-correction term of

$$-(1 + o(1)(f_V + g_V)8tvn^{-3/2}.$$

For open pairs that are partner or internal the only other error term is $\widetilde{O}(t^{-1}e^2) \cdot tvn^{-3/2}$, which we can absorb into the fidelity term. The terminal open pairs (of which there are one or two) contribute an additional error term, depending on the form of π, which we denote by $\delta_{\text{add}} \cdot 8tvn^{-3/2}$.

We claim the following bound:

$$|\delta_{\text{add}}| \leq 0.49\delta_V + O(\delta_Y e^{\delta}).$$

To see this, we first suppose $\pi \neq O$ and consider each of the three types of terminal open pair.

- The only contribution to δ_{add} from an outer open pair is $O(\delta_Y e^{\delta})$.
- The contribution to δ_{add} from a simple open pair is $(1 + o(1))(\delta_{V^I}/2 + \delta_{V^O}/2 + \delta_S)$. We can absorb δ_S into the $O(\delta_Y e^{\delta})$ term. From Definition 2.8 we have

$$c_{V^O} = c_{V^I} = c_V/9, \tag{42}$$

  so $\delta_{V^I}/2 + \delta_{V^O}/2 \leq \delta_V/9$, and we can bound this contribution to δ_{add} by $\delta_V/8 + O(\delta_Y e^{\delta})$.

- The contribution to δ_{add} from a fan end open pair is

$$(1 + o(1))(\tfrac{1}{2}(\delta_{V^*} + \delta_{V^\bullet} + \delta_{V^{*I}} + \delta_{V^{*O}}) + \delta_S) + O(\delta_Y e^\delta).$$

  Again $\delta_S = O(\delta_Y e^\delta)$. The sequences defining V^* and $V^\bullet$ each have $M-1$ fewer symbols than π, but this is compensated for by an additional "O" or "E." Thus Definition 2.8 gives

$$c_{V^*} = c_{V^\bullet} = c_V/9, \text{ and} \tag{43}$$

$$c_{V^{*I}} = c_{V^{*O}} = c_V/81. \tag{44}$$

  Thus $\frac{1}{2}(\delta_{V^*} + \delta_{V^\bullet} + \delta_{V^{*I}} + \delta_{V^{*O}}) \le \delta_V/9 + \delta_V/81$, so we can bound this contribution to δ_{add} by $\delta_V/8 + O(\delta_Y e^\delta)$.

As V can have at most two terminal open pairs, this proves the claim when $\pi \neq O$. If $\pi = O$ then the only contribution is from the simple open pair; recalling the adjustment in Definition 2.8 we have $c_{V^O} = c_{V^I} = 2.2c_V/9$, so the claim also holds in this case.

Combining all the estimates so far gives

$$\begin{aligned}\frac{\mathbb{E}[\Delta_i(\mathcal{D}V) \mid \mathcal{F}_i]}{8tvn^{3/2}} \le & -(1+o(1))o(V)(f_V + g_V) + (1+o(e))e(V)\frac{\delta_{V^+}}{8t^2} \\ & + 0.49\delta_V + O(\delta_Y e^\delta) + (t^{-1} + t^{-2}1_{e(V)>0})\widetilde{O}(e^2).\end{aligned}$$

By Lemma 2.15 we have

$$\frac{\Delta_i(v\delta_V)}{8tvn^{3/2}} \ge \left(\tfrac{e(V)}{8t^2} - o(V) + O(t^{-1}n^{-1})\right)\delta_V + \left(4t\delta_V + (\vartheta'/\vartheta - e(V)t^{-1})2g_V\right)/8t.$$

By Definition 2.8, as V is not a vertex degree we have $c_V = c_{V^+}$, so as in the proof of Lemma 5.6 we have $\delta_{V^+} - \delta_V = 2(g_{V^+} - g_V)$ and $g_{V^+} \le 2tg_V$ (with no V^+ term if $e(V) = 0$). Thus

$$\begin{aligned}\frac{\mathbb{E}[\Delta_i(\mathcal{Z}V) \mid \mathcal{F}_i]}{8tvn^{-3/2}} &= \frac{\mathbb{E}[\Delta_i(\mathcal{D}V) \mid \mathcal{F}_i]}{8tvn^{3/2}} - \frac{\Delta_i(v\delta_V)}{8tvn^{3/2}} \\ &\le \tfrac{e(V)}{8t^2}\cdot 2(g_{V^+} - g_V) + o(V)g_V + 0.49(f_V + 2g_V) - f_V/2 - (\tfrac{\vartheta'/\vartheta}{8t} - \tfrac{e(V)}{8t^2} + \tfrac{1}{2})\cdot 2g_V \\ &\quad + O(\delta_Y e^\delta) + (t^{-1} + t^{-2}1_{e(V)>0})\widetilde{O}(e^2) + o(\delta_{V^+}et^{-2}) + o(\delta_V) \\ &\le g_V\left[o(V) + \tfrac{e(V)}{2t} - \tfrac{\vartheta'/\vartheta}{4t} - \tfrac{1}{50}\right] - \tfrac{f_V}{100} + O(\delta_Y e^\delta) + (t^{-1} + t^{-2}1_{e(V)>0})\widetilde{O}(e^2) + o(\delta_{V^+}et^{-2}) + o(\delta_V).\end{aligned}$$

To conclude the proof, it remains to check that this final expression is negative. This holds as $-g_V\vartheta'/(4t\vartheta)$ dominates when $g_V/t > f_V$ and $-f_V/100$ dominates otherwise. Here we recall that $\vartheta'/\vartheta = K > M^6$ for $t < 1$, and also use the later activation step (see Definition 2.9) for the case $e(V) = 1$ to see that the $t^{-2}1_{e(V)>0}\widetilde{O}(e^2)$ term is negligible. ■

Having verified the trend and boundedness hypotheses—these are Lemmas 6.8 and 6.4, respectively—we can apply Lemma 2.2 and show whp for all stacking variables V we have $V = (1 \pm \delta_V)v$ for $i_V \le i < I$. (Recall that Lemma 3.9 establishes this condition at step i_V.) This proves Theorem 6.1.

7 | INDEPENDENCE NUMBER AND UPPER BOUND

In this section we prove Theorem 1.2 on the independence number and establish the upper bound that completes the proof of Theorem 1.1 on the size of the final graph in the process. We will use union bound arguments that take advantage of our tight control of the evolution of key parameters until the process is very near its end.

We start by giving an intuitive overview of these arguments as applied to the independence number. Suppose we wish to estimate the probability that some set K of $\Theta(\sqrt{n \log n})$ vertices is independent. At any step i, with corresponding time $t = in^{-3/2}$, we would expect that K contains $\approx \hat{q}(t)|K|^2$ open ordered pairs. The total number of open pairs at step i is $Q(i) \approx q(t) = \hat{q}(t)n^2$, so the probability that K remains independent throughout the period in which we track the process should be roughly $(1 - |K|^2/n^2)^{i_{\max}}$. If this were true, we could estimate $Pr(\alpha(G) > k)$ by

$$\binom{n}{k}\left(1 - k^2/n^2\right)^{i_{\max}} < \exp(k \log \tfrac{en}{k} - i_{\max}k^2/n^2),$$

which is $o(1)$ for $k > (1 + o(1))\sqrt{2n \log n}$, as required to prove Theorem 1.2.

However, it is *not* true that every such K has $\approx \hat{q}|K|^2$ open ordered pairs; indeed, if K has a large intersection with the neighborhood of some vertex then K contains significantly fewer open pairs. Thus we require a much more delicate union bound calculation that takes into account the way in which vertex neighborhoods intersect K.

We stress that throughout this section we assume $I > i_{\max}$. Under this assumption, if $i \leq i_{\max}$ the good event $\mathcal{G}_i$ holds and every good V in the three ensembles satisfies $|V - \mathcal{T} V| \leq \delta_V v$. This assumption is valid as the events in the union we define are all intersected with the event $I > i_{\max}$. Formally speaking, in Section 7.2 we bound the probability of the event that $I > i_{\max}$ and the independence number of $G(i_{\max})$ is large, and in Section 7.3 we bound the probability that $I > i_{\max}$ and the maximum degree has the potential to become large in the steps that follow $i_{\max}$.

We also stress that throughout the section "neighbor" means "neighbor in $G(i_{\max})$"' and "$N(x)$" means "$N_{G(i_{\max})}(x)$."

To lighten notation in our calculations, we introduce the following notation for the number of steps in which we track the process and the deterministic prediction for the vertex degrees:

$$m = i_{\max} = \frac{1}{2}\sqrt{1/2 - \varepsilon}\, n^{3/2}(\log n)^{1/2} \quad \text{and} \quad d = 2t_{\max}\sqrt{n} = 2m/n = \sqrt{(1/2 - \varepsilon)n \log n}. \tag{45}$$

In the course of the proof, we will control various polylogarithmic factors using absolute constants

$$0 < \alpha < \gamma < \beta.$$

To clarify the role of these constants we will not substitute actual values, but for concreteness we note that we could let $\alpha = 25, \gamma = 50, \beta = 600$. When these polylog factors are unimportant we will use "tilde" notation as before: recall that $f(n) = \widetilde{O}(g(n))$ and $g(n) = \widetilde{\Omega}(f(n))$ mean that $f(n) \leq (\log n)^A g(n)$ for some absolute constant A.

Our proofs require some preliminary facts established in Section 7.1 (these are mostly density estimates for edges and open pairs). We prove Theorem 1.2 in Section 7.2, and then apply a similar (and easier) argument in Section 7.3 to prove Theorem 1.1.

7.1 | Preliminaries

This subsection contains some density estimates for edges and open pairs, and also some more intricate configurations that will play a crucial role in the argument in Section 7.2. These estimates will be obtained from the critical interval method as described in Section 2.1. We start with an observation that will be used many times in this section to estimate the one-step variances in some extension variable $V = X_{\phi,J,\Gamma}$ due to destruction. This will be applied as in Section 5.2 to bound the one-step conditional variance $\mathrm{Var}_V(i) = \mathrm{Var}(\mathcal{Z}V(i) \mid \mathcal{F}_{i-1})$ via a sum over pairs e in the configuration of the change in $\mathcal{Z}V$ due to the change of status of $f(e)$. Thus if $e = uv$ is an open pair in this configuration we want to estimate the one-step variance Var_e due to closing $f(e)$.

Lemma 7.1. *Consider any extension variable $V = X_{\phi,J,\Gamma}$ and open pair $e \in \Gamma \setminus J$ of Γ. Suppose at step i that the number $N_e^V(i)$ of injections f counted by V destroyed by closing $f(e)$ is bounded as $N_e^V(i) \leq N$, for some constant N. Then*

$$\mathrm{Var}_e := \mathrm{Var}(N_e^V(i) \mid \mathcal{F}_{i-1}) \leq (1+o(1))8tn^{-3/2}NV.$$

Proof. Consider the bipartite graph H with parts (A, B), where A is the set of injections counted by V, $B = Q$ is the set of ordered open pairs, and $f \in A$ is adjacent to $b \in B$ if selecting b as an edge closes $f(e)$. By assumption $d_H(b) \leq N$ for all $b \in B$. We also have $e(H) = 2\sum_{f\in V}(Y_{f(uv)} + Y_{f(vu)}) = (1+o(1))4yV$. Then $\mathrm{Var}_e \leq Q^{-1}\sum_{b\in B} d_H(b)^2 \leq (1+o(1))q^{-1}e(H)N = (1+o(1))8tn^{-3/2}NV$. ■

With this observation in hand, we turn next to some lemmas on counting open pairs.

Definition 7.2. For any set S let $Q_S(i)$ be the number of ordered open pairs in S at step i. For any sets A, B let $Q_{AB}(i)$ be the number of open pairs ab with $a \in A$, $b \in B$ at step i.

Lemma 7.3. *Whp for any set S of size s, step $i \leq i_{\max}$ and $\psi \geq n^{-\varepsilon/5}$,*

(i) if $s \geq n^{1/4}$ and any vertex x has $|N(x) \cap S| \leq L^{-10}\psi^2\hat{q}s$ then $Q_S = (1 \pm \psi)\hat{q}s^2$,
(ii) if $s \geq L^{11}\psi^{-2}\sqrt{n}$ then $Q_S = (1 \pm \psi)\hat{q}s^2$,
(iii) if $s < 2L^{12}\sqrt{n}$ then $Q_S < L^{13}s\hat{q}\sqrt{n}$.

Proof. First consider statements (i) and (ii). We use critical window analysis for $t \geq n^{-0.4}$ to prove the bound $Q_S = (1 \pm \delta_O)\hat{q}s^2$, where $\delta_O = (1 + t/L)\psi/2$. This suffices as $\delta_O \leq \psi$. We use the window $[(1 + \delta_O - g_O)\hat{q}s^2, (1 + \delta_O)\hat{q}s^2]$, where $g_O = \psi/(40L^2)$.

First we use coupling to the Erdős-Rényi process to show that whp Q_S does not enter the critical window at $t = n^{-0.4}$. This follows from the trivial upper bound $Q_S \leq s^2$, and the lower bound $Q_S \geq s^2 - 5n^{0.2}s$, obtained by subtracting the number of paths of length 2 starting in S in the random graph.

Next we establish the trend hypothesis that $\mathcal{Z}Q_S = Q_S - \hat{q}s^2 - \delta_O\hat{q}s^2$ is a supermartingale while Q_S is in its critical window. (Note that our tracking variable in this case is the deterministic function $\hat{q}s^2$.) The expected change in Q_S is

$$\mathbb{E}[\Delta_i Q_S \mid \mathcal{F}_i] = -2Q^{-1}\sum_{ab\in Q_S}(Y_{ab} + Y_{ba} + 1) = -8tn^{-3/2}(1 \pm O(\delta_Y))Q_S.$$

We also note that $\Delta_i(\hat{q}s^2) = (-8tn^{-3/2} + O(L^2n^{-3}))\hat{q}s^2$ and $\Delta_i(\delta_O\hat{q}s^2) = (1+o(1))((L+t)^{-1} - 8t)n^{-3/2}\delta_O\hat{q}s^2$. When Q_S is in the critical interval we have

$$\begin{aligned}\mathbb{E}[\Delta_i\mathcal{Z}Q_S \mid \mathcal{F}_i] &\leq -8tn^{-3/2}\hat{q}s^2(1 + \delta_O - g_O - O(\delta_Y)) \\ &\quad + \hat{q}s^2\left(8tn^{-3/2} - O(L^2n^{-3}) - (1+o(1))((L+t)^{-1} - 8t)n^{-3/2}\delta_O\right). \\ &\leq -8tn^{-3/2}\hat{q}s^2 \cdot (\delta_O - g_O - O(\delta_Y) + (\tfrac{1}{8t(L+t)} - 1)\delta_O)\end{aligned}$$

Using $\frac{\delta_O}{8t(L+t)} \geq 2g_O$ and $\delta_Y \leq n^{-\varepsilon/4} = o(g_O)$ by (14), when Q_S is in the critical interval we have $\mathbb{E}[\Delta_i \mathcal{Z} Q_S \mid \mathcal{F}_i] \leq 0$, so the trend hypothesis holds.

To complete the proof of statements (i) and (ii) we will apply Freedman's inequality and take a union bound over S. To account for the number $\binom{n}{s}$ of events in the union, it suffices to establish the following strengthened form of the bounded hypothesis (2) and (3), where we write N_O and Var_O for the maximum one-step change and conditional variance of Q_S.

$$g_O(t)^2(\hat{q}(t)s^2)^2 = \omega\left(\mathrm{Var}_O(i)(n\log n)^{3/2}s\right), \tag{46}$$

$$g_O(t)\hat{q}(t)s^2 = \omega\left(N_O(i)(\log n)s\right). \tag{47}$$

Since $g_O = \psi/(40L^2)$, it suffices to show $N_O \leq 2L^{-10}\psi^2\hat{q}s$, as by Lemma 7.1 this also implies $\mathrm{Var}_O \leq L^{-4}n^{-3/2}s^{-1}(2L^{-2}\psi\hat{q}s^2)^2$. To see this bound on N_O we use $N_O = O(y)$ for statement (ii), or $N_O \leq |N(x) \cap S| + |N(y) \cap S|$ and our assumption on neighborhoods in S for statement (i).

It remains to prove (iii), which is a one-sided bound rather than a dynamic concentration statement, but we can still apply a modified form of the critical interval method. Writing $F_O = (1 + t/L)L^{13}s\hat{q}\sqrt{n}/2$, it suffices to show $Q_S \leq F_O$ for all S with high probability. Note that the bound is trivial for $t \leq 1$, as $s < 2L^{12}\sqrt{n}$ implies $Q_S \leq s^2 < F_O$. For $t \geq 1$ we use critical window analysis with the window $[F_O - G_O, F_O]$, where $G_O = F_O/(40L^2)$. (Here we use capital letters F, G to distinguish our notation for absolute errors from our usual notation f, g for relative errors.)

When Q_S is in the critical window we estimate $\mathbb{E}[\Delta_i Q_S \mid \mathcal{F}_i] \leq -(1 + o(1))8tn^{-3/2}(F_O - G_O)$. We write $\mathcal{Z}Q_S = Q_S - F_O$ and note that $F_O' = ((L+t)^{-1} - 8t)F_O$. Again using $\frac{F_O}{8t(L+t)} \geq 2G_O$, we obtain the trend hypothesis

$$\mathbb{E}[\Delta_i \mathcal{Z} Q_S \mid \mathcal{F}_i] \leq -(1 + o(1))8tn^{-3/2} \cdot (F_O - G_O + (\tfrac{1}{8t(L+t)} - 1)F_O) \leq 0.$$

For the boundedness hypothesis, accounting for the union bound as in (i) and (ii), and noting that $\mathcal{Z}Q_S(i) < -G_O(t)$ at the step before this variable enters the critical interval, it suffices to show

$$G_O(t)^2 = \omega\left(\mathrm{Var}_O(i)(n\log n)^{3/2}s\right) \quad \text{and} \quad G_O(t) = \omega\left(N_O(i)(\log n)s\right). \tag{48}$$

We use the bound $N_O \leq 2y \leq L^{-9}s^{-1}G_O$. By Lemma 7.1 this implies

$$\mathrm{Var}_O = O\left(tn^{-3/2} \cdot F_O \cdot L^{-9}s^{-1}G_O\right) = O\left(L^{-6}s^{-1}G_O^2n^{-3/2}\right),$$

and the desired inequalities follow. ∎

Lemma 7.4. *Suppose $r, s \geq n^{1/4}$, $\psi \geq n^{-\varepsilon/5}$, and $h \leq L^{-10}\psi^2\hat{q}\min\{r, s\}$. Then whp we have $Q_{RS} = (1 \pm \psi)\hat{q}rs$ for any sets R, S of respective sizes r, s such that any vertex that has a neighbor in one of these sets has at most h neighbors in the other.*

Note that Lemma 7.4 is simply a bipartite version of Lemma 7.3(i). The proof is essentially the same, so we omit it, noting that the condition on h is needed for the boundedness hypothesis.

Next we establish some density estimates.

Definition 7.5. For a set S, let η_S denote the number of edges of $G(i_{\max})$ in S.

Lemma 7.6. *Whp for any set S of size s*

(i) if $s \geq L^{12}\sqrt{n}$ then $\eta_S < L^2 n^{-1/2} s^2$,
(ii) if $s < 2L^{12}\sqrt{n}$ then $\eta_S < L^{15} s$.

Proof. For (i), we estimate the probability that some such S spans $M := L^2 n^{-1/2} s^2$ edges, taking a union bound over S and the steps at which the edges are chosen, for which there are $\binom{n}{s}\binom{m}{M}$ choices. For a specified step at time t, the probability of choosing an edge in S is $Q_S(t)/Q(t) = (1+o(1))s^2/n^2$, using Lemma 7.3(ii) with $\psi = L^{-1/2}$. Thus the failure probability p_0 satisfies

$$p_0 \leq \binom{n}{s}\binom{m}{M}((1+o(1))s^2/n^2)^M.$$

Noting that $M \geq L^{14}s$, the required estimate $p_0 = o(1)$ follows from

$$\log p_0 \leq O(s\log n) + M\log\tfrac{em}{M} + M\log((1+o(1))s^2/n^2) = O(s\log n) + M(O(1) - \log L) \leq -sL^{14}.$$

For (ii), we estimate the probability of choosing an edge in S as $Q_S(t)/Q(t) < 2L^{13}sn^{-3/2}$ by Lemma 7.3(iii). Then the failure probability p_0 satisfies

$$p_0 \leq \binom{n}{s}\binom{m}{L^{15}s}(2L^{13}sn^{-3/2})^{L^{15}s},$$

so $s^{-1}\log p_0 \leq O(\log n) + L^{15}\log\frac{2em}{L^2n^{3/2}} \leq -L^{15}$, giving $p_0 = o(1)$. ∎

Next we deduce a bound on the number of vertices of large degree in a given set. For the following definition we emphasize that vertices in S can belong to $D_d(S)$.

Definition 7.7. Let $D_d(S)$ be the set of vertices that have degree at least d in S.

Lemma 7.8. *Whp for any set S of size s*

(i) if $s \geq L^{12}\sqrt{n}$ and $d > 8L^2n^{-1/2}s$ then $|D_d(S)| < 8L^2n^{-1/2}s^2/d$,
(ii) if $s < L^{12}\sqrt{n}$ and $d > 4L^{15}$ then $|D_d(S)| < 4L^{15}s/d$.

Proof. For (i), suppose on the contrary that there is $T \subseteq D_d(S)$ of size $8L^2n^{-1/2}s^2/d$. Then $S \cup T$ is a set of size at most $2s$ that spans at least $d|T|/2 > L^2n^{-1/2}(2s)^2$ edges, which contradicts Lemma 7.6(i). Similarly, for (ii), if there is $T \subseteq D_d(S)$ of size $4L^{15}s/d$ then $|S \cup T| \leq 2s \leq 2L^{12}\sqrt{n}$ and $\eta_{S\cup T} \geq d|T|/2 > 2L^{15}s \geq L^{15}|S \cup T|$, which contradicts Lemma 7.6(ii). ∎

We conclude this preliminary subsection with an estimate for a more involved configuration required for the proof of Lemma 7.12, using the constants $0 < \alpha < \gamma < \beta$ declared in (7). To motivate the following definition, we remark that it will be applied with $H \subseteq N(x)$, that is, the neighborhood of x in $G(i_{\max})$, which will justify the assumed bounds on degrees and open degrees into H for $a \neq x$, and also that H contains no edges. Furthermore, it will be applied at a step $i < i_{\max}$ at which H only contains vertices y such that xy is open and yet to be chosen as an edge, so there will be no edges between x and H.

Definition 7.9. Let $H \subseteq V$ and $x \in V \setminus H$. We say (x, H) is *neighborly* at step $i < i_{\max}$ if $G(i)$ has no edges within $H \cup \{x\}$ and for any vertex $a \neq x$ at most L^4 edges ab with $b \in H$ and at most $2x = 2\hat{q}^2 n$

open pairs ab with $b \in H$. We let W_{xH} denote the number of ordered triples (a, b, c) of vertices such that ax is open, $\{b, c\} \subseteq H$ and ab, ac are edges.

Lemma 7.10. *Whp for every neighborly (x, H) with $|H| = h$ where $L^{\alpha} < h < L^{-\beta}\sqrt{n}$ we have $W_{xH} < 4L^{-\alpha}h\hat{q}\sqrt{n}$.*

Proof. We will apply the critical interval method, although we cannot do so directly for W_{xH} as the boundedness hypothesis may fail due to vertices with large open degree into H; thus we will make some subtle alterations to the structures that we count.

We start with some definitions. We say that a vertex a is *obese* with respect to H at time t if at least $\hat{q}\sqrt{n}L^{\gamma}$ pairs ab with $b \in H$ are open. (Our extravagant nomenclature here is explained by reference to the definition of "heavy" below.) For any obese vertex a we declare some subset of the open pairs ab with $b \in H$ *inactive* so that the active open degree into H is $\lfloor \hat{q}\sqrt{n}L^{\gamma} \rfloor$.

We stress that the status of an open pair as active or inactive can change back and forth in the course of the process, but once a pair is chosen as an edge its status as active or inactive remains the same for the rest of the process.

For $j \in \{0, 1, 2\}$ let W^j_{xH} denote the number of ordered triples (a, b, c) of vertices such that ax is open, $\{b, c\} \subseteq H$, the pairs ab and ac are both active, and their status depends on j: if $j = 0$ then both are open, if $j = 2$ then both are edges, and if $j = 1$ then ab is open and ac is an edge. Thus W^2_{xH} has the same definition as W_{xH}, with the additional condition that ab and ac are active at the steps they are chosen as edges.

First we show that there is a negligible difference between W_{xH} and W^2_{xH}, and so it will suffice to bound the latter. Let O be the set of vertices that are obese with respect to H at time t. We claim that whp for any H we have

$$|O| < 2hL^{13-\gamma} = o(h). \tag{49}$$

To see this, suppose on the contrary there is $O' \subseteq O$ of size $2hL^{13-\gamma}$. Then $|H \cup O'| < 2h$ and $Q_{H\cup O'} \geq L^{13}h\hat{q}\sqrt{n}$. However, this contradicts Lemma 7.3(iii), so (49) holds.

Applying Lemma 7.3(iii) again, we bound the number of open pairs in $H \cup O$ by $Q_{H\cup O} < L^{13}\hat{q}\sqrt{n} \cdot 3h/2$. Thus the probability at any given step that we choose an edge between an obese vertex and H is at most $2hL^{13}n^{-3/2}$. For each set H let $\mathcal{O}_H$ be the event that the process chooses at least hL^{15} edges between H and obese vertices (recalling that the set of obese vertices may change as the process evolves). By the union bound, the probability that any $\mathcal{O}_H$ holds is at most

$$\sum_{h=L^{\alpha}}^{n^{1/2}L^{-\beta}} \binom{n}{h}\binom{m}{hL^{15}}(2hL^{13}n^{-3/2})^{hL^{15}} \leq \sum_{h=L^{\alpha}}^{n^{1/2}L^{-\beta}} \binom{n}{h}\left(\frac{O(1)}{L}\right)^{hL^{15}} = o(1).$$

Thus we can assume that no event $\mathcal{O}_H$ holds. Then the degree bound for neighborly (x, H) implies $W_{xH} - W^2_{xH} < hL^{19}$, which is negligible by comparison with the desired bound on W_{xH}.

For the remainder of the proof, we will show $W_j \leq F_j := (1 + t/L)w_j/2$ for $j = 0, 1, 2$, where

$$w_0 := L^{-\alpha-4}hx\hat{q}\sqrt{n}, \; w_1 := L^{-\alpha-2}hy\hat{q}\sqrt{n} \text{ and } w_2 := 4L^{-\alpha}h\hat{q}\sqrt{n}.$$

This will suffice to prove the lemma, as we will have $W_{xH} < W^2_{xH} + hL^{19} < F_2 + hL^{19} < w_2$. Similarly to the proof of Lemma 7.3(iii), these are one-sided bounds rather than dynamic concentration statements, but we can still use a modified form of the critical interval method. For W^j_{xH} we use the critical windows $[F_j - G_j, F_j]$, where $G_j = w_j/(40L^2)$.

First we claim that our variables do not enter their critical windows for $n^{-1/4} \le t \le 1$ (assuming $I > i_{\max}$). For $j = 0$ this follows from the trivial bound $W^0_{xH} \le nh^2 \ll w_0(1)$, recalling that β is large compared with α. For $j = 1$ we can bound W^1_{xH} by picking $\{b, c\} \subseteq H$ then a vertex counted by $Y_{bc}(i)$, so by (14) we obtain $W^1_{xH} \le O(y)h^2 \ll w_1$. For $j = 2$ we bound W^2_{xH} by picking $\{b, c\} \subseteq H$ then a common neighbor, for which there are at most $O(L^4)$ choices by Definition 2.12(iii), so $W^2_{xH} = O(L^4h^2) \ll w_2$. Thus the claim holds.

Next we will prove the trend hypothesis, that is, that $\mathcal{Z}W^j_{xH} = W^j_{xH} - F_j$ is a supermartingale while W^j_{xH} is in its critical window. Below we will analyze the contributions to $\mathbb{E}[\Delta_i \mathcal{Z}W^j_{xH} \mid \mathcal{F}_i]$ separately according to each of the pairs ax, ab, ac. When we calculate the expected change due to closing of ab or ac we will ignore correction terms due to changes that do not actually occur when a is obese and these closures simply change the status of some other open pair from inactive to active. To justify this, we first give upper bounds on these correction terms, which we will later see are negligible compared with the main terms.

For $a \in O$ let A_a denote the set of $b \in H$ such that ab is open and active. By (49), the contribution to $\mathbb{E}[\Delta_i W^0_{xH}]$ due to closing a pair ab or ac where a is obese is at most

$$2Q^{-1} \sum_{a \in O} \sum_{b \in A_a} (Y_{ab} + Y_{ba})|A_a| \le 5yq^{-1} \cdot 2hL^{13-\gamma} \cdot (\hat{q}n^{1/2}L^{\gamma})^2 \ll 8tn^{-3/2}F_0L^{-2}. \tag{50}$$

Similarly, the contributions to $\mathbb{E}[\Delta_i W^1_{xH}]$ due to closing a pair ac where a is obese is at most

$$2Q^{-1} \sum_{a \in O} \sum_{b \in A_a} (Y_{ab} + Y_{ba})L^4 \le 5yq^{-1}L^4 \cdot 2hL^{13-\gamma} \cdot \hat{q}n^{1/2}L^{\gamma} \ll 8tn^{-3/2}F_1L^{-2}. \tag{51}$$

In the calculation of the expected change in $\mathcal{Z}W^j_{xH} = W^j_{xH} - F_j$ we write

$$\Delta_i(F_j) = (1 + o(1))F'_j n^{-3/2} \quad \text{and} \quad F'_j \ge ((L+t)^{-1} - (3-j)8t)F_j.$$

For each open pair $\alpha\beta$ we have a destruction term of

$$2Q^{-1} \sum_{f \in W^j_{xH}} (Y_{f(\alpha\beta)} + Y_{f(\beta\alpha)} + 1) \ge (1 + o(1))8tn^{-3/2}(F_j - G_j),$$

when W^j_{xH} is in the critical interval. This gives self-correction against a corresponding $8tn^{-3/2}F_j$ term in $\Delta_i(F_j)$. For each edge we have a creation term of

$$2Q^{-1}W^{j-1}_{xH} \le (1 + o(1))2q^{-1}F_{j-1},$$

where $2q^{-1}F_0 = L^{-2}t^{-1}n^{-3/2}F_1$ and $2q^{-1}F_1 = tL^{-2}n^{-3/2}F_2$.

Next we account for fidelity corrections. As there are no edges within $H \cup \{x\}$ there is no creation fidelity term (it is not possible to add an edge and simultaneously close an open pair in the configuration). For destruction fidelity, we first consider configurations for $j = 0, 1$ in which selecting an edge az simultaneously closes the open pairs ab and ax. There are at most h choices for c, then $2v$ choices for a where $v = x$ for $j = 0$ or $v = y$ for $j = 1$, then L^4 choices for z in the common neighborhood of b and x, then $2y$ choices for $b \in Y_{az}$. This gives a correction term $O(q^{-1}hvL^4y) \ll 8tn^{-3/2}F_jL^{-2}$. For $j = 0$ we also need to consider configurations in which selecting az simultaneously closes ab and ac. There are at most h choices of b, then $2y$ choices of z in Y_{xb}, then $2x$ choices of a in X_{ab}, then L^4 choices of a neighbor c of z in H (as (x, H) is neighborly). This gives a correction term $O(hyxL^4) \ll 8tn^{-3/2}F_0L^{-2}$. Using $\frac{F_j}{8t(L+t)} \geq 4G_j$, we obtain

$$\mathbb{E}[\Delta_i \mathcal{Z} W^0_{xH} \mid \mathcal{F}_i] \leq -(1+o(1))8tn^{-3/2} \cdot (3(F_0 - G_0) + (\tfrac{1}{8t(L+t)} - 3)F_0) \leq 0.$$
$$\mathbb{E}[\Delta_i \mathcal{Z} W^1_{xH} \mid \mathcal{F}_i] \leq -(1+o(1))8tn^{-3/2} \cdot (2(F_1 - G_1) - \tfrac{1}{8L^2t^2}F_1 + (\tfrac{1}{8t(L+t)} - 2)F_1) \leq 0.$$
$$\mathbb{E}[\Delta_i \mathcal{Z} W^2_{xH} \mid \mathcal{F}_i] \leq -(1+o(1))8tn^{-3/2} \cdot (F_2 - G_2 - \tfrac{1}{8L^2}F_2 + (\tfrac{1}{8t(L+t)} - 1)F_2) \leq 0.$$

Note that the correction terms (50) and (51) for inactive edges and the fidelity terms are indeed negligible in this calculation, so the trend hypothesis holds.

It remains to establish the boundedness hypothesis. Note that since we can restrict our attention to $t \geq 1$, the functions G_j are approximately nonincreasing. As we are proving one-sided bounds with a union bound over the choice of x and H, it suffices to establish the boundedness hypothesis as set forth in (48) with h playing the role of s. We add an additional wrinkle here. Recall that Freedman's inequality (Lemma 2.1) only requires a bound on the *positive* change in the random variable in question. For each pair e in the collection ax, ab, ac let N_e^+ bound the *positive* one-step change in $\mathcal{Z}W^j_{xH}$ due to the change in the status of e and let Var_e denote the one-step variance of $\mathcal{Z}W^j_{xH}$ that can be attributed to the change in status of e. To apply Freedman's inequality, since $G_j = w_j/(40L^2)$, it suffices to show

$$N_e^+ \leq w_j/(hL^5) \quad \text{and} \quad \mathrm{Var}_e \leq w_j^2/(hL^8n^{3/2}). \tag{52}$$

In some cases we will show the stronger statement

$$N_e < w_j/(hL^{10}), \tag{53}$$

where N_e is the absolute value of the one-step change in $\mathcal{Z}W^j_{xH}$. Note that (53) clearly implies (52): the bound on N_e^+ is immediate and the bound for Var_e follows by Lemma 7.1.

First we note that the required bounds for creation are straightforward. Indeed, for W^1_{xH} the bound on active open degrees gives $N_e \leq \hat{q}\sqrt{n}L^\gamma \ll w_1/(hL^{10})$, and for W^2_{xH} the assumption that (x, H) is neighborly gives $N_e \leq L^4 \ll w_2/(hL^{10})$.

For destruction we obtain negative changes in $\mathcal{Z}W^j_{xH}$, so we only need to bound Var_e. First we introduce some additional definitions. We say that a vertex a is *heavy* with respect to H at time t if at least $\hat{q}\sqrt{n}L^{-\gamma}$ pairs ab with $b \in H$ are open. Let $T = T_{xH}$ be the set of heavy vertices a such that xa is open. As $|T|\hat{q}\sqrt{n}L^{-\gamma} \leq \sum_{u \in H} X_{ux} < 2xh$, we have

$$|T| < 2hx/(\hat{q}\sqrt{n}L^{-\gamma}) = 2hL^\gamma\hat{q}\sqrt{n}.$$

Let U be the set of vertices z such that zx is open and z has at least $\hat{q}\sqrt{n}L^{-3\gamma}$ neighbors in T. By Lemma 7.8 we have

$$|U| < \begin{cases} 8hL^{4\gamma+15} & \text{if } |T| < L^{12}\sqrt{n}, \\ 32h^2L^{5\gamma+2}\hat{q} & \text{otherwise.} \end{cases}$$

Here we used $\hat{q}\sqrt{n}L^{-3\gamma} > 4L^{15}$ and $\hat{q}\sqrt{n}L^{-3\gamma} > 8L^2n^{-1/2} \cdot 2hL^\gamma\hat{q}\sqrt{n}$, which follows from our choice of β to be large relative to γ, to get the lower bounds on d required for Lemma 7.8.

Now consider destruction for the variables W^j_{xH} for $j = 0, 1$. We write $\Delta_i W^j_{xH} = \Delta_i V_1 + \Delta_i V_2$, where $\Delta_i V_1$ accounts for the change in $V = W^j_{xH}$ that comes from the choice of an edge xz where $z \in U$, and

$\Delta_i V_2$ accounts for the rest. For $\Delta_i V_2$ we will obtain the required bound on Var_e by establishing the bound (53) on N_e. The contribution to N_e from closing ab or ac is bounded by $2y\hat{q}n^{1/2}L^{\gamma} < w_0/(hL^{10})$ for W^0_{xH} (using the bound on active open degrees) and by $2yL^4 < w_1/(hL^{10})$ for W^1_{xH} (as (x,H) is neighborly). Next we consider the contribution from closing xa where a is not heavy. For $j=0$ this is at most

$$(2y)(\hat{q}\sqrt{n}L^{-\gamma})^2 < 2\hat{q}\sqrt{n}L^{1-2\gamma}x < L^{-\alpha-12}x\hat{q}\sqrt{n} = w_0/(hL^{10}),$$

as γ is large relative to α. For $j = 1$, as (x,H) is neighborly, the contribution is at most $(2y)(\hat{q}\sqrt{n}L^{-\gamma})L^4 < L^{-\alpha-12}y\hat{q}\sqrt{n} = w_1/(hL^{10})$, again as γ is large relative to α. Now we consider the contribution from closing of pairs xa where a is heavy. Note that we do not select xz with $z \in U$, as this case will be analyzed in $\Delta_i V_1$, so this contribution is at most

$$\hat{q}\sqrt{n}L^{-3\gamma}(\hat{q}\sqrt{n}L^{\gamma})^2 = L^{-\gamma}x\hat{q}\sqrt{n} \ll w_0/(hL^{10})$$

for $j = 0$ (by the bound on active open degrees), or $\hat{q}\sqrt{n}L^{-3\gamma}(\hat{q}\sqrt{n}L^{\gamma})L^4 \ll w_1/(hL^{10})$ for $j = 1$ (as (x,H) is neighborly and $t \geq 1$). Thus we have the required bound on N_e for $\Delta_i V_2$.

For $j = 0, 1$ it remains to bound Var_e for $\Delta_i V_1$. The probability that an edge xz with $z \in U$ is chosen is at most $2|U|/q$, and the resulting change in W^j_{xH} is at most $(2y)(\hat{q}\sqrt{n}L^{\gamma})^2$ for $j = 0$, or $(2y)(\hat{q}\sqrt{n}L^{\gamma})L^4$ for $j = 1$. Suppose first that $|T| < L^{12}\sqrt{n}$, so that $|U| < 8hL^{4\gamma+15}$. Then for $j = 0$ we have

$$\mathrm{Var}_e \leq 16hL^{4\gamma+15}q^{-1}(2y)^2(\hat{q}\sqrt{n}L^{\gamma})^4 = \widetilde{O}(h\hat{q}^5 n),$$

which suffices to establish (52) as $w_0^2/(hL^8n^{3/2}) = \widetilde{\Omega}(h\hat{q}^6n^{3/2})$. Also, for $j = 1$ we have

$$\mathrm{Var}_e \leq 16hL^{4\gamma+15}q^{-1}(2y)^2(\hat{q}\sqrt{n}L^{\gamma})^2L^8 = \widetilde{O}(h\hat{q}^3),$$

which suffices as $w_1^2/(hL^8n^{3/2}) = \widetilde{\Omega}(h\hat{q}^4n^{1/2})$, recalling that $t \geq 1$. Now suppose $|T| \geq L^{12}\sqrt{n}$, so that $|U| < 32h^2L^{5\gamma+2}\hat{q}$. Then for $j = 0$ we have

$$\mathrm{Var}_e \leq 64h^2L^{5\gamma+2}n^{-2}(2y)^2(\hat{q}\sqrt{n}L^{\gamma})^4 < 256h^2L^{9\gamma+4}\hat{q}^6 n,$$

and for $j = 1$ we have

$$\mathrm{Var}_e \leq 64h^2L^{5\gamma+2}n^{-2}(2y)^2(\hat{q}\sqrt{n}L^{\gamma})^2L^8 < 256h^2L^{7\gamma+12}\hat{q}^4.$$

As $h < L^{-\beta}\sqrt{n}$ and β is large relative to α, γ these bounds suffice to establish (52).

It remains to bound Var_e for destruction of W^2_{xH}. Let W be the set of vertices that are open to x and have at least two neighbors in H. Then $|W| \leq \sum_{a \in H} Y_{xa} < 2yh$. Let U' be the set of vertices that are open to x and have at least $yL^{-\gamma}$ neighbors in W. By Lemma 7.8 we have

$$|U'| < \begin{cases} 8hL^{\gamma+15} & \text{if } |W| < L^{12}\sqrt{n} \\ 32h^2yn^{-1/2}L^{\gamma+2} & \text{otherwise.} \end{cases}$$

Here we used $yL^{-\gamma} > 4L^{15}$ and $yL^{-\gamma} > 8L^2n^{-1/2} \cdot 2yh$ (as β is large relative to γ) to get the lower bound on d required for Lemma 7.8. We write the destruction of W^2_{xH} at step i as $\Delta_i V_1 + \Delta_i V_2$, where

$\Delta_i V_1$ accounts for the change in W^2_{xH} that comes from the choice of an edge xz where $z \in U'$, and $\Delta_i V_2$ accounts for the rest.

For $\Delta_i V_2$ we can obtain the required bound on Var_e from the bound (53) on N_e; indeed, by definition of U' and as (x, H) is neighborly, $N_e < yL^{-\gamma} \cdot L^4 < w_2/(hL^{10})$. For $\Delta_i V_1$, suppose first that $|W| < L^{12}\sqrt{n}$, so that $|U'| < 8hL^{\gamma+15}$. We choose an edge xz with $z \in U'$ with probability at most $2|U|/q$, and as (x, H) is neighborly the resulting change in W^2_{xH} is at most $2y \cdot L^4$, so

$$\mathrm{Var}_e < 8hL^{\gamma+15}q^{-1}(2y)^2L^{16} = \widetilde{O}(hyn^{-3/2}),$$

which suffices as $w_2^2/(hL^8n^{3/2}) = \widetilde{\Omega}(hy^2n^{-3/2})$. On the other hand, if $|W| \geq L^{12}\sqrt{n}$ then

$$\mathrm{Var}_e < 32h^2yn^{-1/2}L^{\gamma+2}q^{-1}(2y)^2L^{16} < 128h^2n^{-1/2}L^{\gamma+19}y^2n^{-3/2},$$

which also suffices to establish (52) as β is large relative to α, γ. ■

7.2 | Proof of Theorem 1.2

We will show whp

$$\alpha(G) < k := (1 + 3\varepsilon)\sqrt{2n \log n}.$$

As $\alpha(G) \leq \alpha(G(i_{\max}))$, it suffices to bound $\alpha(G(i_{\max}))$. We need to estimate the probability that there is an independent set K of size k. As discussed above, we will take a union bound over all such sets K together with certain information about how neighborhoods in $G(i_{\max})$ intersect K.

Let K be a potential independent set of size k. We define a sequence of vertices $x_1, \ldots, x_z$, where each x_ℓ is chosen to maximize the number of neighbors in K that are not also neighbors of some x_j for $j < \ell$. More precisely, the ℓth *hole* is $H_\ell = (N(x_\ell) \setminus \cup_{\ell' < \ell} N(x_{\ell'})) \cap K$, where x_ℓ is chosen to maximize $h_\ell = |H_\ell|$, and we recall our convention that all neighborhoods are defined with respect to $G(i_{\max})$. We stop the sequence if there are no vertices that give more than $L^{2\alpha}$ new neighbors in K. Note that $x_\ell \notin K$ for $\ell \in [z]$, as K is independent. We say that a hole is *large* if it has size more than $L^{-\beta}\sqrt{n}$. We let Z_A be the set of ℓ such that H_ℓ is large,

$$Z_B = [z] \setminus Z_A, \quad A = \cup_{\ell \in Z_A} H_\ell, \quad B = \cup_{\ell \in Z_B} H_\ell, \quad C = K \setminus (A \cup B).$$

For $\ell \in Z_B$ we specify the steps of the process at which the edges between x_ℓ and H_ℓ appear. We write $H_\ell = \{v_{\ell j} : j \in [h_\ell]\}$, where $x_\ell v_{\ell j}$ is selected at step $i_{\ell j}$, and $i_{\ell j}$ is increasing in j. For $\ell \in Z_A$ we specify the entire neighborhood of x_ℓ in $G(i_{\max})$: we write $d_\ell = |N(x_\ell)|$ and $N(x_\ell) = \{v_{\ell j} : j \in [d_\ell]\}$, where $x_\ell v_{\ell j}$ is selected at step $i_{\ell j}$, and $i_{\ell j}$ is increasing in j. We will estimate $\mathbb{P}(\mathcal{E})$, where $\mathcal{E}$ is the event that there is an independent set K with some fixed choices of z; x_ℓ and h_ℓ for $\ell \in [z]$; and d_ℓ for $\ell \in Z_A$. We will refer to these choices of hole sizes, vertices with large neighborhoods in K and vertex degrees as the *initial data* that defines $\mathcal{E}$. Note that by Lemma 7.8(ii) we can assume

$$|Z_A| < 8L^{16+\beta} \quad \text{and} \quad z < 4L^{15-2\alpha}k. \tag{54}$$

For $\ell \in Z_A, j \in [d_\ell]$ we claim that

$$i_{\ell j} = jn/2 \pm n^{3/2-\varepsilon/3} \quad \text{and} \quad d_\ell = d \pm n^{1/2-\varepsilon/3}, \tag{55}$$

where we recall $d = 2t_{\max}\sqrt{n} = 2m/n = \sqrt{(1/2 - \varepsilon)n \log n}$. To see (55), note that if, for example, we had $i = i_{\ell j} < jn/2 - n^{3/2-\varepsilon/3}$ then we would have $Y_{x_\ell}(i) \geq j > 2n^{-1}(i + n^{3/2-\varepsilon/3}) = y_1(t) + 2n^{1/2-\varepsilon/3}$, which contradicts the degree bounds $Y_u(i) = (1 \pm \delta_{Y_1}(t))y_1(t)$ in the event $\mathcal{G}_i$ (see Definition 2.12).

Now, in addition to the initial data, we fix the independent set K, the specific edges $x_\ell v_{\ell j}$ and appearance times $i_{\ell j}$ for $\ell \in Z_A, j \in [d_\ell]$, and likewise for $\ell \in Z_B, j \in [h_\ell]$. We let $\mathcal{E}_K$ be the event that K is independent and all the specified edges appear at the specified steps of the process. Thus $\mathcal{E}$ is a union of events of the form $\mathcal{E}_K$.

To estimate the probability of any given event $\mathcal{E}_K$, for each step i we need to estimate the probability that the selected edge is compatible with $\mathcal{E}_K$, conditional on the history of the process. We say i is a *selection step* if i is one of $i_{\ell j}$ for $\ell \in Z_A, j \in [d_\ell]$ or $\ell \in Z_B, j \in [h_\ell]$; then the selected edge is specified by $\mathcal{E}_K$, so the required probability is simply $2/Q = (1 \pm 2\delta_Q)2q^{-1}$. For other i, the required probability is $1 - N_i/Q$, where N_i is the number of ordered open pairs that cannot be selected at step i when $\mathcal{E}_K$ occurs. If $i = i_{\ell j}$ is a selection step write $N_i = 0$. Then we estimate

$$\mathbb{P}(\mathcal{E}_K) \le \prod_{\ell \in Z_A} \prod_{j=1}^{d_\ell} (1 \pm 2\delta_Q) 2q(t_{\ell j})^{-1} \cdot \prod_{\ell \in Z_B} \prod_{j=1}^{h_\ell} (1 \pm 2\delta_Q) 2q(t_{\ell j})^{-1} \cdot \prod_{i=1}^{m} (1 - N_i/Q). \tag{56}$$

To estimate N_i, we classify open pairs that cannot be selected at step i as follows.

- Let N_{iAi} be the number of ordered open pairs of the form $v_{\ell j} v_{\ell j'}$ for some $\ell \in Z_A, j, j' \in [d_\ell]$.
- Let N_{iAo} be the number of ordered open pairs of the form $x_\ell y$ or $y x_\ell$ where $\ell \in Z_A$ and $y \notin N(x_\ell) \cup K \cup \{x_1, \dots, x_z\}$.
- Let N_{iBi} be the number of ordered open pairs ab such that $B \cap ab \neq \emptyset$ and selecting $e_i = ab$ would close an open pair of the form $x_\ell v_{\ell j}$ for $\ell \in Z_B, j \in [h_\ell]$.
- Let N_{iBo} be the number of ordered open pairs ab such that $B \cap ab = \emptyset$ and selecting $e_i = ab$ would close an open pair of the form $x_\ell v_{\ell j}$ for $\ell \in Z_B, j \in [h_\ell]$.
- Let N_{iK} be the number of ordered open pairs in K that are not contained within any hole.

We refer to pairs counted by N_{iAo} or N_{iBo} as *outer* and those counted by N_{iAi} or N_{iBi} as *inner* (which is indicated by one of the $i's$ in the notation; the other refers to the step i, which we hope will not cause confusion). For $\ell \in Z_A$ we stress that by naming the $v_{\ell j}$'s we have specified all neighbors of x_ℓ (not only those in K), so we cannot select a pair $y x_\ell$ with $y \notin N(x_\ell)$; we also exclude $y \in K \cup \{x_1, \dots, x_z\}$ in the definition of N_{iAo} to facilitate the estimate for overcounting in Lemma 7.15. For N_{iK} we note that all open pairs within K are forbidden (as K is independent) but again to avoid overcounting we only include those not contained within any hole. We write

$$N_i \ge N_{iAi} + N_{iAo} + N_{iB} + N_{iK} - N_{iO},$$

where $N_{iB} = N_{iBi} + N_{iBo}$ and N_{iO} corrects for any open pairs that appear in more than one of the above collections. (We will see that the most significant source of overcounting comes from pairs counted by both N_{iK} and N_{iBi}.) We substitute

$$1 - N_i/Q \le \exp\left\{-(1 - 2\delta_Q) q^{-1} (N_{iAi} + N_{iAo} + N_{iB} + N_{iK} - N_{iO})\right\} \tag{57}$$

in (56), recalling that $\delta_Q = O(n^{-\epsilon/5})$, to obtain

$$\begin{aligned} -\log \mathbb{P}(\mathcal{E}_K) \ge{}& S_{Ai} - T_A + S_B - T_B + S_{Ao} + S_K - S_O \\ &+ \log \frac{n^2}{2} \left(\sum_{\ell \in Z_A} d_\ell + |B| \right) - O(n^{1/2-\epsilon/5}), \text{ where} \end{aligned} \tag{58}$$

$$S_\mu = \sum_{i=1}^{m} N_{i\mu} q^{-1} \qquad \text{for } \mu \in \{Ai, Ao, B, K, O\},$$

$$T_A = \sum_{\ell \in Z_A} \sum_{j=1}^{d_\ell} 4t_{\ell j}^2 \qquad \text{and} \qquad T_B = \sum_{\ell \in Z_B} \sum_{j=1}^{h_\ell} 4t_{\ell j}^2.$$

To estimate the terms in (58), we start by showing in the next two lemmas that $S_{Ai} - T_A$ and $S_B - T_B$ are negligible. (The remaining terms will be used to balance the number of events in our union bound calculation.)

Lemma 7.11. $T_A - S_{Ai} < O(n^{1/2-\varepsilon/5})$.

Proof. We start by giving a lower bound on N_{iAi} for any i that is not a selection step. For $\ell \in Z_A$ let $j_\ell = j_\ell(i)$ be the value of $j \in [d_\ell]$ such that $i_{\ell(j-1)} \le i < i_{\ell j}$, where $i_{\ell 0} := 0$, that is, $j_\ell - 1$ edges have been selected at x_ℓ. Let $S_\ell = \{v_{\ell j}\}_{j=j_\ell+1}^{d_\ell}$ and $s_\ell = |S_\ell| = d_\ell + 1 - j_\ell$; thus $\{x_\ell v : v \in S_\ell\}$ is the set of open pairs at x_ℓ that will later be selected as edges. As we consider the whole neighborhood of x_ℓ (not just the neighborhood in K), the number of ordered open pairs $v_{\ell j} v_{\ell j'}$ with $j > j_\ell, j' \le j_\ell$ is $\sum_{v \in S_\ell} 2Y_{vx_\ell} = (1 \pm \delta_Y) 2ys_\ell$.

We also note that any vertex has at most L^4 neighbors in S_ℓ by the codegree bound in $G(i_{\max})$, which is valid as we assume $I < i_{\max}$. Then by Lemma 7.3(i) whp $Q_{S_\ell} = (1 \pm n^{-\varepsilon/5}) \hat{q} s_\ell^2$ if $s_\ell > n^{1/4}$ and $\hat{q} s_\ell \ge n^{2\varepsilon/5} L^{14}$. Since $\hat{q} \ge n^{-1/2+\varepsilon}$ this holds for $s_\ell > n^{1/2-\varepsilon/2}$, so we can write $Q_{S_\ell} \ge (1 - n^{-\varepsilon/5}) \hat{q} s_\ell (s_\ell - n^{1/2-\varepsilon/2})$, as this bound is trivial for $s_\ell \le n^{1/2-\varepsilon/2}$. The bound on codegrees also implies that the number of open pairs that can be counted by more than one $\ell \in Z_A$ is at most $(|Z_A| L^4)^2 = \widetilde{O}(1)$ by (54), which is negligible. Thus

$$\begin{aligned} N_{iAi} &\ge (1 - n^{-\varepsilon/5}) \sum_{\ell \in Z_A} \left(2ys_\ell + \hat{q} s_\ell (s_\ell - n^{1/2-\varepsilon/2})\right) - O(\hat{q} n^{1-\varepsilon/5}) \\ &= \sum_{\ell \in Z_A} \left(2ys_\ell + \hat{q} s_\ell^2\right) - O(\hat{q} n^{1-\varepsilon/5}). \end{aligned} \tag{59}$$

To estimate $S_{Ai} = \sum_{i=1}^m N_{iAi} q^{-1}$, it is convenient to use the bound (59) for all i, even selection steps (where $N_i = 0$); this is valid as the resulting correction is $\widetilde{O}(n^{-1/2})$, which is negligible. We write $S_{Ai} = S_{Ai1} + S_{Ai2} + \widetilde{O}(n^{1/2-\varepsilon/5})$ according to the contributions of the first and second terms in (59). Then

$$\begin{aligned} S_{Ai1} &= \sum_{i=1}^{m} \sum_{\ell \in Z_A} 2ys_\ell q^{-1} = \sum_{\ell \in Z_A} \sum_{j=1}^{d_\ell} \sum_{i=i_{\ell(j-1)}}^{i_{\ell j}-1} 4tn^{-3/2}(d_\ell + 1 - j) = \sum_{\ell \in Z_A} \sum_{j=1}^{d_\ell} \sum_{i=1}^{i_{\ell j}} 4in^{-3} \\ &= \sum_{\ell \in Z_A} \sum_{j=1}^{d_\ell} 2t_{\ell j}^2 - \sum_{\ell \in Z_A} \sum_{j=1}^{d_\ell} 2t_{\ell j} n^{-3/2} = \frac{T_A}{2} - \widetilde{O}(n^{-1}). \end{aligned}$$

Recalling (55), we note that

$$\begin{aligned} \frac{T_A}{2} = \sum_{\ell \in Z_A} \sum_{j=1}^{d_\ell} 2t_{\ell j}^2 &< |Z_A| \sum_{j=1}^{d+n^{1/2-\varepsilon/3}} 2\left(jn^{-1/2}/2 + n^{-\varepsilon/3}\right)^2 \\ &< |Z_A| \sum_{j=1}^{d} j^2 (2n)^{-1} + \widetilde{O}(n^{1/2-\varepsilon/3}). \end{aligned} \tag{60}$$

We also have

$$S_{Ai2} = \sum_{i=1}^{m} \sum_{\ell \in Z_A} \hat{q} s_\ell^2 q^{-1} = \sum_{\ell \in Z_A} \sum_{j=1}^{d_\ell} \sum_{i=i_{\ell(j-1)}}^{i_{\ell j}} n^{-2}(d_\ell - j)^2,$$

which is minimized when each d_ℓ is as small as possible, and then each $i_{\ell j}$ occurs as early as possible, so $S_{Ai2} \geq |Z_A| \sum_{j=1}^{d} (2n)^{-1} j^2 - \widetilde{O}(n^{1/2-\epsilon/3}) \geq T_A/2 - \widetilde{O}(n^{1/2-\epsilon/3})$ by (60). The lemma follows. ∎

Lemma 7.12. $T_B - S_B \leq O(L^{-2} n^{1/2})$.

Proof. Similarly to the proof of Lemma 7.11, we start by giving a lower bound on N_{iB} for any i that is not a selection step. For $\ell \in Z_B$ let $S_\ell = S_\ell(i)$ be the set of $v_{\ell j}$ with $j \in [h_\ell]$ such that $x_\ell v_{\ell j}$ is still open. We write $s_\ell = |S_\ell|$. Each $v_{\ell j}$ in S_ℓ contributes $2Y_{v_{\ell j} x_\ell} = (1 \pm \delta_Y)2y$ to N_{iBi} and $2Y_{x_\ell v_{\ell j}} = (1 \pm \delta_Y)2y$ to N_{iBo}; however, we need to account for open pairs that may be counted by more than one pair $x_\ell v_{\ell j}$.

We claim that there is no overcounting for inner pairs. To see this, note that if $v_{\ell j} v_{\ell' j'}$ is counted for $x_\ell v_{\ell j}$ and for $x_{\ell'} v_{\ell' j'}$ then $x_\ell v_{\ell' j'}$ and $x_{\ell'} v_{\ell j}$ are both edges, but this cannot occur by the hole construction procedure. Furthermore, there is no overcounting between N_{iBi} and N_{iBo}, as inner pairs intersect K but outer pairs do not (as K is independent).

Thus the claim holds, and it remains to consider overcounting for outer pairs. This may occur for $x_\ell v_{\ell j}$ and $x_\ell v_{\ell j'}$ with $\ell \in Z_B$ and $j, j' \in S_\ell$. The number of such overcounted pairs is at most $W_{x_\ell S_\ell}$, which we will estimate by Lemma 7.10. To see that this lemma applies, we note that $s_\ell \leq h_\ell < L^{-\beta}\sqrt{n}$ as holes H_ℓ with $\ell \in Z_B$ are not large. We also note that (x_ℓ, S_ℓ) is neighborly, as $S_\ell \subseteq N(x_\ell)$ and all pairs $x_\ell y$ with $y \in S_\ell$ are open, so $G(i)$ has no edges within $H_\ell \cup \{x_\ell\}$ and for any vertex $a \neq x_\ell$ at most L^4 edges ab with $b \in H_\ell$ and at most $2x$ open pairs ab with $b \in H_\ell$. If $s_\ell \geq L^\alpha$ then Lemma 7.10 gives $W_{x_\ell S_\ell} < L^{-\alpha} s_\ell \hat{q} \sqrt{n}$. Summing over $\ell \in Z_B$, using $|Z_B| \leq z \leq 4L^{15-2\alpha}k$ from (54) and $\sum_{\ell \in Z_B} s_\ell \leq k$ we obtain

$$N_{iBo} \geq (1-\delta_Y)2y \sum_{\ell \in Z_B} (s_\ell - L^\alpha) - \sum_{\ell \in Z_B} L^{-\alpha} s_\ell \hat{q}\sqrt{n} \geq 2y \sum_{\ell \in Z_B} s_\ell - L^{17-\alpha} k \hat{q} \sqrt{n}.$$

Including N_{iBi}, we deduce

$$N_{iB} \geq (1-\delta_Y)4y \sum_{\ell \in Z_B} s_\ell - L^{17-\alpha} k \hat{q}\sqrt{n} = 4y \sum_{\ell \in Z_B} s_\ell - O(L^{-3}\hat{q}n), \tag{61}$$

as α is large. As $S_B = \sum_{i=1}^{m} N_{iB} q^{-1}$, we have

$$S_B + O(L^{-2}n^{1/2}) = \sum_{i=1}^{m} \sum_{\ell \in Z_B} 4y s_\ell q^{-1} \geq \sum_{\ell \in Z_B} \sum_{j=1}^{h_\ell} \sum_{i=i_{\ell(j-1)}}^{i_{\ell j}} 8tn^{-3/2} s_\ell$$

$$= \sum_{\ell \in Z_B} \sum_{j=1}^{h_\ell} \sum_{i=1}^{i_{\ell j}} 8in^{-3} = T_B - \widetilde{O}(n^{-1}).$$

Similarly to Lemma 7.11, there is a negligible correction due to using the bound (61) at selection steps. The lemma follows. ∎

Lemmas 7.11 and 7.12 reduce (58) to

$$-\log \mathbb{P}(\mathcal{E}_K) \geq S_{Ao} + S_K - S_O + \log \frac{n^2}{2}\left(\sum_{\ell \in Z_A} d_\ell + |B|\right) - O(n^{1/2}L^{-2}), \tag{62}$$

We continue to estimate the terms in (62) over the next three lemmas.

Lemma 7.13. $S_{Ao} \geq 2|Z_A|m/n - \widetilde{O}(n^{1/2-\varepsilon/5})$.

Proof. If i is not a selection step then by control of open degrees

$$N_{iAo} \geq 2\sum_{\ell \in Z_A}(X_{x_\ell} - d_\ell - k - z) \geq 2|Z_A|\hat{q}n - \tilde{O}(\hat{q}n^{1-\varepsilon/5}).$$

As $S_{Ao} = \sum_{i=1}^{m} N_{iAo}q^{-1}$ the lemma follows. ■

For N_{iK} we will require more precise estimates for the contribution from open pairs with one vertex in the smaller holes, and so we need to account for this contribution further into the process. Accordingly, we define the following thresholds for hole sizes. We write

$$h^* = h^*(i) = \min\{n^{2/5}, L^{-50}\hat{q}\sqrt{n}\},$$

and let $\ell^* = \ell^*(i) \in [z+1]$ be such that $h_\ell \geq h^*$ for $1 \leq \ell < \ell^*$ and $h_\ell < h^*$ for $\ell^* \leq \ell \leq z$.

We also let z' be such that $h_\ell \geq n^{2/5}$ for $\ell \leq z'$ and $h_\ell < n^{2/5}$ otherwise. Thus $\ell^* \geq z'$ and equality holds at the beginning of the process. By Lemma 7.8(ii) we have

$$z' < 4L^{15}k/n^{2/5} = \widetilde{O}(n^{1/10}). \tag{63}$$

We let $J_1 = J_1(i) = \cup_{\ell \leq \ell^*} H_\ell$ and $J_2 = J_2(i) = \cup_{\ell > \ell^*} H_\ell$; thus (J_1, J_2) is a partition of $A \cup B$.

We write $N_{iK} \geq \sum_{\ell=1}^{z'} N_{iKH_\ell} + N_{iKJ_2} + N_{iKC}$, where each N_{iKX} counts ordered open pairs counted by N_{iK} with first vertex in X.

Lemma 7.14. *If i is not a selection step then $N_{iK} \geq \sum_{\ell=1}^{z'} N_{iKH_\ell} + N_{iKJ_2} + N_{iKC}$, where*

(i) $N_{iKX} \geq \hat{q}k|X|$ for $X \in \{J_2, C\}$, and
(ii) $N_{iKH_\ell} > (1 - L^{-5})\hat{q}h_\ell k/2$ if $\ell \leq z'$ and $\hat{q} \geq n^{-1/6}$.

Proof. We write $N_{iKJ_2} = Q'_{J_2} + Q_{J_1J_2} + Q_{J_2C}$, where Q'_{J_2} counts ordered open pairs in J_2 that are not contained within any hole. To estimate Q_{J_2} we note that any vertex has degree at most h^* in J_2 by the hole construction procedure. By Lemma 7.3(i) whp $Q_{J_2} = (1 \pm L^{-5})\hat{q}|J_2|^2$ if $\hat{q}|J_2| \geq L^{20}h^*$, so we can write $Q_{J_2} \geq (1 - L^{-5})\hat{q}|J_2|(|J_2| - L^{-30}\sqrt{n})$. Then

$$Q'_{J_2} \geq Q_{J_2} - h^*|J_2| \geq (1 - L^{-5})\hat{q}|J_2|(|J_2| - 2L^{-30}\sqrt{n}).$$

For the second term we consider $Q_{J_1J_2} \geq Q_{J_1J'_2}$ where $J'_2 = J_2 \setminus N(T)$ and T is the set of vertices with at least $L^{20}h^*$ neighbors in J_1. We can assume $|T| < 4L^{-5}|J_1|/h^* < 6L^{-4}\sqrt{n}/h^*$ by Lemma 7.8, so $|N(T) \cap J_2| < 6L^{-4}\sqrt{n}$. We apply Lemma 7.4 with $R = J_1$ and $S = J'_2 = J_2 \setminus N(T)$, noting that if a vertex

x has a neighbor in S then $x \notin T$, so x has at most $L^{20}h^*$ neighbors in J_1. If $\hat{q}\min\{|J_1|, |J_2'|\} \geq L^{40}h^*$ this gives whp $Q_{J_1J_2'} = (1 \pm L^{-5})\hat{q}|J_1||J_2'|$, so as $h^* \leq L^{-50}\hat{q}\sqrt{n}$ we have

$$Q_{J_1J_2'} \geq (1 - L^{-5})\hat{q}(|J_1| - L^{-4}\sqrt{n})(|J_2| - 7L^{-4}\sqrt{n}).$$

We can apply the same argument to estimate $Q_{J_2C} \geq Q_{J_2C'}$ where $C' = C \setminus N(T')$ and T' is the set of vertices with at least $L^{20+2\alpha}$ neighbors in J_2. We can assume $|T'| < 4L^{-5-2\alpha}|J_2| < 6L^{-4-2\alpha}\sqrt{n}$ by Lemma 7.8, so $|N(T') \cap C| < 6L^{-4}\sqrt{n}$ as any vertex has at most $L^{2\alpha}$ neighbors in C. Applying Lemma 7.4 with $R = J_2$ and $S = C' = C \setminus N(T')$, whp $Q_{J_2C'} = (1 \pm L^{-5})\hat{q}|J_2||C'|$ if $\hat{q}\min\{|J_2|, |C'|\} \geq L^{40+2\alpha}$, so we can write $Q_{J_2C'} \geq (1 - L^{-5})\hat{q}(|J_2| - L^{-4}\sqrt{n})(|C| - 7L^{-4}\sqrt{n})$. In total, as $|J_1| + |J_2| + |C| = k$ and $\hat{q}kL^{-4}\sqrt{n} = O(L^{-3}\hat{q}n)$ we obtain

$$N_{iKJ_2} \geq Q'_{J_2} + Q_{J_1J_2'} + Q_{J_2C'} \geq \hat{q}k|J_2| - O(L^{-3}\hat{q}n).$$

We now turn to $N_{iKC} \geq Q_C + Q_{A\cup B,C}$. As any vertex has at most $L^{2\alpha}$ neighbors in C, by Lemma 7.3(i) whp $Q_C \geq (1 - L^{-5})\hat{q}|C|(|C| - L^{-4}\sqrt{n})$. Next we estimate $Q_{A\cup B,C} \geq Q_{A\cup B,C''}$ where $C'' = C \setminus N(T'')$ and T'' is the set of vertices with at least $L^{20+2\alpha}$ neighbors in $A \cup B$. As in the argument for $Q_{J_2C'}$, we have $Q_{A\cup B,C''} = (1 \pm L^{-5})\hat{q}|A \cup B||C''|$ if $\hat{q}\min\{|A \cup B|, |C''|\} \geq L^{40+2\alpha}$, so

$$N_{iKC} \geq Q_{A\cup B,C''} + Q_C \geq \hat{q}k|C| - O(L^{-3}\hat{q}n).$$

This completes the proof of (i). For (ii) we need to estimate N_{iKH_ℓ} when $\hat{q} \geq n^{-1/6}$ and $\ell \leq z'$ (i.e., $h_\ell \geq n^{2/5}$). We write $X = \{\ell' \neq \ell : h_{\ell'} \geq 2n^{1/4}\}$ and $N_{iKH_\ell} = \sum_{\ell'\in X} Q_{H_\ell H_{\ell'}} + Q_{H_\ell K'}$, where $K' = K \setminus \bigcup_{\ell'\in X} H_{\ell'}$. We first apply Lemma 7.4 for each $\ell' \in X$ to $R = H_\ell \setminus N(x_{\ell'})$ and $S = H_{\ell'} \setminus N(x_\ell)$. This is valid by the codegree bound, which implies $|R|, |S| \geq n^{1/4}$ and also that any vertex with a neighbor in one of R or S has at most $L^4 < L^{-20}\hat{q}(2n^{1/4})$ neighbors in the other, as $\hat{q} \geq n^{-1/6}$. Thus $Q_{H_\ell H_{\ell'}} = (1 \pm L^{-5})\hat{q}h_\ell h_{\ell'}$.

Now we estimate $Q_{H_\ell K'} \geq Q_{RK'}$ where $R = H_\ell \setminus N(U)$ and U is the set of $x \neq x_\ell$ with at least $n^{1/5}$ neighbors in K. We have $|U| < 8L^{16}n^{3/10}$ by Lemma 7.8(ii), so $|N(U) \cap H_\ell| < L^{21}n^{3/10}$ by the codegree bound. Next we note that if a vertex x has a neighbor in K' then $x \neq x_\ell$ by the hole construction procedure, so by the codegree bound x has at most $L^4 < n^{1/5}$ neighbors in $R \subseteq H_\ell$. On the other hand, if x has a neighbor in R then $x \notin U$, so x has at most $n^{1/5}$ neighbors in $K' \subseteq K$. By Lemma 7.4, as $\hat{q} \geq n^{-1/6}$ we have $Q_{H_\ell K'} \geq (1 - L^{-5})\hat{q}(h_\ell - L^{21}n^{3/10})(|K'| - n^{2/5})$. As $h_\ell \leq d_\ell < (1 - \varepsilon)k/2$ we have $k - h_\ell - n^{2/5} > k/2$, and (ii) follows. ■

Lemma 7.15. *The overcount at step i is $N_{iO} = O(L^{-3}\hat{q}n)$, so $S_O = \sum_i N_{iO}q^{-1} = O(L^{-2}n^{-1/2})$.*

Proof. Let us consider the possible pairwise overcounting between N_{iAo}, N_{iAi}, N_{iBo}, N_{iBi}, and N_{iK}. Note that by excluding $y \in K \cup \{x_1, \dots, x_z\}$ in the definition of N_{iAo} we ensured that it does not intersect any of the other collections. There is no overcounting between N_{iBo} and $N_{iBi} + N_{iK}$, as pairs counted by the former do not intersect K while pairs counted by the latter do intersect K. There is no overcounting between N_{iBi} and N_{iAi}, as the hole construction procedure ensures that no vertex in a hole H_ℓ with $\ell \in Z_B$ is also a neighbor of some vertex $x_{\ell'}$ such that $\ell' \in Z_A$. It remains to consider the following possible overcounting of pairs:

(i) N_{iAi} with N_{iK},
(ii) N_{iAi} with N_{iBo},
(iii) N_{iK} with N_{iBi}.

For (i), we note that a pair counted by N_{iAi} and N_{iK} has the form yy' where y, y' are both neighbors of some x_ℓ with $\ell \in Z_A$, and are both in K but not in the same hole. By the hole construction procedure at least one is also adjacent to some other $x_{\ell'}$, so by the codegree bound there are $\widetilde{O}(k) = \widetilde{O}(n^{1/2})$ such pairs. For (ii), the overcount between N_{iAi} and N_{iBo} is determined by naming a vertex $b \in B$, a vertex x_ℓ such that $\ell \in Z_A$, and a vertex c that is in the (final) common neighborhood of x_ℓ and b; this overcount is at most $k|Z_A|L^4 = \widetilde{O}(n^{1/2})$.

To bound the most significant overcount (iii), namely that between N_{iK} and N_{iBi}, we introduce the following definition. We say that a hole H_ℓ with $\ell \in Z_B$ is *black* if x_ℓ has more than $L^{30}h_\ell$ neighbors in K. We let XH be the set of such x_ℓ and BH be the set of vertices that belong to black holes. By Lemma 7.6(ii) applied to $S = K \cup XH$ we have $L^{15}|S| > \eta_S \geq \sum_{x_\ell \in XH} L^{30}h_\ell = L^{30}|BH|$, so $|BH| \leq L^{-14}k$. The contribution to N_{iBi} of pairs that would close pairs $x_\ell v_{\ell j}$ with $v_{\ell j} \in BH$ is at most $3y|BH| \leq 3L^{-14}yk \leq 3L^{-13}\hat{q}kn^{1/2}$.

Now consider overcounted pairs that would close pairs that are not incident to black holes. Such a pair has the form $v_{\ell j}v_{\ell' j'}$ where $x_\ell v_{\ell' j'}$ is an edge, so $\ell' < \ell$ by the hole construction procedure. It suffices to show for any fixed x_ℓ that at most $L^{-10}h_\ell\hat{q}\sqrt{n}$ such pairs are also counted by N_{iBi}. Suppose first that $h_\ell \geq n^{2/5}$, so that $\ell' < \ell \leq z' = \widetilde{O}(n^{1/10})$ by (63). By the codegree bound there are at most $z' \cdot L^4 < n^{1/5}$ such edges $x_\ell v_{\ell' j'}$, which are only counted in our estimate for N_{iK} in Lemma 7.14 while $\hat{q} > n^{-1/6}$, so the overcount for such a hole is at most $h_\ell n^{1/5} < h_\ell \hat{q} n^{2/5}$. Now suppose $h_\ell < n^{2/5}$. We recall that open pairs between H_ℓ and $H_{\ell'}$ are only counted in our estimate for N_{iK} in Lemma 7.14 if $H_\ell \subseteq J_2$, that is, if $h_\ell < h^* \leq L^{-50}\hat{q}\sqrt{n}$. Since H_ℓ is not black, the number of choices for $v_{\ell' j'}$ is at most $L^{30}h_\ell < L^{-10}\hat{q}\sqrt{n}$, so such pairs contribute at most $L^{-10}h_\ell\hat{q}\sqrt{n}$. Summing over all holes gives the desired bound. ∎

We are now ready for the union bound calculation that bounds $\mathbb{P}(\mathcal{E})$. Recall that we have fixed the initial data that defines the event $\mathcal{E}$; that is, we have specified z, the vertices $x_1, x_2, \dots, x_z$, the hole sizes $h_1, \dots, h_z$ and the degrees d_ℓ of vertices x_ℓ for $\ell \in Z_A$. We then partition $\mathcal{E}$ into events $\mathcal{E}_K$ as analyzed above, defined by choices of neighborhoods of x_ℓ for $\ell \in Z_A$, vertices in $A \cup B$ (which are named by specifying the vertices in holes), selection steps $i_{\ell j}$, and vertices in C. The number of choices for the data that defines $\mathcal{E}_K$ is at most

$$\left(\prod_{\ell \in Z_A} \binom{n}{d_\ell}\binom{d_\ell}{h_\ell} m^{d_\ell}\right)\left(\prod_{\ell \in Z_B} \binom{n}{h_\ell} m^{h_\ell}\right)\binom{n}{|C|}.$$

To estimate $\mathbb{P}(\mathcal{E})$ we apply (62) to each such choice of $\mathcal{E}_K$, substituting $S_O = O(L^{-2}n^{-1/2})$ from Lemma 7.15 and $S_{Ao} \geq 2|Z_A|m/n - \widetilde{O}(n^{1/2-\epsilon/5})$ from Lemma 7.13 (the latter accounts for the $\exp(-2m/n)$ term in the calculation below). Recalling $|B| = \sum_{\ell \in Z_B} h_\ell$ and $d_\ell = 2m/n \pm n^{1/2-\epsilon/3}$, using $\binom{d_\ell}{h_\ell} < \exp\{O(\log\log n)h_\ell\}$ for $\ell \in Z_A$ and $\log\binom{n}{|C|} < |C|\log n/2 + O(\log\log n)k$, we have

$$\begin{aligned}\mathbb{P}(\mathcal{E}) &\leq \prod_{\ell \in Z_A}\left[\left(\frac{ne}{d_\ell}\cdot\frac{2m}{n^2}\right)^{d_\ell}\exp\{-2m/n + O(\log\log n)h_\ell\}\right] \\ &\quad \cdot \left(\prod_{\ell \in Z_B}\left(\frac{ne}{h_\ell}\cdot\frac{2m}{n^2}\right)^{h_\ell}\right)\binom{n}{|C|} e^{-S_K + O(L^{-2}n^{1/2})} \\ &\leq \exp\left\{\sum_{\ell \in Z_B} h_\ell \log(\sqrt{n}/h_\ell) + |C|\log n/2 - S_K + O(\log\log n)k\right\}. \qquad (64)\end{aligned}$$

It remains to show that S_K is sufficiently large to make the above probability expression small enough for the union bound over the initial data defining $\mathcal{E}$. We first note for Z_A that the counting terms $\left(\frac{ne}{d_\ell}\cdot\frac{2m}{n^2}\right)^{d_\ell} = (e \pm O(n^{-\epsilon/5}))^{d_\ell}$ are canceled to highest order by the probability term $\exp(-2m/n)$ from Lemma 7.13, so we require S_K to dominate the counting terms from the choice of B and C. For B we consider the contributions from each hole as follows.

The contributions corresponding to the hole H_ℓ depends on time when the hole moves out of the set J_1 defined before Lemma 7.14. If $h_\ell \geq n^{2/5}$ (i.e., $\ell \leq z'$) we obtain a term $\hat{q}kh_\ell/2$ in the bound from Lemma 7.14 while $\hat{q} > n^{-1/6}$, that is, up to time $\frac{1}{2}\sqrt{\frac{1}{6}\log n}$. If $h_\ell < n^{2/5}$ we obtain a term $\hat{q}kh_\ell$ from Lemma 7.14 while $\hat{q} > L^{50}h_\ell/\sqrt{n}$, that is, up to time $t_\ell = \frac{1}{2}\sqrt{\log\frac{\sqrt{n}}{L^{50}h_\ell}}$ if this time is less than $i_{\max}$ and up to time $i_{\max}$ otherwise. Let z'' be the smallest index ℓ such that $t_\ell < t_{\max}$ (this corresponds to a threshold for hole sizes that is about $L^{-50}n^{\epsilon}$). As $S_K = \sum_i N_{iK}q^{-1}$, we have

$$S_K \geq |C|\frac{mk}{n^2} + \left(\sum_{\ell=1}^{z'}\frac{h_\ell}{2}\right)n^{3/2}\cdot\frac{1}{2}\sqrt{\frac{1}{6}\log n}\cdot\frac{k}{n^2} + \left(\sum_{\ell=z'+1}^{z''}h_\ell\cdot n^{3/2}\cdot\frac{1}{2}\sqrt{\log\frac{\sqrt{n}}{L^{50}h_\ell}}\right)\cdot\frac{k}{n^2}$$
$$+\left(\sum_{\ell=z''}^{z}h_\ell\right)\frac{mk}{n^2} - O(L^{-2}n^{1/2}). \tag{65}$$

Finally we substitute (65) in (64), grouping terms according to the contribution of each h_ℓ, organized into the same summation ranges as in (65). For each hole H_ℓ with $\ell \in Z_B$ included in one of these ranges we have a counting term $\log\left(\frac{ne}{h_\ell}\cdot\frac{2m}{n^2}\right)^{h_\ell} = h_\ell(\log\frac{\sqrt{n}}{h_\ell} + O(\log\log n))$ from (64) which we pair with a probability term from (65). In the calculations below we also use (i) $\log\frac{\sqrt{n}}{h_\ell} \leq \frac{1}{10}\log n$ for $\ell \leq z'$, (ii) $\sqrt{(\frac{1}{2}\log n)\cdot\log(\frac{\sqrt{n}}{L^{50}h_\ell})} > \log\frac{\sqrt{n}}{h_\ell}$ for $z' < \ell \leq z''$, and (iii) $mk/n^2 > (1+\epsilon)\frac{1}{2}\log n$, which holds (for small ϵ) as $k = (1+3\epsilon)\sqrt{2n\log n}$ and $m = \sqrt{(1/2-\epsilon)\log n}\cdot n^{3/2}/2$. We have

$$\begin{aligned}\log\mathbb{P}(\mathcal{E}) &\leq -\sum_{\ell=1}^{z'}h_\ell\left(\frac{1}{4\sqrt{3}}-\frac{1}{10}\right)\log n - \sum_{\ell=z'+1}^{z''}3\epsilon h_\ell\log\frac{\sqrt{n}}{h_\ell}\\ &\quad -\sum_{\ell=z''+1}^{z}\epsilon h_\ell\frac{1}{2}\log n - \epsilon|C|\frac{1}{2}\log n + O(\log\log n)k\\ &\leq -\frac{\epsilon}{4}k\log n + O(\log\log n)k.\end{aligned}$$

As the number of choices of the initial data that defines $\mathcal{E}$ is $O(n^{2z})$ and $z \leq 4kL^{15-2\alpha}$, where α is large, the probability that any such event $\mathcal{E}$ holds is $o(1)$, which completes the proof. □

7.3 | Proof of the upper bound in Theorem 1.1.

This proof is very similar to that of Theorem 1.2, but much simpler. The lower bound on degrees in G follows from Theorem 2.13, so it remains to show the upper bound. We take a union bound over every vertex x, potential neighborhood A, and set C such that

$$|C| = 5\epsilon\sqrt{n\log n}$$

of the event that

1. A is the neighborhood of x in $G(i_{\max})$,
2. $A \cup C$ spans no edge in $G(i_{\max})$, and
3. vx is open in $G(i_{\max})$ for all $v \in C$.

We view C as vertices that might be added to the neighborhood of v between time $t_{\max}$ and the end of the process. We show that whp there is no triple (x, A, C) with these properties.

We fix x, A, C, write $A = \{v_1, \dots, v_{d'}\}$ for some d' and specify the appearance time i_j for every edge xv_j, where $j < j'$ implies $i_j < i_{j'}$. As in (55), $I < i_{\max}$ implies

$$i_j = jn/2 \pm n^{3/2-\epsilon/3} \quad \text{and} \quad d' = d \pm n^{1/2-\epsilon/3},$$

where we recall $d = 2t_{\max}\sqrt{n} = 2m/n = \sqrt{(1/2-\epsilon)n\log n}$.

Let $\mathcal{F}$ be the event that $A \cup C$ is an independent set in $G(i_{\max})$, all pairs joining x and C are open in $G(i_{\max})$, and all the specified edges appear at the specified steps of the process. To estimate the probability of the event $\mathcal{F}$, for each step i we need to estimate the probability that the selected edge is compatible with this event, conditional on the history of the process. We say i is a *selection step* if i is one of i_j for $j \in [d']$; then the selected edge is specified by $\mathcal{F}$, so the required probability is simply $2/Q = (1 \pm 2\delta_Q)2q^{-1}$. For other i, the required probability is $1 - N_i/Q$, where N_i is the number of ordered open pairs that cannot be selected at step i when $\mathcal{F}$ occurs. If $i = i_j$ is a selection step write $N_i = 0$. Then we estimate

$$\mathbb{P}(\mathcal{F}) \le \prod_{j=1}^{d'} (1 \pm 2\delta_Q)2q(t_j)^{-1} \cdot \prod_{i=1}^{m} (1 - N_i/Q),$$

where $t_j = i_j/n^{3/2}$. We write $N_i = N_{iA} + N_{iC}$, where N_{iA} counts the ordered open pairs within A and N_{iC} counts those in $A \cup C$ with at least one vertex in C. We have

$$-\log \mathbb{P}(\mathcal{F}) \ge S_A - T_A + S_C + d' \log \frac{n^2}{2} - O(n^{1/2}), \tag{66}$$

where $S_\mu = \sum_{i=1}^m N_{i\mu} q^{-1}$ for $\mu \in \{A, C\}$ and $T_A = \sum_{j=1}^{d'} 4t_j^2$.

Following the argument in the previous section for estimating $S_{Ai} - T_A$, we have the following estimate on $S_A - T_A$. We include a proof here in the interest of presenting a complete and self-contained proof of the upper bound in Theorem 1.1.

Lemma 7.16. $S_A - T_A = \widetilde{O}(n^{1/2-\epsilon/3})$.

Proof. We first estimate N_{iA} when i is not a selection step. Let $S = S(i) = \{v_j \in A : i_j > i\}$ and $s = |S|$; thus $S(i)$ is the set of vertices y in A such that yx is open and is yet to be joined to x. The number of ordered open pairs $v_j v_{j'}$ with $j > i, j' \le i$ is $\sum_{v \in s} 2Y_{vx} = (1 \pm \delta_Y)2ys$. Next note that any vertex has at most L^4 neighbors in S, by the bound on codegrees in $G(i_{\max})$, which applies as $I > i_{\max}$. Then by Lemma 7.3(i) whp $Q_s = (1 \pm n^{-\epsilon/5})\hat{q}s^2$ if $s > n^{1/4}$ and $\hat{q}s \ge n^{2\epsilon/5}L^{14}$. Since $\hat{q} \ge n^{-1/2+\epsilon}$ this holds for $s > n^{1/2-\epsilon/2}$, so we can write $Q_s \ge (1 - n^{-\epsilon/5})\hat{q}s(s - n^{1/2-\epsilon/2})$. Thus

$$N_{iA} \ge (1 - n^{-\epsilon/5})\left(2ys + \hat{q}s(s - n^{1/2-\epsilon/2})\right) = 2ys + \hat{q}s^2 - \widetilde{O}(\hat{q}n^{1-\epsilon/5}).$$

Now we estimate $S_A = \sum_{i=1}^m N_{iA}q^{-1}$, which we write as $S_A = S_{A1} + S_{A2} + \widetilde{O}(n^{1/2-\epsilon/5})$ according to the contributions of the first and second terms in the estimate for N_{iA}, and as before we incur a negligible error by using this bound even at selection steps. Thus

$$
\begin{aligned}
S_{A1} &= \sum_{i=1}^{m} 2ysq^{-1} = \sum_{j=1}^{d'} \sum_{i=i_{j-1}}^{i_j-1} 4tn^{-3/2}(d'+1-j) = \sum_{j=1}^{d'} \sum_{i=1}^{i_j} 4in^{-3} \\
&= \sum_{j=1}^{d'} 2t_j^2 - \sum_{j=1}^{d'} 2t_j n^{-3/2} = \frac{T_A}{2} - \widetilde{O}(n^{-1}), \text{ and} \\
S_{A2} &= \sum_{i=1}^{m} \hat{q}s^2q^{-1} = \sum_{j=1}^{d'} \sum_{i=i_{j-1}}^{i_j-1} n^{-2}(d'+1-j)^2 \\
&\geq \sum_{j=1}^{d} (2n)^{-1}j^2 - \widetilde{O}(n^{1/2-\varepsilon/3}) \geq T_A/2 - \widetilde{O}(n^{1/2-\varepsilon/3}).
\end{aligned}
$$

The lemma follows. ■

To estimate S_C we require the crucial claim that

$$|N(u) \cap C| < L^2 n^{\varepsilon} \tag{67}$$

for any vertex u. Indeed, if this failed for some u then at time $t_{\max}$ we have $Y_{xu} > 2y$. However, this would contradict our estimate on Y-variables. (We can assume xu is a non-edge as x is open to C, and we recall that we track Y_{xu} whether xu is open or closed.) Thus the claim holds.

While $\hat{q}|C| > L^{15}n^{\varepsilon}$, which as $|C| = \widetilde{\Theta}(\sqrt{n})$ holds up to time $(1+o(1))t_{\max}$, we can apply Lemmas 7.3(i) and 7.4 to obtain $Q_C \geq (1-L^{-1})\hat{q}|C|^2$ and $Q_{AC} \geq (1-L^{-1})\hat{q}|A||C|$. When i is not a selection step this gives $N_{iC} = 2Q_{AC} + Q_C \geq (1-L^{-1})\hat{q}(2|A||C| + |C|^2)$, so

$$S_C = \sum_{i=1}^{m} N_{iC}q^{-1} > (1-o(1))(2|A|+|C|)|C|m/n^2 = (1-o(1))\left(1 - 2\varepsilon + 5\varepsilon\sqrt{\frac{1}{2} - \varepsilon}\right)|C|\tfrac{1}{2}\log n.$$

Now we substitute Lemma 7.16 in (66), and take the union over all possible choices of the data that specifies an event $\mathcal{F}$, namely the choices of x, d', A, C and the collection of times at which the edges joining x to A appear. Thus we bound the probability p_0 that any triple (x, A, C) as above exists by

$$p_0 < n\sum_{d'} \binom{n}{d'}\binom{n}{|C|} m^{d'} \left(\tfrac{2}{n^2}\right)^{d'} \exp\left\{-(1-o(1))\left(1 - 2\varepsilon + 5\varepsilon\sqrt{\frac{1}{2} - \varepsilon}\right)|C|\tfrac{1}{2}\log n + O(n^{1/2})\right\}.$$

Here we note that the counting term $\binom{n}{d'}m^{d'}\left(\frac{2}{n^2}\right)^{d'} = \exp[(1+o(1))d]$ is of lower order than the main counting term $\binom{n}{|C|} = \exp[(1+o(1))|C|\frac{1}{2}\log n]$, and this is more than compensated for by the probability term: assuming $\varepsilon < 1/4$, we obtain

$$p_0 < n\sum_{d'} \exp\left\{-\varepsilon|C|\tfrac{1}{5}\log n\right\}.$$

Thus the required bound on degrees holds with high probability. □

8 | CONCLUDING REMARKS

We have determined $R(3, t)$ to within a factor of $4 + o(1)$, so we should perhaps hazard a guess for its asymptotics: we are tempted to believe the construction rather than the bound, that is, that $R(3, t) \sim t^2/4 \log t$. We only proved an upper bound on the independence number of the graph G produced by the triangle-free process, so in principle it might give a better lower bound on $R(3, t)$. However, we believe that this is not the case: we conjecture that the bound on the independence number in Theorem 1.2 is asymptotically best possible.

Another natural direction for future research is to provide an asymptotically optimal analysis in greater generality for the H-free process. No doubt the technical challenges will be formidable, given the difficulties that arise in the case of triangles. But on an optimistic note, it is encouraging that one can build on two different proofs of this case.

ACKNOWLEDGMENTS

We thank the anonymous referees for many helpful comments and suggestions.

REFERENCES

1. M. Ajtai, J. Komlós, and E. Szemerédi, *A note on Ramsey numbers*, J. Combin. Theory Ser. A. **29** (1980), 354–360.
2. N. Alon, *Explicit ramsey graphs and orthonormal labelings*, Electron. J. Combin. **1** (1994), R12.
3. N. Alon, S. Ben-Shimon, and M. Krivelevich, *A note on regular Ramsey Graphs*, J. Graph Theory. **64** (2010), 244–249.
4. N. Alon and J. Spencer, *The probabilistic method*, 2nd ed., Wiley, New York, 2000.
5. T. Bohman, *The triangle-free process*, Adv. Math. **221** (2009), 1653–1677.
6. T. Bohman, A. Frieze, and E. Lubetzky, *A note on the random greedy triangle-packing algorithm*, J. Combin. **1** (2010), 477–488.
7. T. Bohman, A. Frieze, and E. Lubetzky, *Random triangle removal*, Adv. Math. **280** (2015), 379–438.
8. T. Bohman and P. Keevash, *The early evolution of the H-free process*, Invent. Math. **181** (2010), 291–336.
9. B. Bollobás and O. Riordan, *Random graphs and branching processes*. Handbook of Large-scale Random Networks, Bolyai Soc. Math. Stud., Springer, Berlin, Vol. 18, 2009, pp. 15–115.
10. B. Bollobás and O. Riordan, *Constrained graph processes*, Electron. J. Combin. **7** (2000), R18.
11. P. Erdős, *Graph theory and probability, II*, Can. J. Math. **13** (1961), 346–352.
12. P. Erdős, S. Suen, and P. Winkler, *On the size of a random maximal graph*, Random Structures Algorithms. **6** (1995), 309–318.
13. G. Fiz Pontiveros, S. Griffiths, and R. Morris, *The triangle-free process and* $R(3, k)$, Mem. Amer. Math. Soc. **263** (2020), 125.
14. D.A. Freedman, *On tail probabilities for martingales*, Ann. Probab. **3** (1975), 100–118.
15. A. Frieze and M. Karoński, *Introduction to random graphs*, Cambridge University Press, 2016.
16. S. Gerke and T. Makai, *No dense subgraphs appear in the triangle-free graph process*, Electron. J. Combin. **18** (2011), R168.
17. J.H. Kim, *The Ramsey number* $R(3, t)$ *has order of magnitude* $t^2/\log t$, Random Structures Algorithms. **7** (1995), 173–207.
18. J.H. Kim and V.H. Vu, *Concentration of multivariate polynomials and its applications*, Combinatorica. **20** (2000), 417–434.
19. D. Osthus and A. Taraz, *Random maximal H-free graphs*, Random Structures Algorithms. **18** (2001), 61–82.
20. M. Picollelli, *The diamond-free process*, Random Structures Algorithms. **45** (2014), 513–551.
21. M. Picollelli, *The final size of the* C_4*-free process*, Combin. Probab. Comput. **20** (2011), 939–955.
22. M. Picollelli, *The final size of the* C_ℓ*-free process*, SIAM Discrete Math. **28** (2014), 1276–1305.

23. A. Ruciński and N. Wormald, *Random graph processes with degree restrictions*, Combin. Probab. Comput. **1** (1992), 169–180.
24. J. Shearer, *A note on the independence number of triangle-free graphs*, Discrete Math. **46** (1983), 83–87.
25. J. Spencer, *Maximal trianglefree graphs and Ramsey* $R(3,k)$. Unpublished manuscript. Available at: http://www.cs.nyu.edu/spencer/papers/ramsey3k.pdf.
26. J. Spencer, *Asymptotic lower bounds for Ramsey functions*, Discrete Math. **20** (1997), 69–76.
27. J. Spencer, *Counting extensions*, J. Combin. Theory Ser. A. **55** (1990), 247–255.
28. L. Warnke, *Dense subgraphs in the H-free process*, Discrete Math. **333** (2011), 2703–2707.
29. L. Warnke, *The* C_ℓ*-free process*, Random Structures Algorithms. **44** (2014), 490–526.
30. L. Warnke, *When does the* K_4*-free process stop?*, Random Structures Algorithms. **44** (2014), 355–397.
31. G. Wolfovitz, *Lower bounds for the size of random maximal H-free graphs*, Electron. J. Combin. **16** (2009), R4.
32. G. Wolfovitz, *The K4-free process*, arXiv: 1008.4044.
33. G. Wolfovitz, *Triangle-free subgraphs in the triangle-free process*, Random Structures Algorithms. **39** (2011), 539–543.

How to cite this article: Bohman T, Keevash P. Dynamic concentration of the triangle-free process. *Random Struct Alg.* 2021;58:221–293. https://doi.org/10.1002/rsa.20973