The Ronald O. Perelman Center for Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
21-029

CONTAGION AND EQUILIBRIA IN DIVERSIFIED FINANCIAL NETWORKS

VICTOR AMELKIN
Amazon.com

SANTOSH VENKATESH
University of Pennsylvania

RAKESH VOHRA
University of Pennsylvania

November 30, 2021

https://ssrn.com/abstract=3974645

CONTAGION AND EQUILIBRIA IN DIVERSIFIED FINANCIAL NETWORKS

VICTOR AMELKIN
Amazon.com

SANTOSH VENKATESH
Department of Electrical & Systems Engineering, University of Pennsylvania

RAKESH VOHRA
Warren Center for Network & Data Sciences, Department of Economics, and Department of Electrical & Systems Engineering, University of Pennsylvania

Diversified cross-shareholding networks are thought to be more resilient to shocks, but diversification also increases the channels by which a shock can spread. To resolve these competing intuitions we introduce a stochastic model of a diversified cross-shareholding network in which a firm's valuation depends on its cash endowment and the shares it owns in other firms. We show that a concentration of measure phenomenon emerges: *almost all realized network instances drawn from any probability distribution in a wide class are resilient to contagion if endowments are sufficiently large*. Furthermore, the size of a shock needed to trigger widespread default *increases* with the exposure of firms to each other. Distributions in this class are characterized by the property that a firm's equity shares owned by others are weakly dependent yet lack "dominant" shareholders.

KEYWORDS: financial network, random network, contagion, systemic risk, equilibrium, dynamics, concentration of measure.

1. INTRODUCTION

Cross-shareholding networks, incorporating interdependencies between firms holding each other's shares, are pervasive (Shi, Townsend and Zhu, 2019). In such settings interconnectivity is traditionally held to encourage resilience via diversification but also has the countervailing possibility of acting as a conduit by which shocks to one firm can be transmitted to others.

As in Elliott, Golub and Jackson (2014) we suppose a firm's value depends on the shares it owns in other firms as well as its cash endowment (in the form of held primitive assets). When a firm's value falls below a failure threshold (insolvency), it discontinuously imposes losses on its counter-parties, e.g., as default costs. The counter-parties in turn may push other firms below the failure threshold. Thus, an initial *small* shock to a small number of firms has the potential to trigger widespread insolvency.

Prior work sought to characterize the structure of cross-shareholding networks that encourage the propagation and amplification of shocks. The absence of a closed form description of the equilibrium profile of firm valuations in cross-shareholding networks limited attention to specific networks, such as paths, cycles, stars, or cliques. An equilibrium here is a combination of firm values consistent with the firms' cash endowments and equity share cross-ownership. As a firm's value is a non-linear function of its cash endowment and the returns on its investment in other firms, there can be many equilibria. Prior work deals with this by selecting a "maximal" or "minimal" equilibrium. Therefore, one may justifiably wonder to what extent

Victor Amelkin: vctr@seas.upenn.edu
Santosh Venkatesh: venkates@seas.upenn.edu
Rakesh Vohra: rvohra@seas.upenn.edu

This research was made possible by grant 2017 PRE 301 from the Rockefeller Foundation.

the behavior observed in specific networks will extend to other networks and whether their extreme equilibria are informative about behavior at other equilibria.

Here we focus on diversified networks in which equity shares are widely distributed. On the one hand, they should be more resilient to shocks because of diversification, but it also increases the channels by which a shock can spread. Which of the two forces will dominate? Diversification is usually modeled by assuming that each firm holds *equal* amounts of shares in all other firms (called a regular clique). Instead of identifying *fully diversified networks* with the regular clique, we define them as arising from a probability distribution over equity shares. The properties of the distribution—rather than the features of any specific realized profile of share cross-ownership—encode what it means to be fully diversified. Specifically, the equity shares of a firm that are held by other firms are exchangeable random variables drawn from a suitably regular distribution where "regularity" is captured by a weak moment condition designed to obviate pathologies; we assume that the vectors of share allocations are independent across firms. We impose no conditions on cash endowments.

Exchangeability means that no one firm is distinguished from the others in some relevant way. It does *not* require firms to hold identical quantities of a given firm's shares. The moment condition ensures that the differences in shares held are not so dramatic as to result in ownership being concentrated in the hands of a few firms. Shares being widely distributed *and* the absence of concentration capture the essence of diversification. We call a network generated in this way a fully diversified network. No two realized fully diversified networks will be identical. However, as the size of the network increases, we show that the equilibria of *almost all* fully diversified networks are concentrated around the equilibria of the regular clique, making the latter "typical": even though, *ex post*, any realization of a fully diversified network will generally be very different from a regular clique, we can infer its equilibrium behavior from the equilibria of the regular clique.

To come to a clearer understanding of how a fully diversified network will respond to a shock to a firm's value, we abstract away from differences between firms by assuming each has the same endowment and exposure (fraction of a firm's equity owned by the firms in the network). As the equilibria of a fully diversified network are concentrated around the equilibria of the regular clique, two parameters suffice to determine equilibrium valuations: the endowment and the exposure.

By exploiting the fact that the equilibria of cross-shareholding networks lend themselves to an interpretation as the rest points of a natural dynamic we determine how *all* equilibria respond to a shock, not just the best (or the worst) equilibria that are the focus in, say, Elliott, Golub and Jackson (2014). The dynamic's state space is comprised of vectors of firm valuations divided into three zones: *optimal* (the number of solvent firms is "large"), *safe* (the number of solvent firms is "middling"), and *risky* (the number of solvent firms is "small"). Suppose a shock displaces firm values from their equilibrium state. The new profile of firm valuations becomes the initial state of the dynamic. If the initial state is located in the optimal zone, the dynamic converges to the "best" equilibrium profile (where every firm is solvent). If the initial state is located in the risky zone, the dynamic converges to the "worst" equilibrium profile (every firm is insolvent, which, essentially, corresponds to the economy's collapse). If the initial state is located in the safe zone, the dynamic converges to a profile of firm valuations where, often, initially solvent firms remain solvent, and initially insolvent firms remain insolvent.

We also describe how the "sizes" of these zones change as we vary both the endowment and exposure. With exposure fixed and a sufficiently large endowment, there is only one equilibrium, the "best" one in which all firms are solvent. As the endowment declines (holding exposure fixed), the "worst" equilibrium where all firms are insolvent emerges. This is not a sign of fragility as it requires a substantial shock to the "best" equilibrium to send the system into the "worst" equilibrium, that is, the basin of attraction of the "best" equilibrium is large. As the endowment declines further, equilibria where a majority of the firms are insolvent appear. However, the basin of attraction of the "best" equilibrium continues to remain large.

As the endowment continues to diminish, the basin of attraction of the "worst" equilibrium expands, absorbing equilibria with a small number of solvent firms, while that of the "best" equilibrium shrinks. As the endowment shrinks further, equilibria where a majority of firms are solvent appear. In this domain characteristic of smaller endowments, a very small shock to the "best" equilibrium will result in firms' valuations either returning to the best equilibrium or converging to a state that is "close" to the initial state. By "close" we mean that, while a firm's valuation will change, the firm will preserve its (in)solvency status, that is, if its valuation exceeded the failure threshold, it will not drop below it; and if it was initially below that threshold, it will not rise above it. The situation is perilous, however, as a medium to large shock, will propel the dynamics into the large basin of attraction of the worst equilibrium leading to collapse. The greatest peril emerges when the endowment becomes small: in that domain all equilibria vanish except the "worst" which now dominates the landscape.

If instead, we hold the endowment fixed, an increase in exposure increases the basin of attraction of the "best" equilibrium. Hence, a small shock to the profile of firm valuations in the best equilibrium will result in the system returning to its original state. A decrease in exposure, while shrinking the basins of attraction of the best and worst equilibria introduces many equilibria with a "middling" number of solvent firms. In this case, a small shock to the best equilibrium could cause the state of the system to fall out of its basin of attraction. However, the system will not terminate in the worst equilibrium but get "trapped" by one with a "middling" number of solvent firms.

We also investigate the impact of a shock to the endowment when there is a large number of firms. We find that increasing exposure makes firms slightly more sensitive to each other's endowment shocks. We also find that as the exposure increases, the magnitude of an endowment shock to any one needed to drive the system to the worst equilibrium decreases. However, the impact of exposure is of second order in the asymptotic regime: both these effects are strongly muted by the fact that there is a large number of firms. In particular the magnitude of the shock needed to force the system into the worst equilibrium scales with the number of firms. Hence, small shocks to the endowment are unlikely to generate widespread failure. Our story in a slogan: *if, for any given exposure, the endowment reaches a critical level then almost all diversified networks are resilient to shocks and contagion*.

A key technical contribution of independent interest is to the theory of random matrices: we show a concentration of measure phenomenon for random matrices that possess weak dependencies in the form of exchangeable column elements.

The remainder of this paper is organized as follows. Prior work is summarized in Section 2. Essential technical preliminaries are in Section 3. The model is defined in Section 4 and the characterization of its equilibria can be found in Section 5. Section 6 is devoted to the equilibria of the regular clique. Section 7 characterizes equilibria in random equity networks and fleshes out the concentration result which justifies the focus on the regular clique. Dynamical considerations are introduced in Section 8, leading naturally to Section 9 in which resilience to shocks is discussed. And to conclude, extensions to Erdös–Rényi topologies, stochastic block models, and sparse networks are sketched in Section 10. Proofs of the technical results are deferred to appendices (A–F) to keep the flow of presentation of the main ideas as unencumbered as possible.

2. PRIOR WORK & CONNECTIONS TO THE RESULTS OF THIS PAPER

The role that network structure plays in systemic risk has attracted considerable attention in the last twenty years. The relevant papers are those that consider a network of firms linked through debt, equity, or both. Their focus is to link the topology of the network (degree distribution, centrality) to the likelihood of contagion.

Acemoglu, Ozdaglar and Tahbaz-Salehi (2015), for example, consider a network where firms are connected via debt. An underlying network of debtor and creditor relationships is fixed (these are the links) and the magnitude of the clearing payments associated with these links are determined in equilibrium. They find that when shocks have small magnitude, higher network density (fraction of links relative to the maximum possible number of links) improves robustness of the system, with the regular clique being the best and the ring being the worst among regular networks. However, they find that for larger shocks, denser networks, such as the regular clique, facilitate the spread of defaults.

Our model of cross-shareholding networks follows Elliott, Golub and Jackson (2014) and focuses on equity rather than debt. Elliott, Golub and Jackson (2014) characterize the equilibria of the model—these are the firm valuations consistent with asset holding and equity share cross-holding by the firms. Like Acemoglu, Ozdaglar and Tahbaz-Salehi (2015), their focus is on relating network structure as measured by exposure (called integration in their paper) and diversification (the number of firms that hold one's shares) to contagion. Elliott, Golub and Jackson (2014) [see also Jackson and Pernoud (2019, Sec. 2.4)] argue that exposure and diversification at an "intermediate" level make networks susceptible to contagion. *This non-monotonicity is absent in our model.*

A direct comparison of our results with those of Elliott, Golub and Jackson (2014) is not possible because we do not measure diversification in terms of the number of firms that hold one's shares. In our analysis we don't fix the number of counterparties that a firm has equity stakes in *a priori* but let it depend on the realized distribution of equity shares. Our analog to the notion of "density" in the cited work is exposure, the relative share of a firm's equity owned by other firms in the network. As remarked earlier, in our setting we find that the impact of exposure on stability is as follows: holding endowment fixed, an increase in exposure increases the basin of attraction of the "best" equilibrium.

Nevertheless, in Section 10 we outline a variant of our model based on Erdös–Rényi random graphs that does permit a comparison with the cited work and a resolution of the apparently contradictory findings. We argue that the "intermediate level" of integration identified in Elliott, Golub and Jackson (2014) at which non-monotonic behavior is manifested occurs when the underlying network is *very* sparse, that is to say, network topologies potentially susceptible to contagion are characterized by bounded, typically small, node degrees. In Section 10 we deal with the connected domain in Erdös–Rényi networks and their stochastic block model extensions which, even at the junction where connectivity just emerges, are much richer in interconnections than the sparse "intermediate zone" of Elliott, Golub and Jackson (2014). In the connected regime we consider, network behavior is resolutely monotone with an emergent concentration of measure as the network size becomes large.

Unlike Elliott, Golub and Jackson (2014) our focus is not just on the extremal equilibria. We provide a complete characterization of the distribution of *all* equilibria.

We study the stability of the equilibria and their susceptibility to shocks and contagion using a natural dynamic in the spirit of Eisenberg and Noe (2001). In this latter paper the authors consider a network of firms linked through debt rather than equity, characterize consistent clearing vectors of debt repayments between firms, and provide an algorithm—referred to as the "fictitious sequential default" algorithm—for determining such vectors. Eisenberg and Noe (2001) interpreted the algorithm as a process of dynamic adjustment of firm valuations. They do not provide results that relate network structure to contagion.

Finally, our work is related to Dasaratha (2020), which characterizes eigenvector centrality—defined via a recurrence, similar to our equilibria, albeit, in a linear form—of a wide range of networks. These centrality issues are elegantly addressed using existing concentration results for random matrices. These results are of independent interest but not applicable to our setting. Our focus is on the actual equilibria and not centrality, our model has added complexities in view of non-linearities, and, from a technical point of view, significant hurdles are created by the fact that the relevant network matrices in our case do *not* have independent entries. The bulk of the existing literature on concentration deals with matrices with independent entries and those results will not carry through to our setting. Thus, a technical contribution of this paper is a concentration result for random matrices with dependent entries.

3. NOTATIONAL PRELIMINARIES

All vectors are interpreted as column-vectors unless stated otherwise. We reserve the special notation $\mathbb{1}_n$ for the vector of all ones of length n and drop the subscript and write simply $\mathbb{1}$ when we can safely do so without ambiguity when the dimension is clear from context. In the same spirit, we will also write $\mathbb{1}_{n\times m} = \mathbb{1}_n \mathbb{1}_m^\intercal$ for the $n \times m$ matrix of ones. We will in analogous fashion reserve the notation $\mathbb{0} = \mathbb{0}_n$ for the vector of length n all of whose elements are equal to zero and write $\mathbb{0}_{n\times m}$ for the $n \times m$ matrix of zeros.

The names of generic vectors and matrices are displayed in boldface. Vector inequalities are understood elementwise: if $\mathbf{x} = (x_1, \dots, x_n)^\intercal$ and $\mathbf{y} = (y_1, \dots, y_n)^\intercal$ are two n-dimensional real vectors, then the vector inequality $\mathbf{x} \le \mathbf{y}$ means that $x_i \le y_i$ for each i. The expression $\operatorname{diag}(\mathbf{x})$ represents the $n \times n$ diagonal matrix with the elements of $\mathbf{x}$ residing on its main diagonal.

If $\mathbf{A}$ is a generic matrix, we write $\mathbf{A}_{i*}$ for its ith row and $\mathbf{A}_{*j}$ for its jth column. If $\mathbf{A}$ is any real square matrix, its matrix exponential $\exp(\mathbf{A})$ is defined as $\exp(\mathbf{A}) := \sum_{\ell=0}^{\infty} \frac{\mathbf{A}^\ell}{\ell!}$.

If $\mathbf{A}$ is symmetric then $\operatorname{spec}(\mathbf{A})$ stands for its spectrum—its set of eigenvalues—and $\lambda_{\text{max}}(\mathbf{A})$ denotes its largest eigenvalue. More generally, if $\mathbf{A}$ is an $n \times m$ real rectangular matrix, then $\sigma_{\text{max}}(\mathbf{A})$ and $\sigma_{\text{min}}(\mathbf{A})$ represent its largest and smallest singular values, respectively.

We utilize the standard definitions of the vector ℓ^p-norm, $\|\mathbf{x}\|_p$, for $1 \le p \le \infty$. In the special case of $\|\mathbf{x}\|_1$ where $\mathbf{x} \in \{0,1\}^n$, the ℓ^1-norm counts the number of 1s in the vector and we will take a liberty with notation and write $\|\mathbf{x}\|_1 = |\mathbf{x}|$ for the ℓ^1-norm in this context.

As is customary, we recycle notation for the matrix norms induced by vector norms. Suppose $\mathbf{A} = [a_{ij}]_{i,j=1}^n$ is any real square matrix of order n. For $1 \le p \le \infty$, identify the operator p-norm $\|\mathbf{A}\|_p$ with the norm induced by the corresponding vector ℓ^p-norm via

$$\|\mathbf{A}\|_p := \sup_{\mathbf{x} \ne \mathbb{0}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}.$$

We will encounter the special cases $p = 1$, 2, and ∞: in these cases the operator norms take on simple aspects. The operator 1-norm may be identified with the maximum absolute column sum: $\|\mathbf{A}\|_1 = \max_{1\le j\le n} \sum_{i=1}^n |a_{ij}|$; the operator 2-norm (also called the spectral norm) may be identified with the largest singular value: $\|\mathbf{A}\|_2 = \sigma_{\text{max}}(\mathbf{A})$; and the operator ∞-norm may be identified with the maximum absolute row sum: $\|\mathbf{A}\|_\infty = \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|$.

Finally, in the usual abuse of notation, we will also use this terminology for random variables in L^p-spaces: if X is a random variable in some probability space and $X \in L^p$ for some $p \ge 1$, the L^p-norm of X is $\|X\|_p := \mathbf{E}\big(|X|^p\big)^{1/p}$ where $\mathbf{E}(\cdot)$ is expectation with respect to the underlying probability measure.

4. A CROSS-SHAREHOLDING MODEL

We consider a network of n interconnected financial firms. Each firm i has a cash endowment e_i and valuation V_i. We think of the vector of cash endowments $\mathbf{e} = (e_1, \dots, e_n)^\intercal \in \mathbb{R}_+^n$ as a fixed parameter of the model. Our focus is on the vector of firm valuations $\mathbf{V} = (V_1, \dots, V_n)^\intercal$ which represents the state of the system. We suppose that $\mathbf{V}$ takes values in some bounded box $\mathcal{D} \subset \mathbb{R}^n$ serving as the state space or the domain of firm valuations.

The valuation V_i of each firm i depends upon its cash endowment and the dividends on its investments in other firms. Interdependencies between firms holding each other's shares are captured by a matrix of cross-holdings $\mathbf{C} = [C_{ij}]_{i,j=1}^n$ where $C_{ij} \in [0, 1)$ is the share of firm j's equity owned by firm i. For each j, we will suppose that the totality of the shares of firm j owned by the firms in the network is strictly less than one, $C_{1j} + \cdots + C_{nj} < 1$, and interpret the deficit $\widehat{C}_j := 1 - \sum_i C_{ij} > 0$ as the share of firm j owned by outside investors external to the network. Identify $\widehat{\mathbf{C}} := \operatorname{diag}\bigl(\widehat{C}_1, \dots, \widehat{C}_n\bigr)$.

We may interpret the cross-holdings matrix $\mathbf{C} = [C_{ij}]_{i,j=1}^n$ as the adjacency matrix of an edge-weighted directed graph with self-loops allowed. We set $C_{ij} = 0$ if the edge (i, j) is absent in the graph.

An *equilibrium profile of firm valuations* (or, simply, *equilibrium*), $\mathbf{V}$, is identified with a solution of the fixed point equation

$$\mathbf{V} = \mathbf{e} + \mathbf{C}\mathbf{V} - \beta \mathbb{1}_{\{\mathbf{V} \leq \tau \mathbb{1}\}} \tag{1}$$

where τ and β are fixed positive values specified in the model, and the vector indicator $\mathbb{1}_{\{\mathbf{V} \leq \tau \mathbb{1}\}}$ picks out indices i for which $V_i \leq \tau$: the ith component of $\mathbb{1}_{\{\mathbf{V} \leq \tau \mathbb{1}\}}$ is equal to 1 if $V_i \leq \tau$ and is equal to 0 if $V_i > \tau$. We interpret the fixed point equation (1) as saying that the valuation V_i of firm i is comprised of its cash endowment e_i, together with the dividends $\sum_j C_{ij} V_j$ it receives on its investment in other firms, but subject potentially to a shock penalty β (representing, for example, a fire sale or debt write-down) applied if its valuation is below the fixed threshold τ.

We will refer to the firms whose valuations fall below τ—and who suffer the fixed penalty β—as *insolvent*. The language, while admittedly taking liberties with the meaning of the term, has the virtue of being vivid.

DEFINITION 1—Solvent and Insolvent Firms: Given a solvency threshold τ, fixed in the model, we say that a firm $i \in \{1, \dots, n\}$ is *solvent* if $V_i > \tau$, and *insolvent* otherwise.

Equation (1) is the model of Elliott, Golub and Jackson (2014, p. 9, eq. (4)) and we refer the reader there for a detailed discussion. There are three points of departure in our paper from their formulation, one major and two minor, and we touch upon these briefly before proceeding.

i. *(Major)* A focus of Elliott, Golub, and Jackson was the number of counterparties each firm has and its influence on the impact of integration and diversification on the risk of contagion. In our paper we shift focus to the actual amount of equity that is shared. A major point of departure then becomes a nuanced consideration of the rôle of the wide range of possible cross-holdings: we accomplish this by a consideration of *random* selections of cross-holdings of shares (random networks). As we shall see, this change in perspective leads to insights of a very different character.
ii. *(Minor)* In Elliott, Golub, and Jackson, firm valuations are thought of as arising from the values of primitive assets or factors of production. These may be viewed as investments in external assets that generate a flow of cash over time. The particulars of these primitive assets are not critical in our setting and we have rolled the net contribution for each firm i into a fixed cash holding or endowment e_i which may be thought of in the Elliot, Golub, Jackson framework as the return from a single primitive asset.

iii. *(Minor)* Elliott, Golub, and Jackson distinguish between the book value of the firm and its market value which is the value held by outside investors. They differ by a scale factor. The book value is what we call the firm valuation denoted $\mathbf{V}$. The vector of market valuations is $\mathbf{v} := \widehat{\mathbf{C}}\mathbf{V}$. We use $\mathbf{V}$, in the insolvency indicator $\mathbb{1}_{\{\mathbf{V}\le\tau\mathbb{1}\}}$ on the right in (1) to trigger penalties. In Elliott, Golub, and Jackson, the insolvency indicator is replaced by $\mathbb{1}_{\{\mathbf{v}\le\tau'\mathbb{1}\}}$. The two formulations are equivalent once τ is scaled appropriately.

There is a natural dynamic which allows us to interpret the fixed-point equation (1) as equilibria of a dynamical system. If firm valuations are allowed to adjust with time, we write $\mathbf{V}_t$ and $V_{i,t}$ to reflect the values of $\mathbf{V}$ and V_i, respectively, at time t. In the dynamic setting we interpret $V_{i,t}$ as the "estimated" valuation of firm i at time t, with $V_{i,\infty}$ denoting its equilibrium valuation (assuming convergence).

In discrete-time, the dynamical system takes the form

$$\mathbf{V}_{t+1} = \mathbf{e} + \mathbf{C}\mathbf{V}_t - \beta\mathbb{1}_{\{\mathbf{V}_t\le\tau\mathbb{1}\}}. \tag{2}$$

In this formulation, firm valuations $\mathbf{V}_t$ adjust in time based on cash endowments $\mathbf{e}$ and the dividends $\mathbf{C}\mathbf{V}_t$ on the investments in other firms at their current valuations $\mathbf{V}_t$, subject to a potential penalty of β applied whenever the current valuation drops below the solvency threshold τ. The non-linear plunge in a firm's valuation can be interpreted as indicative of distress if we interpret the process (2) as describing the actual change in firm valuations in time. Alternatively, we can interpret it as a "fictitious" dynamic [in the language of Eisenberg and Noe (2001)] that an external party uses to arrive at a consistent set of firm valuations starting from an initial estimate.

The continuous time counterpart of (2) is given by

$$\frac{d}{dt}\mathbf{V}_t = \dot{\mathbf{V}}_t = \mathbf{e} - (\mathbf{I}-\mathbf{C})\mathbf{V}_t - \beta\mathbb{1}_{\{\mathbf{V}_t\le\tau\mathbb{1}\}}. \tag{3}$$

The equilibria of both (2) and (3) are solutions of the fixed point equation (1), but the continuous-time version (3) is analytically easier to work with in a consideration of the dynamics.

5. EQUILIBRIA IN FIXED NETWORKS

The equilibria of the model are the solutions of the equation (1). An alternative characterization is available in view of the following observation.

LEMMA 1—[Proof ↗]: *The matrix* $\mathbf{I}-\mathbf{C}$ *is non-singular.*

We defer the proof of this and the other assertions in this section to Appendix A.

In view of Lemma 1, we may rewrite the fixed point equation (1) as

$$\mathbf{V} = (\mathbf{I}-\mathbf{C})^{-1}(\mathbf{e} - \beta\mathbb{1}_{\{\mathbf{V}\le\tau\mathbb{1}\}}). \tag{4}$$

When $\beta = 0$, the model possesses a unique equilibrium $(\mathbf{I}-\mathbf{C})^{-1}\mathbf{e}$. In general, the system (4) can have many solutions. A partitioning of the state space enumerates the possibilities.

Each partition of the collection of firms into two sets, one identified with solvency and the other with insolvency, identifies a unique subset of the state space that we call an *orthant* for reasons that will become clear directly. Let us define $\mathbf{k} = (k_1,\dots,k_n)^\intercal \in \{0,1\}^n$ to be an *indicator of solvency* of the corresponding firms ($1 \equiv$ solvent, $0 \equiv$ insolvent): for a given indicator of solvency $\mathbf{k}$, firm i is solvent if $k_i = 1$ and insolvent if $k_i = 0$. We further define $\mathbb{K}^n(\mathbf{k})$ to be an orthant in $\mathbb{R}^n$, anchored at point $\tau\mathbb{1}$ rather than the origin, identifying the space of firm valuations with solvency characterized by $\mathbf{k}$:

$$\mathbb{K}^n(\mathbf{k}) := \{\mathbf{x} : x_i > \tau \text{ if } k_i = 1 \text{ and } x_i \leq \tau \text{ if } k_i = 0\}.$$

The language, of course, is an extension of quadrants in two dimensions: the 2^n orthants thus engendered as $\mathbf{k}$ sweeps across $\{0,1\}^n$ partition $\mathbb{R}^n$. We may group the 2^n orthants into $n+1$ equivalence classes, $\mathcal{R}_0$, $\mathcal{R}_1$, $\dots$, $\mathcal{R}_n$, where equivalence class $\mathcal{R}_k$ consists of the $\binom{n}{k}$ orthants $\mathbb{K}^n(\mathbf{k})$ for which $|\mathbf{k}| = k$. That is to say, $\mathcal{R}_k$ is the collection of orthants each of which is characterized by a distinct set of exactly k solvent firms. When we are concerned only with the number k of solvent firms in an orthant, we write $\mathbb{K}(k)$ to refer to *any* one orthant $\mathbb{K}^n(\mathbf{k})$, $|\mathbf{k}| = k$, from class $\mathcal{R}_k$, dropping dimensionality n if clear from context.

Fix any solvency-identifying index vector $\mathbf{k} \in \{0,1\}^n$. If $\mathbf{V}$ is any point in the corresponding orthant $\mathbb{K}^n(\mathbf{k})$ then $\mathbb{1}_{\{\mathbf{V}\leq\tau\mathbb{1}\}} = \mathbb{1} - \mathbf{k}$, as the indicator picks out precisely those indices i for which $V_i \leq \tau$. It follows that, if a solution to the fixed point equation (1) exists in orthant $\mathbb{K}^n(\mathbf{k})$, it must be given by

$$\mathbf{V} = \mathbf{V}(\mathbf{k}) := (\mathbf{I} - \mathbf{C})^{-1}\big(\mathbf{e} - \beta(\mathbb{1} - \mathbf{k})\big) \qquad \big(\mathbf{k} \in \{0,1\}^n\big). \tag{5}$$

This is a *putative*[1] equilibrium. It will be *feasible* if, and only if, $\mathbf{V}(\mathbf{k}) \in \mathbb{K}^n(\mathbf{k})$.

As the orthants $\mathbb{K}^n(\mathbf{k})$ partition the space, it follows that, if the fixed point equation (1) has any feasible solution(s) at all, then it (they) must lie in the collection of putative equilibria $\big\{\, \mathbf{V}(\mathbf{k}) : \mathbf{k} \in \{0,1\}^n \,\big\}$. But is there any feasible equilibrium at all in this collection? Yes: existence is guaranteed by the Knaster–Tarski theorem.

THEOREM 1—Equilibrium Existence [Proof]: *The fixed point equation* (1) *has at least one solution.*

Candidates for extremal equilibria are easy to identify. Set

$$\begin{aligned} \mathbf{V}^{\text{sup}} &:= \mathbf{V}(\mathbb{1}) = (\mathbf{I} - \mathbf{C})^{-1}\mathbf{e}, \\ \mathbf{V}^{\text{inf}} &:= \mathbf{V}(\mathbb{0}) = (\mathbf{I} - \mathbf{C})^{-1}(\mathbf{e} - \beta\mathbb{1}). \end{aligned} \tag{6}$$

Then $\mathbf{V}^{\text{sup}}$ is a candidate for an equilibrium in the "best" orthant $\mathbb{K}(n) = \mathbb{K}^n(\mathbb{1}) = \{\, \mathbf{V} : \mathbf{V} > \tau\mathbb{1} \,\}$ in which all firms are solvent; it is feasible if, and only if, $\mathbf{V}^{\text{sup}} > \tau\mathbb{1}$. Likewise, $\mathbf{V}^{\text{inf}}$ is a candidate for an equilibrium in the "worst" orthant $\mathbb{K}(0) = \mathbb{K}^n(\mathbb{0}) = \{\, \mathbf{V} : \mathbf{V} \leq \tau\mathbb{1} \,\}$ in which all firms are insolvent; it is feasible if, and only if, $\mathbf{V}^{\text{inf}} \leq \tau\mathbb{1}$. If we set

$$f(\mathbf{V}) := \mathbf{e} + \mathbf{C}\mathbf{V} - \beta\mathbb{1}_{\{\mathbf{V}\leq\tau\mathbb{1}\}},$$

then the natural iteration

$$\mathbf{V}_{t+1} = f\big(\mathbf{V}_t\big) \qquad (t \geq 0)$$

converges to one of the extremal fixed point candidates depending on whether the iteration is started at a "very large" or a "very small" initial vector of valuations $\mathbf{V}_0$. Necessary and sufficient conditions for the feasibility of these extremal equilibria may be deduced from the defining equations (6).

[1]While, formally, putative equilibria may not be feasible, they affect the shape of the model state's trajectory in the dynamic version of our model.

THEOREM 2—Extremal Equilibria [Proof]: *Suppose equal cash endowments* $\mathbf{e} = e\mathbb{1}$ *and that every firm* i *holds a positive amount of equity shares, that is to say,* $C_{ij} > 0$ *for at least one* j*. Then:*

a) *For* $\mathbf{V}^{\text{sup}}$ *to be an equilibrium it is necessary that* $e > \tau \cdot \sigma_{\min}(\mathbf{I} - \mathbf{C})$*; for it to be an equilibrium it is sufficient that* $e > \tau$.

b) *For* $\mathbf{V}^{\text{inf}}$ *to be an equilibrium it is necessary that* $e \leq \tau + \beta$*; for it to be an equilibrium it is sufficient that* $e \leq \tau \cdot \frac{\sigma_{\min}(\mathbf{I}-\mathbf{C})}{\sqrt{n}} + \beta$.

The bounds in Theorem 2 are crude—this is the price of extreme generality. A more refined understanding of the feasible equilibria of the model will have to devolve upon the imposition of more structure.

6. THE EQUILIBRIA OF THE REGULAR CLIQUE

We now fix the cash endowment of each firm at the same nominal value e. Suppose additionally that a fixed fraction $c \in (0,1)$ of each firm's equity is distributed across the network, with the residual equity proportion $1 - c$ held by outside investors. We do this so as to focus on the role of network structure on equilibrium outcomes.

An egalitarian distribution of equity will result in each firm distributing a fraction $\frac{c}{n}$ of its equity to every firm in the network. We call this a *regular clique*. The matrix of cross-holdings (or *adjacency matrix*) takes on a particularly simple form in this setting with every element equal to $\frac{c}{n}$ or, in matrix form, $\mathbf{C} = \mathbf{C}^0 := \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$.

The simplicity of the regular clique is appealing for an initial foray into the model structure by way of building intuition into the nature of equilibria. The symmetries inherent in the model permit explicit analytical characterizations. But we accrue much much more by a careful examination of this special case. As we shall see in the following section, the regular clique is typical of the entire class of fully diversified networks. Hence, by focusing on its properties we learn about the properties of all fully diversified networks.

The dependence of equilibria upon the model parameters in the regular clique is simple. The fixed point condition (1) is

$$\mathbf{V} = e\mathbb{1} + \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\mathbf{V} - \beta\mathbb{1}_{\{\mathbf{V}\leq\tau\mathbb{1}\}}.$$

The penalty offset on the right ranges over 2^n possibilities depending on which subset of firms are solvent. Some notation will help navigate the labyrinth.

Recall that each index vector $\mathbf{k} = (k_1, \dots, k_n)^\intercal \in \{0,1\}^n$ is associated with a unique *orthant* $\mathbb{K}^n(\mathbf{k})$ in $\mathbb{R}^n$ consisting of those points $\mathbf{x} = (x_1, \dots, x_n)^\intercal$ satisfying the inequalities $x_i > \tau$ if $k_i = 1$ and $x_i \leq \tau$ if $k_i = 0$. If $\mathbf{V} \in \mathbb{K}^n(\mathbf{k})$ then $\mathbb{1}_{\{\mathbf{V}\leq\tau\mathbb{1}\}} = \mathbb{1} - \mathbf{k}$ and so any given $\mathbf{k}$ engenders a *putative* equilibrium $\mathbf{V} = \mathbf{V}^0(\mathbf{k})$ indexed by $\mathbf{k}$ and satisfying the fixed point equation

$$\mathbf{V}^0(\mathbf{k}) = \big((e-\beta)\mathbb{1} + \beta\mathbf{k}\big) + \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\mathbf{V}^0(\mathbf{k}). \tag{7}$$

This will be a *feasible* equilibrium in orthant $\mathbb{K}^n(\mathbf{k})$ if, and only if, the solution satisfies $\mathbf{V}^0(\mathbf{k}) \in \mathbb{K}^n(\mathbf{k})$. The collection of feasible equilibria as $\mathbf{k}$ ranges over all possibilities identifies the collection of equilibrium points of the regular clique.

The symmetries inherent in the situation lead to explicit closed form solutions. For $\mathbf{k} \in \{0,1\}^n$, write $|\mathbf{k}| := k_1 + \cdots + k_n$ for the ℓ^1-norm. For a given cash endowment e, the *feasibility constraint* for a given $\mathbf{k}$ is captured by the system of inequalities

$$\tau(1-c) + \big(1 - \tfrac{|\mathbf{k}|}{n}\big)\beta c < e \leq \tau(1-c) + \beta - \tfrac{|\mathbf{k}|}{n}\beta c \qquad (\text{if } 1 \leq |\mathbf{k}| \leq n-1), \tag{8}$$

$$e \le \tau(1-c) + \beta \qquad \text{(if } |\mathbf{k}| = 0\text{)}, \tag{9}$$

$$e > \tau(1-c) \qquad \text{(if } |\mathbf{k}| = n\text{)}. \tag{10}$$

The reason for the terminology will be apparent shortly. With a view to keeping the theorem statement uncluttered, it will also be useful to define

$$v(s) := \frac{e - (1-s)\beta c}{1-c} \qquad (0 \le s \le 1). \tag{11}$$

THEOREM 3—Equilibria for Regular Cliques [Proof]: *For any given* $\mathbf{k} \in \{0,1\}^n$*, the fixed point equation* (7) *has a unique solution* $\mathbf{V}^0(\mathbf{k})$ *with components given by*

$$V_i^0(\mathbf{k}) = \begin{cases} v\left(\frac{|\mathbf{k}|}{n}\right) & \text{if } k_i = 1, \\ v\left(\frac{|\mathbf{k}|}{n}\right) - \beta & \text{if } k_i = 0, \end{cases} \tag{12}$$

where $v(\cdot)$ *is defined in* (11)*. For it to be a feasible equilibrium of the regular clique in orthant* $\mathbb{K}^n(\mathbf{k})$ *it is necessary and sufficient that the corresponding feasibility constraint in the system of inequalities (8–10) be satisfied. Setting* $\mathbf{k} = \mathbb{0}$ *and* $\mathbb{1}$ *in turn, it follows a fortiori that the extremal equilibria are given putatively by*

$$\begin{aligned} \mathbf{V}^{\text{inf}} &= \mathbf{V}^0(\mathbb{0}) = \big(v(0) - \beta\big)\mathbb{1} = \frac{e-\beta}{1-c}\mathbb{1}, \\ \mathbf{V}^{\text{sup}} &= \mathbf{V}^0(\mathbb{1}) = v(1)\mathbb{1} = \frac{e}{1-c}\mathbb{1}. \end{aligned} \tag{13}$$

They are feasible if the corresponding feasibility constraints (9, 10), respectively, hold.

The proof, presented in Appendix B, follows from the fact that the matrix $\mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$ is almost diagonal.

The feasibility constraints (8–10) depend only on $|\mathbf{k}|$. Therefore, we group the orthants into the $n+1$ equivalence classes, $\mathcal{R}_0, \mathcal{R}_1, \ldots, \mathcal{R}_n$, where, as defined earlier, the equivalence class $\mathcal{R}_k$ consists of the $\binom{n}{k}$ orthants $\mathbb{K}^n(\mathbf{k})$ for which $|\mathbf{k}| = k$. When cash endowments are equal, the symmetry of the regular clique enjoins that if there is a feasible equilibrium in any orthant $\mathbb{K}^n(\mathbf{k})$ in the equivalence class $\mathcal{R}_k$, then there will be feasible equilibria in all the orthants $\mathbb{K}^n(\mathbf{k})$ in the class. The form of the solution (33) reinforces this observation.

How does the set of equilibrium profiles of valuations change as e changes? The "inverted Z" in Fig. 1 shows the relationship between the cash endowment e and the presence of feasible equilibria in the equivalence classes $\mathcal{R}_k$. For each $1 \le k \le n-1$, the corresponding feasibility constraint (8) identifies an endowment interval of length $\beta(1-c)$; the feasibility constraints for the bookend cases $k = 0$ and $k = n$ form the foot and head of the inverted Z. For any given e, the intersection of the vertical line at e with the horizontal feasibility constraint intervals identifies those equivalence classes containing distinct feasible equilibria.

We can use the inverted Z to understand what happens as we shrink the endowment holding exposure fixed. Fig. 2 attempts to convey snapshots of the evolution of the number and distribution of equilibria as e decreases from left to right, top to bottom. Each vertical pair of figures, from (a) to (f), corresponds to a particular value of e. In each vertical figure pair, the top figure identifies, for each e, the equivalence classes $\mathcal{R}_k$ of orthants that contain feasible equilibria. The bottom figure in each figure pair schematically illustrates which orthants are populated

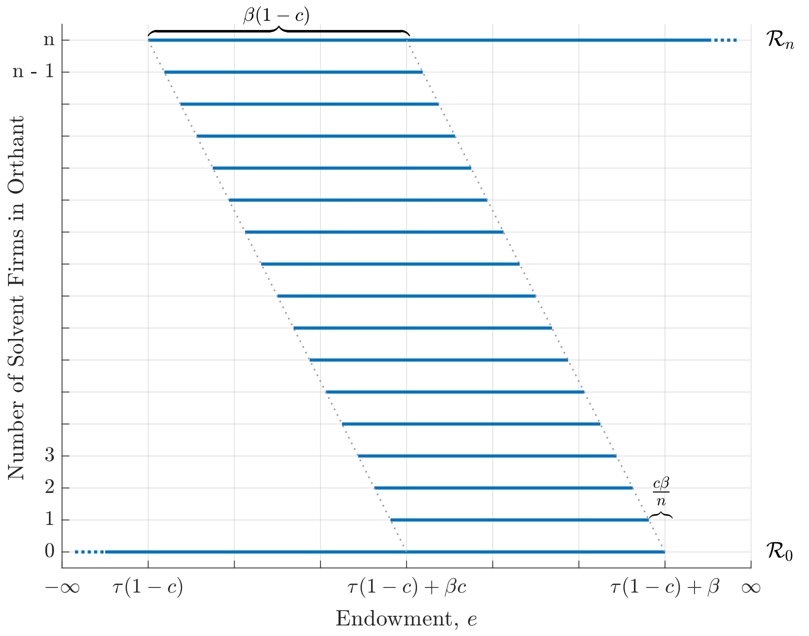

FIGURE 1.—A schematic of endowment versus solvency index illustrating the feasibility constraints (8, 9, 10) for the existence of equilibria in the equivalence classes, $\mathcal{R}_k$, of orthants. Each horizontal line identifies an endowment interval for which orthants in the identified equivalence class contain distinct feasible equilibria.

with equilibria. The two-dimensional grid is a convenient[2] visualization of the orthants: the square in the northeast corner represents the sole orthant $\mathbb{K}(n) = \mathbb{K}^n(\mathbb{1})$ in the equivalence class $\mathcal{R}_n$; the square in the southwest corner represent the sole orthant $\mathbb{K}(0) = \mathbb{K}^n(\mathbb{0})$ in the equivalence class $\mathcal{R}_0$. The orthants that are home to equilibria where at least half the firms are insolvent appear below the off-diagonal. Above the off-diagonal are the orthants home to equilibria where at most half the firms are insolvent.

We start with the case of "largest" e depicted in (a). In this case there is only one orthant containing an equilibrium: it is the "best" equilibrium point $\mathbf{V}^{\text{sup}}$ in orthant $\mathbb{K}^n(\mathbb{1})$ which is the sole representative in equivalence class $\mathcal{R}_n$ at the northeast corner of the grid.

As e declines, the "best" equilibrium $\mathbf{V}^{\text{sup}}$ acquires a companion "worst" equilibrium $\mathbf{V}^{\text{inf}}$ in orthant $\mathbb{K}^n(\mathbb{0})$ which is the sole representative in the class $\mathcal{R}_0$ at the southwest corner of the grid.

Decreasing e further, results in equivalence classes $\mathcal{R}_k$ with a minority of solvent firms becoming populated with equilibria near the southwest corner of the grid as illustrated in (b). As e continues to decline, more and more orthants with a majority of insolvent firms acquire equilibria, and in (c) orthants in the classes $\mathcal{R}_k$ near the southwest corner are all populated with equilibria while the other half of the state space of orthants are almost bereft of equilibria excepting the "best" equilibrium point $\mathbf{V}^{\text{sup}}$ in the northeast corner.

One should not confuse the presence of a plethora of equilibria in which a large proportion of firms are insolvent as a sign of fragility. As we show in Section 8, once we associate a dynamic with these equilibria, the basin of attraction of the "best" equilibrium remains large. Hence, a small (or even moderate) shock to the profile of valuations at the "best" equilibrium does not

[2]The square grid distorts the true picture of the actual number of orthants in each class $\mathcal{R}_k$: as noted earlier, the actual number of orthants in a given equivalence class varies binomially rather than linearly with index.

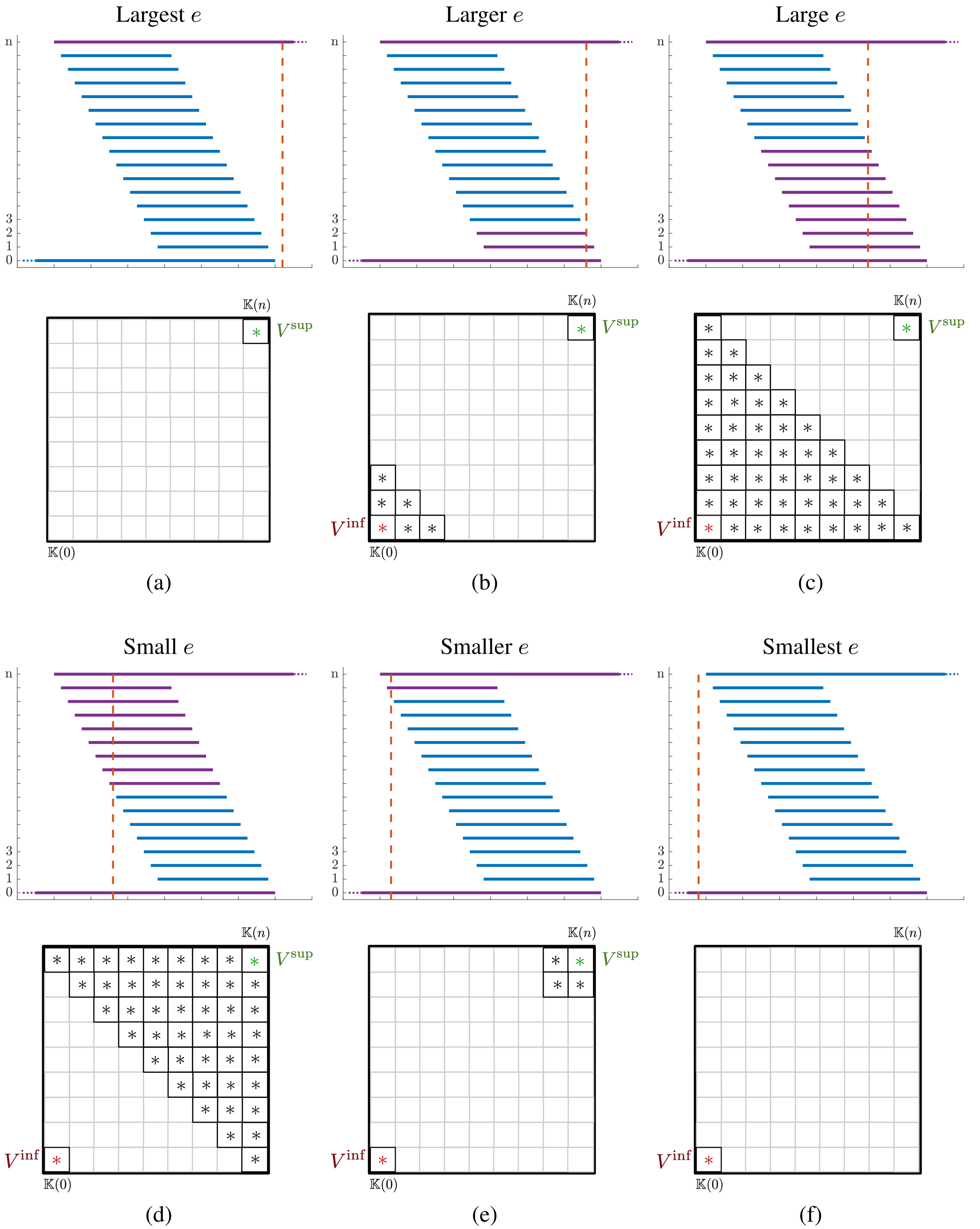

FIGURE 2.—Snapshots of the evolution of the number of equilibria as the cash endowment decreases. The top row identifies equivalence classes $\mathcal{R}_k$ that contain feasible equilibria. The bottom row schematically illustrates the distribution of equilibria across orthants.

result in widespread default. In fact, the system is resilient in the sense that after a small shock to the "best" equilibrium, the system returns to its initial position.

As the endowment continues to decline in (d)–(f), the equilibrium points from the orthants closest to $\mathcal{R}_0$ vanish, but the population of equilibria in the half of the space of orthants closest to $\mathcal{R}_n$ increases.

Decreasing e further extinguishes all equilibria except the "worst" and "best" equilibria, $\mathbf{V}^{\text{inf}}$ and $\mathbf{V}^{\text{sup}}$. A further decrease in endowment (f) eliminates the "best" equilibrium so that, eventually, only the "worst" equilibrium $\mathbf{V}^{\text{inf}}$ remains.

7. EQUILIBRIA IN RANDOM NETWORKS

This section contains the main result of this paper. In it we justify the careful consideration of the regular clique by showing that, in a formal sense, it is the archetypal fully diversified cross-holding network. We accomplish this by a consideration of a *random* selection of cross-shareholding matrix $\mathbf{C}$. This permits the *simultaneous* consideration of all instances of cross-holdings, most of a highly irregular character. Its benefit rests in the idea that randomness creates probabilistic symmetries that can be exploited analytically. And its utility devolves from the phenomenon of concentration of measure wherein, as we shall see, "most" networks, very different from each other in their realizations, behave very similarly in the disposition and nature of their equilibria.

7.1. *The Scaffolding and a Preview of the Results*

In order to understand the impact of interdependencies on contagion and cascades of failures, we consider networks where firms, while potentially having very different investment portfolios, follow similar (random) investment strategies. The setting is sanitized to be sure but it permits us to focus on the interaction between diversification and contagion. With this objective in mind, we hold other parameters in the system, such as the cash holdings of each firm (more generally, the primitive assets or factors of production) and the fraction of each firm's value owned by outside shareholders external to the network, fixed at some nominal values. Our models are qualitatively different from settings which are topologically non-uniform—included among these are Erdös–Rényi topologies, stochastic block models, core-periphery networks, or scale-free-like settings—but many of our results have natural extensions in these directions. We outline one of these in Section 10.

We will need to build up some technical machinery but some motivation is always welcome before diving into dry technicalities. Suppose, as in the previous section, that the cash holdings of each firm are fixed at the nominal value e and that a fixed fraction c of each firm's equity is distributed across the network. The notion that firms have similar random investment strategies is captured by a consideration of a random matrix of cross-holdings $\mathbf{C}$ (a *random clique*) whose columns are independent and identically distributed and share a common distribution. The notion of probabilistic balance is captured in the idea that the distribution of a firm's equity across network entities is an *exchangeable* system satisfying a suitable regularity constraint to avoid pathologies: we will define what this means precisely later in this section in a notion that we call *asymptotically diffuse*. Recall that, for $\mathbf{k} \in \{0,1\}^n$, the vector $\mathbf{V}^{(0)}(\mathbf{k})$ represents the putative equilibrium in orthant $\mathbb{K}^n(\mathbf{k})$ of the regular clique $\mathbf{C}^0$. Let the analogous putative equilibrium corresponding to an asymptotically diffuse random clique $\mathbf{C}$ be $\mathbf{V}(\mathbf{k})$ (if it exists).

The following theorem provides an inviting flavor of the kind of result that is now within reach. In rough terms: *the feasible equilibria of the random clique converge point-wise to the feasible equilibria of the regular clique*. Or, with a little more precision in language:

THEOREM 4—Equilibria in Large Random Networks: Simplified Version: *Select $\epsilon > 0$ and $\delta > 0$ as small as desired. Then, for all sufficiently large n, and any selected index vector $\mathbf{k} \in \{0,1\}^n$, with probability at least $1-\delta$, if $\mathbf{V}^0(\mathbf{k}) \in \mathbb{K}^n(\mathbf{k})$ then so is $\mathbf{V}(\mathbf{k})$ and, moreover, each component $V_i(\mathbf{k})$ of the equilibrium of the random clique is ϵ-close to the corresponding component $V_i^0(\mathbf{k})$ of the equilibrium of the regular clique.*

We will actually be able to say much much more but we will need to first detour through some preliminaries.

7.2. *The Probabilistic Skeleton*

Suppose that the elements of the stochastic (column) vector $\mathbf{X} = (X_1, \dots, X_n)^\intercal$ are exchangeable, nonnegative-valued random variables summing to 1. These will represent, up to a scale factor, the distribution of a given firm's equity across the entities in the network. If $F_n = F_n(x_1, \dots, x_n)$ denotes the (joint) distribution of $X_1, \dots, X_n$ then: exchangeability enjoins that $F_n = F_n(x_1, \dots, x_n)$ is invariant with respect to permutations of its coordinates; positivity enjoins that F_n is supported in the first orthant $x_1, \dots, x_n \geq 0$; and the sum constraint enjoins that F_n is singular and concentrated on the hyperplane $x_1 + \cdots + x_n = 1$. As an immediate consequence of exchangeability, we see that the component-wise expectations and variances must be equal, as indeed must all moments. For the same reason, all mixed moments are also invariant with respect to permutations of the spacings $X_1, \dots, X_n$ but more can be said: the sum constraint $X_1 + \cdots + X_n = 1$ enjoins indeed that *the spacings are negatively correlated*. A rich variety of identities arise out of this simple observation. We will defer these to Lemma 4 in Appendix C and satisfy ourselves here with a pared down, second moment variation on the theme for immediate reference—the technical details of the full monty will become important only in the proof of the main theorem and are not essential here.

PROPOSITION: *Suppose $X_1, \dots, X_n$ are exchangeable, non-negative, and sum to 1. Then, they have common mean $1/n$, and the common pairwise covariance γ is related to the common variance s^2 through the relation $\gamma = -\frac{1}{n-1}s^2$. In words: the variables are negatively correlated and asymptotically weakly dependent.*

The following examples provide several natural ways of generating such families of exchangeable random variables.

EXAMPLE 1) *The de Finetti spacings.* Identify $X_1, \dots, X_n$ with the spacings engendered by throwing $n-1$ points uniformly at random in the unit interval. The corresponding joint distribution was obtained by Bruno de Finetti [see Venkatesh (2013, §IX.2, p. 282)]. Write x_+ for the positive part of x. Then, for any non-negative $x_1, \dots, x_n$,

$$\mathbf{P}\{X_1 > x_1, \dots, X_n > x_n\} = \big[(1 - x_1 - \cdots - x_n)_+\big]^{n-1}, \tag{14}$$

whence the distribution function $F_n(x_1, \dots, x_n) = \mathbf{P}\{X_1 \leq x_1, \dots, X_n \leq x_n\}$ can be written down by inclusion and exclusion. An elementary integration now show that, if $1 \leq i \leq n$ and $i' \neq i$, then

$$\mathbf{E}(X_i) = \frac{1}{n} \text{ and } \operatorname{Var}(X_i) = \frac{n-1}{n^2(n+1)} \sim \frac{1}{n^2}. \tag{15}$$

These assertions are proved in Lemma 10 in Appendix C; we will see indeed in Corollary 6 that $\mathbf{E}\big((X_i - 1/n)^\nu\big) = \mathcal{O}(n^{-\nu})$ for every non-negative integer ν. ▶

The de Finetti distribution (14) is an archetype for our model of a fully diversified cross-holding network; it provides perhaps the most intuitive way to distribute points in the interval in an unbiased fashion. It may not be surprising that other formulations lead to similar results.
EXAMPLE 2) *Spacings induced via the Haar measure on the sphere.* Suppose $\mathbf{Z} = (Z_1, \dots, Z_n) \sim \text{Haar}(\mathbb{S}^{n-1})$ is drawn from the Haar measure on the n-dimensional unit sphere $\mathbb{S}^{n-1} := \{\, \mathbf{z} : \|\mathbf{z}\|_2 = 1 \,\}$. The Haar measure formalizes the notion of "uniformly random selection" on the sphere $\mathbb{S}^{n-1}$: if $\mathbb{A}$ is any Borel-measurable subset of $\mathbb{S}^{n-1}$, then

$$\mathbf{P}\{\mathbf{Z} \in \mathbb{A}\} = \frac{\text{Area}(\mathbb{A})}{\text{Area}(\mathbb{S}^{n-1})} = \frac{\displaystyle\int_{\mathbb{A}} d\Omega}{\displaystyle\int_{\mathbb{S}^{n-1}} d\Omega} = \frac{n\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} \int_{\mathbb{A}} d\Omega \tag{16}$$

where $d\Omega$ represents a differential element with respect to geodesic distance on the unit sphere $\mathbb{S}^{n-1}$. The final step is a consequence of the observation that the area of the unit sphere is given from elementary considerations by [Venkatesh (2013, Theorem XIV.7.2, page 496)]

$$A_n := \text{Area}(\mathbb{S}^{n-1}) = \int_{\mathbb{S}^{n-1}} d\Omega = \frac{2\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}. \tag{17}$$

The Haar measure exhibits rotation invariance: if Q is any orthogonal transformation on $\mathbb{R}^n$ and $\mathbb{A}$ is any Borel set in $\mathbb{S}^{n-1}$ then $\mathbf{P}\{\mathbf{Z} \in \mathbb{A}\} = \mathbf{P}\{\mathbf{Z} \in Q\mathbb{A}\}$. This is the natural analogue on the sphere of the characteristic property of translation invariance of Lebesgue measure on the real line.

Set $X_i = Z_i^2$ for $1 \le i \le n$. Then $X_1, \dots, X_n$ are exchangeable, non-negative variables representing spacings of the unit interval with distribution inherited from (16). In particular, the marginal distribution is given by

$$\mathbf{P}\{X_1 \ge x\} = \mathbf{P}\left\{Z_1 \ge \sqrt{x}\,\right\} = \frac{2A_{n-1}}{A_n} \int_x^1 \frac{(1-t)^{\frac{n-2}{2}}}{2\sqrt{t}}\, dt \qquad (0 \le x \le 1).$$

This assertion is proved in Lemma 15 in Appendix C. Another routine integration now shows that

$$\mathbf{E}(X_i) = \frac{1}{n} \text{ and } \text{Var}(X_i) = \frac{2(n-1)}{n^2(n+2)} \sim \frac{2}{n^2}.$$

The Haar-induced spacings are approximately one-and-a-half times more dispersive than the de Finetti spacings but the moments share the same asymptotic order: we shall see in Corollary 8 in Appendix C that $\mathbf{E}\big((X_i - 1/n)^\nu\big)$ has asymptotic order $n^{-\nu}$ for every $\nu \ge 0$. ▶

The construction in the previous example exploited the fact that $Z_1^2 + \cdots + Z_n^2 = 1$. If the variables are not already normalized properly, it is easy enough to normalize by scaling. This leads to a rich class of distributions of spacings.
EXAMPLES: 3) *Spacings engendered by the exponential density.* Suppose $Z_1, \dots, Z_n$ are independent, exponentially distributed random variables with unit mean. If we set

$$X_i = \frac{Z_i}{Z_1 + \cdots + Z_n} \qquad (1 \le i \le n), \tag{18}$$

then $X_1, \dots, X_n$ have the de Finetti distribution (14) [Venkatesh (2013, §IX.11, Problem 21, p. 312)].

4) *Periodogram analysis.* Suppose $U_1, \dots, U_n, V_1, \dots, V_n$ are independent, standard normal random variables and let $Z_i = U_i^2 + V_i^2$. The variables $X_1, \dots, X_n$ defined anew by (18) again have the de Finetti distribution (14) [Venkatesh (2013, §IX.12, Problem 24, p. 365)]. ▶

More generally, the formulation (18) provides a systematic way of generating exchangeable variables with a sum constraint.

EXAMPLES: 5) *General spacings engendered by scaling.* Suppose $Z_1, \dots, Z_n$ is drawn by independent sampling from an absolutely continuous distribution G with support in the positive half-line. Let $g = G'$ be the corresponding density. Form $X_1, \dots, X_n$ via (18). As in the case of the spacings of the uniform, the variables $X_1, \dots, X_n$ are exchangeable (by virtue of the fact that the underlying product measure $G^{\otimes n}$ is invariant with respect to permutations of coordinates). Again, the variables $X_1, \dots, X_n$ are linearly dependent with $X_1 + \cdots + X_n = 1$. If we introduce the nonce variable $X_0 = Z_1 + \cdots + Z_n$, however, then the variables $X_0, X_1, \dots, X_{n-1}$ are linearly independent. The Jacobian of the transformation

$$Z_1 = X_0 X_1, \ \dots, \ Z_{n-1} = X_0 X_{n-1}, \ Z_n = X_0(1 - X_1 - \cdots - X_{n-1})$$

is $J = x_0^{n-1}$, whence the density of $X_0, X_1, \dots, X_{n-1}$ is

$$x_0^{n-1} \cdot g(x_0 x_1) \cdots g(x_0 x_{n-1}) g\bigl(x_0(1 - x_1 - \cdots - x_{n-1})\bigr).$$

Integrating out x_0 yields the explicit form for the density

$$f_{n-1}(x_1, \dots, x_{n-1}) = \int_0^\infty x_0^{n-1} g(x_0 x_1) \cdots g(x_0 x_{n-1}) \cdot g\bigl(x_0(1 - x_1 - \cdots - x_{n-1})\bigr)\, dx_0 \quad (19)$$

of $X_1, \dots, X_{n-1}$ with support in the regular probability simplex $\mathbb{P}^{n-1}$. No further analytical simplification is possible in general unless g is amenable.

6) *Spacings engendered by folded normals.* If $U_1, \dots, U_n$ are independent, standard normal, then $Z_1 = |U_1|, \dots, Z_n = |U_n|$ are independent with the common folded normal density $g(x) = 2\phi(x) = \sqrt{\frac{2}{\pi}}\, e^{-x^2/2}$ with support only in the positive half-line $[0, \infty)$. In this case (19) simplifies to the explicit form

$$f_{n-1}(x_1, \dots, x_{n-1}) = \left(\frac{4}{\pi}\right)^{\frac{n}{2}} \cdot \frac{\Gamma\left(\frac{n}{2}\right)}{2} \cdot \frac{1}{\left[x_1^2 + \cdots + x_{n-1}^2 + (1 - x_1 - \cdots - x_{n-1})^2\right]^{\frac{n}{2}}}$$

with support again in the regular probability simplex $\mathbb{P}^{n-1}$. ▶

All systems of exchangeable spacings $X_1, \dots, X_n$ are negatively correlated and satisfy the Proposition at the start of this section. Not all exchangeable systems of spacings will have moments that decay as quickly as the de Finetti system (15). Pathological cases arise in "winner take all" scenarios such as coordinate or frame-based spacings.

EXAMPLE 7) *Coordinate spacings lead to pathologies.* Suppose the distribution $F_n(x_1, \dots, x_n)$ is atomic and places equal mass on each of the n atoms $(1, 0, \dots, 0)$, $(0, 1, \dots, 0)$, and $(0, 0, \dots, 1)$. In other words, precisely one of $X_1, \dots, X_n$ is equal to one, the others all taking value 0. Then each X_i individually is a Bernoulli variable taking value 1 with probability $1/n$ and value 0 with probability $1 - 1/n$. It follows that

$$\mathbf{E}(X_i) = \frac{1}{n}, \ \operatorname{Var}(X_i) = \frac{1}{n}\left(1 - \frac{1}{n}\right), \ \operatorname{Cov}(X_i, X_j) = \frac{-1}{n^2}$$

for $1 \le i \le n$ and $j \ne i$. As each variable is individually Bernoulli, we may conclude in fact that $\mathbf{E}\bigl((X_i - 1/n)^\nu\bigr) = \mathcal{O}(n^{-1})$ uniformly for $\nu \ge 1$. ▶

Examples like the one above are pathological. In the context of cross-holdings they concentrate all the available equity of each firm in a single entity, albeit randomly selected: on average, each firm owns the majority stake in a single other firm. This manifests itself in a centered νth moment decaying glacially at rate n^{-1}. De Finetti-type distributions, on the other hand, spread shares more uniformly across the network, a manifestation of which is a centered νth-moment which decays much more rapidly at rate $n^{-\nu}$. With a view to understanding the impact of diversification we will add a smoothness constraint to the system to enforce a certain amount of diversification of equity: this will effectively enjoin that no firm, or small number of firms, corners all shares. The specific form this will take is in the form of an asymptotically uniform eighth-moment constraint.

With asymptotics in n in mind, temporarily introduce a superscript $^{(n)}$ to make explicit the size of the network under consideration. We are then really dealing with a triangular array

$$\begin{array}{ll}
\mathbf{X}^{(1)} & X_1^{(1)} \\
\mathbf{X}^{(2)} & X_1^{(2)}, X_2^{(2)} \\
\mathbf{X}^{(3)} & X_1^{(3)}, X_2^{(3)}, X_3^{(3)} \\
\mathbf{X}^{(4)} & X_1^{(4)}, X_2^{(4)}, X_3^{(4)}, X_4^{(4)} \\
\multicolumn{2}{l}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots} \\
\mathbf{X}^{(n)} & X_1^{(n)}, X_2^{(n)}, X_3^{(n)}, X_4^{(n)}, \dots, X_n^{(n)} \\
\multicolumn{2}{l}{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots}
\end{array}$$

of non-negative random variables where, for each n, the variables $X_1^{(n)}, \dots, X_n^{(n)}$ in the nth row are exchangeable and sum to 1. By exchangeability, $\mathbf{E}\bigl(X_1^{(n)}\bigr) = \frac{1}{n}$ for each n. We will want a little control over the next few moments as well to head off pathological settings like those of Example 7. This leads to a consideration of a dispersive family of distributions that we call *asymptotically diffuse* with an eighth-moment constraint.

In the usual notation, for $p \ge 1$ write $\|X\|_p := \mathbf{E}\bigl(|X|^p\bigr)^{1/p}$ for the L^p-norm of X.

DEFINITION 2: Suppose $\mathcal{X} = \bigl\{ \bigl(X_1^{(n)}, \dots, X_n^{(n)}\bigr) : n \ge 1 \bigr\}$ is a triangular array of row-wise exchangeable, non-negative valued random variables with row sum unit. We say that the array (or more precisely, its distribution in a suitable probability space) is *asymptotically diffuse* if

$$\left\| X_1^{(n)} - \frac{1}{n} \right\|_8 \le \frac{A}{n} \qquad (n \ge 1)$$

for some absolute positive constant A.

The terminology is intended to capture the idea that in an asymptotically diffuse system the unit interval is parceled out equitably across all the spacings and no small group of spacings is typically dominant; or, in the economic context, all parties have a *prima facie* share in the wealth.

If a triangular array is asymptotically diffuse then, in view of the monotonicity of L^p-norms, it follows that

$$\left\| X_1^{(n)} - \frac{1}{n} \right\|_\nu \le \frac{A}{n}, \text{ or, equivalently, } \mathbf{E}\left(\left| X_1^{(n)} - \frac{1}{n} \right|^\nu \right) \le \frac{A^\nu}{n^\nu} \qquad (\text{for } 1 \le \nu \le 8). \tag{20}$$

The de Finetti distributional family of Example 1 is an archetype of such a system: more generally, an asymptotically diffuse triangular array is of the de Finetti-type in the behavior of the low-order moments up through the eighth moment.

7.3. *The Random Network Model*

With preliminaries out of the way, let $\mathcal{X}_1 = \{\mathbf{X}_1^{(n)}, n \geq 1\}$, $\mathcal{X}_2 = \{\mathbf{X}_2^{(n)}, n \geq 1\}$, …, $\mathcal{X}_j = \{\mathbf{X}_j^{(n)}, n \geq 1\}$, … be a sequence of triangular arrays drawn by independent sampling from a common asymptotically diffuse distribution. In particular, for each j and n, the components of the vector $\mathbf{X}_j^{(n)} = \bigl(X_{1j}^{(n)}, \dots, X_{nj}^{(n)}\bigr)^\intercal$ (which constitutes the nth row of the jth array $\mathcal{X}_j$ reassembled in a vector column) are exchangeable, nonnegative-valued random variables summing to 1 and satisfying the eighth-moment constraint $\bigl\| X_{1j}^{(n)} - \frac{1}{n} \bigr\|_8 \leq \frac{A}{n}$ for some absolute constant A.

The idea is to use the nth rows of these arrays to generate the corresponding random cross-holdings matrix for that value of n. The process is as follows: let $0 < c < 1$ be a fixed positive value representing the fraction of each firm's equity collectively owned by all the firms. For each n, set $C_{ij}^{(n)} = cX_{ij}^{(n)}$ and identify the corresponding random matrix of cross-holdings

$$\bigl[C_{ij}^{(n)}\bigr]_{i,j=1}^{n} = \mathbf{C}^{(n)} = c\,\bigl[\mathbf{X}_1^{(n)}\ \mathbf{X}_2^{(n)} \cdots \mathbf{X}_n^{(n)}\bigr]$$

obtained by identifying the columns in turn with the nth rows of the triangular arrays $\mathcal{X}_1$, …, $\mathcal{X}_n$. For each n, the matrix $\mathbf{C}^{(n)}$ has independent columns, each column $\mathbf{C}_j^{(n)} = c\mathbf{X}_j^{(n)}$ consisting of a family of exchangeable, non-negative valued variables summing to c and representing the distribution of ownership of firm j's equity across the network. The model captures the idea of distributing a fixed fraction c of each firm's equity among all the firms in the network in a stochastically unbiased fashion.

In view of the Proposition of Section 7.2, for each j, the variables $C_{1j}^{(n)}$, …, $C_{nj}^{(n)}$ have common mean c/n, are negatively correlated, and are weakly dependent.[3] In view of the fact that the generating distribution is asymptotically diffuse, we can say a little more: for $1 \leq i \leq n$ and $i' \neq i$,

$$\mathbf{E}\bigl(C_{ij}^{(n)}\bigr) = \frac{c}{n},\ \operatorname{Var}\bigl(C_{ij}^{(n)}\bigr) = \mathcal{O}\left(\frac{1}{n^2}\right),\ \bigl|\operatorname{Cov}\bigl(C_{ij}^{(n)}, C_{i'j}^{(n)}\bigr)\bigr| = \mathcal{O}\left(\frac{1}{n^3}\right),$$

$$\mathbf{E}\left(\left|C_{ij}^{(n)} - \frac{1}{n}\right|^{\nu}\right) = \mathcal{O}\left(\frac{1}{n^{\nu}}\right) \qquad (1 \leq \nu \leq 8).$$

Figure 3 illustrates the archetypal setting of Example 1 when the cross-holdings of shares of firm j are engendered by throwing $n-1$ points uniformly at random in the interval $[0, c]$.[4]

7.4. *Empirical Insights from Small Networks*

We summarize some empirical insights gleaned from small random networks that motivate the theoretical analysis.

---

[3] Dependence is technically weak in the sense that the covariance of the random variables is an order of magnitude smaller than their variance. It is qualitatively weak in that the behavior of the model over random networks will show to be similar to its behavior over deterministic networks.

[4] It is worth noting at this juncture that these considerations generalize without difficulty to sparser network models inherited from Erdös–Rényi topologies where equity is distributed only along edges that are present. We will indicate how this can be done in Section 10 but defer a full consideration of this and other extensions for elsewhere.

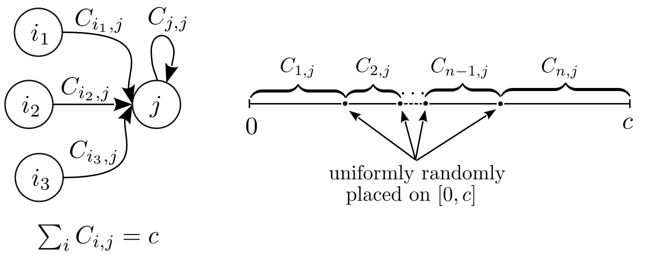

FIGURE 3.—A de Finetti allocation of share equity in which cross-holdings are assigned according to the spacings engendered by throwing $n-1$ points uniformly at random in the interval $[0, c]$. The superscript $^{(n)}$ has been suppressed for legibility.

An enumeration of equilibrium points in random and regular cliques of $n = 16$ firms indicates that their respective sets of equilibrium points behave similarly with changing endowment e. In the notation introduced in Section 6, let $\mathcal{R}_k$ denote the equivalence class of orthants consisting of the regions in which there are exactly k solvent firms and let $\mathbb{K}_1(k), \ldots, \mathbb{K}_{\binom{n}{k}}(k)$ be an enumeration of the distinct orthants in the class. In Fig. 4 the fraction of orthants containing an equilibrium point in a given equivalence class $\mathcal{R}_k$ is reported as a function of k for the regular clique and an instance of a random clique. The endowment e decreases from left to right and top to bottom.

The symmetry of regular cliques means that if any orthant in the equivalence class $\mathcal{R}_k$ contains an equilibrium point, then so do all orthants in the class. In (relatively small) random cliques, however, we often see that, due to the lack of symmetry, some orthants in an equivalence class $\mathcal{R}_k$ may contain equilibria but not others.

In Section 8 we shall see that the setting of small endowments (the bottom three rows in Fig. 4) are of primary interest in evaluating systemic risks in such networks. In such cases random cliques do not admit the full range of equilibria in a given equivalence class as in the corresponding regular clique. This gap typically occurs within a short transient range of values k. One may conjecture that this transient range where there is a gap between the outcomes for the regular clique and a random clique would rapidly diminish as the network's size grows. Fig. 5 and Fig. 6 confirm this though there are visible diminishing returns as the shrinkage of the transient region slows down with increasing n.

Fig. 5 presents a visualization of data for an instantiation of a random clique in a larger network of size $n = 100$. Exhaustive enumeration becomes untenable as network size grows and we have resorted to statistical sampling: a sample of 100 orthants is selected at random in each equivalence class $\mathcal{R}_k$ for $1 \leq k \leq 99$ and each of the selected orthants is tested for the presence of an equilibrium point. Successive graphs in the figure show how the fraction of the orthants containing equilibria in each equivalence class $\mathcal{R}_k$ is affected by decreasing endowment. The transient region where there is a qualitative difference in the equilibria distribution in the classes $\mathcal{R}_k$ in the two cases is shrinking but still visible.

To get a better understanding of the evolution of the gap between the presence of equilibria in the regular clique vis-à-vis a random clique we formalize the notion of the "size" of the transient as the ratio, with respect to n, of the number of equivalence classes where the share of orthants with equilibria differs by ten percent between the regular clique and an instantiation of a random clique. Fig. 6 shows the decrease in the transient region as n increases.

The take away from our simulations is that the random clique, in spite of huge instantiation variability, appears to behave increasingly like the regular clique in terms of the number and disposition of equilibria when n becomes large. This, in fact, is the content of our main theorem.

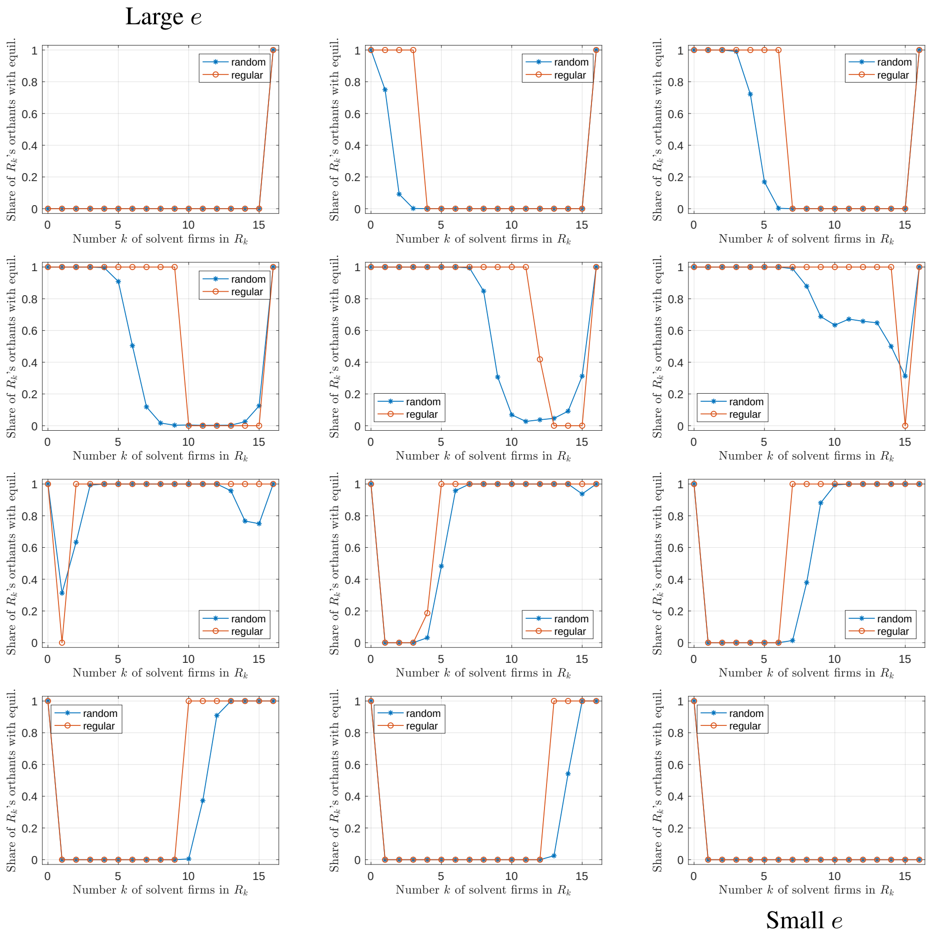

FIGURE 4.—The dependence of the set of equilibrium points on endowment e in random and regular cliques of size $n = 16$. Each figure corresponds to a different value of endowment e, starting with “large” values in the left top corner, and ending with “small” values in the bottom right corner.

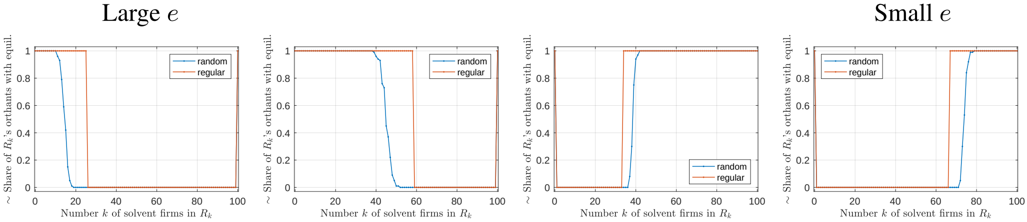

FIGURE 5.—The dependence of the set of equilibrium points on endowment e in random and regular cliques of size $n = 100$. Endowments decrease from left to right.

7.5. *Concentration of Measure in Large Networks*

The insight from our analysis of equilibria in small networks is that, when the network size grows, the set of equilibrium profiles of random clique instances is often similar to the set of

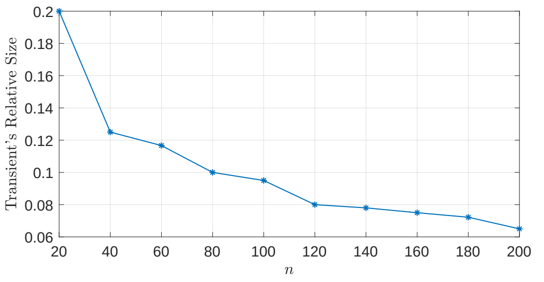

FIGURE 6.—The dependence of the relative size of the transient region where the equilibrium population in a random clique differs from that of the corresponding regular clique as the network size increases.

equilibrium profiles of the regular clique. *Our main theorem in this paper establishes indeed that, in spite of the fact that typical instances of a random clique are highly irregular, for almost all network instances, the equilibria of a random clique do indeed converge to the corresponding equilibria of a regular clique*. In a formal sense, (almost) all roads do lead to Rome.

This is the setting: suppose that the cash holdings (primitive assets, factors of production) of each firm is fixed at a nominal value e and that a fixed fraction c of each firm's equity is distributed across the network (with the residual proportion $1-c$ of each firm's equity held by outside investors).

Fix any positive integer n and suppose $\mathbf{k} \in \{0,1\}^n$ is any solvency-identifying index vector $\mathbf{k} = (k_1, \dots, k_n) \in \{0,1\}^n$; the corresponding orthant $\mathbb{K}^n(\mathbf{k})$ in $\mathbb{R}^n$ consists of those points $\mathbf{x} = (x_1, \dots, x_n)$ satisfying $x_i > \tau$ if $k_i = 1$ and $x_i \leq \tau$ if $k_i = 0$. Consider the case of the regular clique with $n \times n$ cross-holdings matrix $\mathbf{C}^{(0,n)} := \frac{c}{n} \mathbb{1}_{n\times n} = \frac{c}{n} \mathbb{1}_n \mathbb{1}_n^\intercal$ where all the elements are equal to c/n. Theorem 3 now identifies the explicit form of the potential equilibrium solution $\mathbf{V}^{(0,n)}(\mathbf{k})$ in orthant $\mathbb{K}^n(\mathbf{k})$; the solution is feasible if, and only if, the corresponding feasibility constraint in the system of inequalities (8–10) is satisfied.

Now suppose $\mathbf{C}^{(n)}$ is a sequence of random cross-holding matrices generated by an asymptotically diffuse distribution. Write $\mathbf{I}^{(n)}$ for the identity matrix of order n. Then, for any given n, the unique putative equilibrium of the random clique in orthant $\mathbb{K}^n(\mathbf{k})$, if it exists, is given by

$$\mathbf{V}^{(n)}(\mathbf{k}) = \left(\mathbf{I}^{(n)} - \mathbf{C}^{(n)}\right)^{-1} \left((e-\beta)\mathbb{1}^{(n)} + \beta \mathbf{k}\right), \qquad \mathbf{k} \in \{0,1\}^n.$$

The solution will exist if $\mathbf{I}^{(n)} - \mathbf{C}^{(n)}$ is non-singular (which happens almost surely); the solution will be feasible if $\mathbf{V}^{(n)}(\mathbf{k}) \in \mathbb{K}^n(\mathbf{k})$.

The following result vastly strengthens Theorem 4 and is the main theorem of this paper. It asserts that, in a formal sense, the equilibria of almost all instances of a random clique coincide with the corresponding equilibria of a regular clique.

THEOREM 5—Equilibria in Large Random Networks: Full Version [Proof]: *For any sequence of index vectors, $\left\{ \mathbf{k}^{(n)} \in \{0,1\}^n, n \geq 1 \right\}$, we have*

$$\sup_{1\leq i\leq n} \left| V_i^n\left(\mathbf{k}^{(n)}\right) - V_i^{0,n}\left(\mathbf{k}^{(n)}\right) \right| \to 0$$

almost surely as $n \to \infty$.

Introduce the nonce notations $I_1^{(n)} = \left\{ i : k_i^{(n)} = 1 \right\}$ and $I_0^{(n)} = \left\{ i : k_i^{(n)} = 0 \right\}$ for the index sets labelling putatively solvent and insolvent firms, respectively. Recall from (11) that

$$v(s) := \frac{e - (1-s)\beta c}{1-c}.$$

COROLLARY 1: *Suppose* $\frac{1}{n}\left|\mathbf{k}^{(n)}\right| := \frac{1}{n}\left(k_1^{(n)} + \cdots + k_n^{(n)}\right) \to s$ *for some* $0 \le s \le 1$. *Then, for every* $\varepsilon > 0$*,*

$$\sup_{i \in I_1^{(n)}} \left|V_i^n\left(\mathbf{k}^{(n)}\right) - v(s)\right| < \varepsilon \text{ and } \sup_{i \in I_0^{(n)}} \left|V_i^n\left(\mathbf{k}^{(n)}\right) - v(s) - \beta\right| < \varepsilon$$

almost surely as $n \to \infty$*. The putative equilibria* $\mathbf{V}^{(n)}\left(\mathbf{k}^{(n)}\right)$ *are asymptotically feasible for all sufficiently large* n *if, and only if, the cash position satisfies*

$$\begin{aligned}
\tau(1-c) + (1-s)\beta c < e < \tau(1-c) + \beta - s\beta c \qquad & \textit{if } 0 < s < 1,\\
e < \tau(1-c) + \beta \qquad & \textit{if } s = 0,\\
e > \tau(1-c) \qquad & \textit{if } s = 1.
\end{aligned}$$

In rough terms: given a sequence $\left\{ \mathbf{k}^{(n)}, n \ge 1 \right\}$ with $\frac{1}{n}\left|\mathbf{k}^{(n)}\right| \to s$, for almost all instantiations of asymptotically diffuse cross-holdings, the identified asymptotic fraction s of solvent firms have equilibrium valuations ϵ-close to $v(s)$, while the identified asymptotic fraction $1-s$ of insolvent firms have equilibrium valuations ϵ-close to $v(s) - \beta$. Convergence is in the sup-norm—*the equilibrium valuations of all the firms converge simultaneously*.

The theorem relies on subtleties in the concentration of measure and we have assembled the complex proof in stages in three appendices in an effort to keep the key ideas plainly in view: preparatory materials are gathered in Appendices B and C, and the pieces are stitched together in Appendix D.

8. DYNAMICS

We have described the set of equilibria and how they depend on endowment and exposure. To determine which, if any, of these equilibria are stable (do they attract or repel a state $\mathbf{V}_t$ at time t?) we need an associated dynamic. We present the results without adornment here and relegate proofs to Appendix F.

8.1. *The General Setting*

Begin by a consideration of the general form of the model where $\mathbf{e}$ is a vector of firm endowments and $\mathbf{C}$ a fixed matrix of cross-shareholdings. It will be analytically more convenient to work with the continuous-time form (3) of the model reproduced here for ease of reference:

$$\dot{\mathbf{V}}_t = \mathbf{e} - (\mathbf{I} - \mathbf{C})\mathbf{V}_t - \beta \mathbb{1}_{\{\mathbf{V}_t \le \tau \mathbb{1}\}}. \tag{3}$$

As described earlier, the dynamical equation can be interpreted in two ways: as the actual change in the firm valuations as a result of dividend cross-payment and shock absorption; or as a fictitious dynamics that an external party—such as the market—uses to compute the firms' valuations starting with some initial firm market valuation estimates $\mathbf{V}_0$ using locally available information.

Our goal is to understand the form and properties of functions $\mathbf{V}_t$ that satisfy (3). We refer to such functions $\mathbf{V}_t$ as solutions or (state) trajectories of (3).

The dynamical system (3) has discontinuities in its right-hand side at the boundaries of orthants $\mathbb{K}^n(\mathbf{k}) \subset \mathcal{D}$. Thus, we cannot expect the system to have smooth solutions. When the behavior of solutions at the surface of a vector field discontinuity is unknown, it is customary to replace the original dynamical system with a differential inclusion, replace the standard derivative with the corresponding Filippov map, and look for the generalized Filippov solutions $\mathbf{V}_t$ (Filippov, 1988, Ch. 2, Sec. 4). We will understand a solution $\mathbf{V}_t$ as a Carathéodory solution, that is, $\mathbf{V}_t$ is an absolutely continuous function defined for all $t > 0$ and differentiable almost everywhere (see Filippov (1988, Sec. 2.4.1)). Each such solution has no standard derivative at the boundary between orthants $\mathbb{K}^n(\mathbf{k})$. We may select among solutions by imposing the conservative assumption that, at each orthant boundary point, $\mathbf{V}_t$ moves as defined by the right-hand side of (3), that is, in the direction of the adjacent region with the smallest number of solvent firms. The boundedness of the right-hand side of (3) guarantees existence of a solution, while our assumption regarding solution behavior at the region boundary establishes uniqueness [Cortés (2009), Filippov (1988, Sec. 2.7)].

With these conventions in place, with $\mathbf{k}$ any fixed index vector of solvency suppose that the trajectory $\mathbf{V}_t$ is governed by the dynamical system

$$\dot{\mathbf{V}}_t = \mathbf{e} - (\mathbf{I} - \mathbf{C})\mathbf{V}_t - \beta(\mathbb{1} - \mathbf{k}) \qquad (t > 0) \tag{21}$$

for any fixed vector of endowments $\mathbf{e}$ and cross-shareholding matrix $\mathbf{C}$. We begin by a consideration of the plausible assertion that, if the trajectory $\mathbf{V}_t$ never escapes the orthant $\mathbb{K}^n(\mathbf{k})$, then it will converge to a feasible equilibrium in it (if one exists). Eq. (5), reproduced here for convenience, provides an explicit representation for the putative limit point

$$\mathbf{V} = \mathbf{V}(\mathbf{k}) := (\mathbf{I} - \mathbf{C})^{-1}\big(\mathbf{e} - \beta(\mathbb{1} - \mathbf{k})\big).$$

Suppose that the equilibrium $\mathbf{V}(\mathbf{k})$ is feasible and suppose additionally that the initial point $\mathbf{V}_0$ of the trajectory lies in the orthant $\mathbb{K}^n(\mathbf{k})$.

THEOREM 6—Dynamics in Orthants Containing an Equilibrium [Proof]: *If the trajectory $\mathbf{V}_t$ starts in orthant $\mathbb{K}^n(\mathbf{k})$, then, for all $t > 0$ until the trajectory escapes $\mathbb{K}^n(\mathbf{k})$ or, if such an escape never happens, for all $t > 0$,*

$$\mathbf{V}_t = \exp\big(-(\mathbf{I} - \mathbf{C})t\big)\big(\mathbf{V}_0 - \mathbf{V}(\mathbf{k})\big) + \mathbf{V}(\mathbf{k}), \tag{22}$$

and, a fortiori, if the trajectory does not escape $\mathbb{K}^n(\mathbf{k})$, $\lim_{t\to\infty} \mathbf{V}_t = \mathbf{V}(\mathbf{k})$.

For a discussion of the plausibility of the no-escape assumption see Appendix E.

In rough terms, Theorem 6 can be interpreted as saying that, if initial firm valuations $\mathbf{V}_0$ are not too distant from an equilibrium $\mathbf{V}(\mathbf{k})$, then the dynamic will converge to that equilibrium. The theorem is silent about (i) the "typical" firm valuation dynamics in random networks, and (ii) the role of the assumption that the state never exits the orthant under consideration and the complementary question of dynamical flow across orthant boundaries.

To come to a better understanding of these issues we focus, in the remainder of this section, on regular cliques.

8.2. *The Regular Clique*

As before, consider the fixed adjacency matrix $\mathbf{C} = \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$ and fix a common endowment $\mathbf{e} = e\mathbb{1}$. Rewriting (33) for ease of reference, the putative equilibrium $\mathbf{V} = \mathbf{V}^0(\mathbf{k})$ in any given orthant $\mathbb{K}^n(\mathbf{k})$ has an explicit form with components given by

$$V_i^0(\mathbf{k}) = \begin{cases} \frac{e-\left(1-\frac{|\mathbf{k}|}{n}\right)\beta c}{1-c} & \text{if } k_i = 1, \\ \frac{e-\left(1-\frac{|\mathbf{k}|}{n}\right)\beta c}{1-c} - \beta & \text{if } k_i = 0. \end{cases}$$

Introduce the nonce notation $\operatorname{avg}(\mathbf{V}) := \frac{1}{n}\mathbb{1}^\intercal\mathbf{V}$ for the arithmetic mean of the components of $\mathbf{V}$. By specializing Theorem 6 to the regular clique with a common endowment, we obtain an explicit rate of convergence for confined trajectories. As before, suppose that the putative regular clique equilibrium $\mathbf{V}^0(\mathbf{k})$ is feasible for some index vector of solvency $\mathbf{k}$. Suppose additionally that the initial point of the trajectory, $\mathbf{V}_0 = (V_{1,0}, \ldots, V_{n,0})^\intercal$, lies in the orthant $\mathbb{K}^n(\mathbf{k})$, and that the trajectory $\mathbf{V}_t$ is specified by (21) with $\mathbf{C} = \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$ and $\mathbf{e} = e\mathbb{1}$:

$$\dot{\mathbf{V}}_t = e\mathbb{1} - \left(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\right)\mathbf{V}_t - \beta(\mathbb{1} - \mathbf{k}) \qquad (t > 0). \tag{23}$$

THEOREM 7—Dynamics in Orthants Containing an Equilibrium for Regular Cliques [Proof ↗]: *If the trajectory $\mathbf{V}_t = (V_{1,t}, \ldots, V_{n,t})^\intercal$ starts in orthant $\mathbb{K}^n(\mathbf{k})$, then, for all $t > 0$ until the trajectory escapes $\mathbb{K}^n(\mathbf{k})$ or, if such an escape never happens, for all $t > 0$,*

$$\begin{aligned} V_{i,t} = {} & \left[\operatorname{avg}(\mathbf{V}_0) + \beta\left(1 - \tfrac{|\mathbf{k}|}{n}\right) - V_i^0(\mathbf{k})\right] \exp\left(-(1-c)t\right) \\ & - \left[\operatorname{avg}(\mathbf{V}_0) + \beta\left(1 - \tfrac{|\mathbf{k}|}{n}\right) - V_{i,0}\right] \exp(-t) + V_i^0(\mathbf{k}), \end{aligned} \tag{24}$$

and, a fortiori, the trajectory approaches the equilibrium exponentially fast.

Of fundamental importance to the resilience of the regular clique to contagion is the size of the basins of attraction of the "best" and "worst" equilibria, $\mathbf{V}^{\sup}$ and $\mathbf{V}^{\inf}$, respectively. For the regular clique with a common endowment, these extremal equilibria are given in (34) by

$$\mathbf{V}^{\sup} = \mathbf{V}^0(\mathbb{1}) = \frac{e}{1-c}\mathbb{1},$$

$$\mathbf{V}^{\inf} = \mathbf{V}^0(\mathbb{0}) = \frac{e-\beta}{1-c}\mathbb{1}.$$

Simulations suggest that the basins of attraction of the extremal equilibria engulf the corresponding extreme orthants $\mathbb{K}^n(\mathbb{0})$ and $\mathbb{K}^n(\mathbb{1})$, and, in practice, are tangibly larger. For instance, whenever either $\mathbf{V}^{\inf}$ or $\mathbf{V}^{\sup}$ is the only equilibrium, then, unsurprisingly, the entire space is this equilibrium's basin of attraction. This is an immediate consequence of the fact that $\mathbf{V}^{\inf}$ and $\mathbf{V}^{\sup}$ are attractive and unique. Simulations suggest that the basins of attraction continue to be dominant even in intermediate cases—when non-extremal equilibria are present—as well: see Fig. E.2 in Appendix E for the attraction basin diagram in two dimensions.

The latter intermediate case is most interesting, so let us focus on it, assuming that $\tau(1 - c) < e \leq \tau(1-c) + \beta$. In this regime, as per Theorem 3, we are guaranteed that the extremal equilibria $\mathbf{V}^{\inf}$ and $\mathbf{V}^{\sup}$ are present. Let $k_{\text{low}} + 1$ and $k_{\text{high}} - 1$ denote, respectively, the smallest and the largest numbers of solvent firms in any non-extreme equilibrium. Formally, k_{low} is the largest index k for which $e \leq \tau(1-c) + \left(1 - \frac{k}{n}\right)\beta c$, and k_{high} is the smallest index k for which $e > \tau(1-c) + \beta - \frac{k}{n}\beta c$. With k_{low} and k_{high} defined this way, Theorem 3 tells us that all the non-extremal orthants $\mathbb{K}^n(k)$ with either $1 \leq k \leq k_{\text{low}}$ or $k_{\text{high}} \leq k \leq n-1$ contain no equilibria: orthants in the immediate neighborhood of the extremal orthants are equilibria-free.

In this regime, simulations show the following dynamics across orthants. If k_{low} exists, and $\mathbf{V}_0 \in \mathbb{K}^n(k)$ with $k \leq k_{\text{low}}$, then $\lim_t \mathbf{V}_t = \mathbf{V}^{\text{inf}}$. Likewise, if k_{high} exists, and $\mathbf{V}_0 \in \mathbb{K}^n(k)$ with $k \geq k_{\text{high}}$, then $\lim_t \mathbf{V}_t = \mathbf{V}^{\text{sup}}$. Simply put, the orthants with at least k_{high} solvent firms are subsumed into the attraction basin of $\mathbf{V}^{\text{sup}}$, while the orthants with at most k_{low} solvent firms are subsumed into the attraction basin of $\mathbf{V}^{\text{inf}}$. The only observed deviations from this behavior are when the trajectory started in orthants with $k \approx k_{\text{low}}$ or $k \approx k_{\text{high}}$ and accidentally escaped into an orthant with $k_{\text{low}} < k < k_{\text{high}}$ or the other way around—see Appendix E for a discussion of such escapes.

A figure may help clarify the notational potage: the distribution of equilibria and the dynamics of firm valuations are shown schematically in Fig. 7 for intermediate values of endowment e. In this intermediate range, orthants in the immediate neighborhood of the extremal orthants

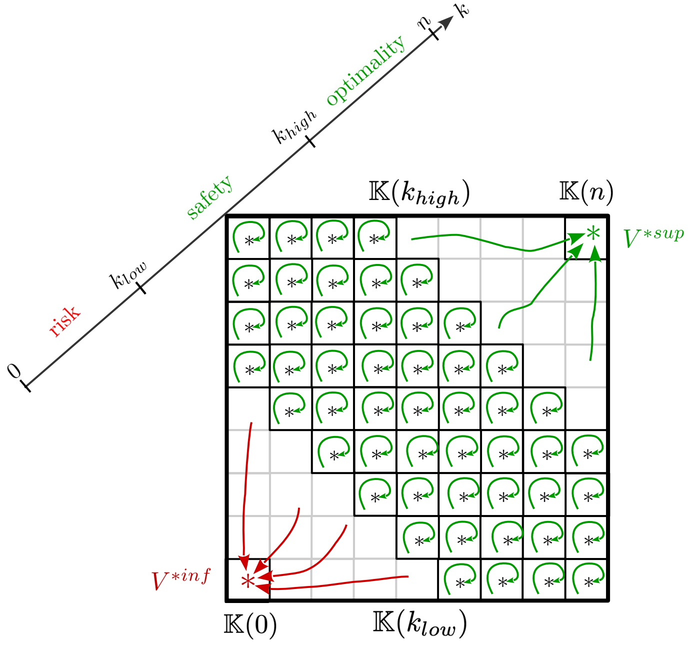

FIGURE 7.—A cartoon depiction of firm valuation dynamics in a regular clique for the case of intermediate cash endowments.

do not contain equilibria. The diagram should not be taken too literally: the sizes of the attraction basins of the two extremal equilibria and the positioning and size of the "safe" zone depend upon the size of endowment and the exposure—at the extreme, one or the other basin of attraction dominates the landscape. Bearing this caveat in mind, we distinguish three qualitatively different zones in the model's state space.

i. *Risky:* Orthants $\mathbb{K}(k)$ with $k \leq k_{\text{low}}$ are "risky" in that, if the initial vector of firm valuations $\mathbf{V}_0$ is in one of these orthants, then oftentimes the trajectory $\mathbf{V}_t$ converges to the worst equilibrium point $\mathbf{V}^{\text{inf}}$, that is, each firm ends up insolvent. Orthants in this regime are the source of systemic risk.

ii. *Optimal:* In contrast, orthants $\mathbb{K}(k)$ with $k \geq k_{\text{high}}$ are "ultimately safe" in that, even if some firms start as insolvent in $\mathbf{V}_0 \in \mathbb{K}(k)$, in typical trajectories every firm ends up solvent with $\mathbf{V}_t$ converging to the best equilibrium point $\mathbf{V}^{\text{sup}}$.

iii. *Safe:* The third part of the state space—"the safety band"—is comprised of the orthants $\mathbb{K}(k)$ with a "middling" number of solvent firms $k_{\text{low}} < k < k_{\text{high}}$. In these orthants, the system is stable, in that, if $\mathbf{V}_0$ belongs to one of these orthants, then $\mathbf{V}_\infty = \lim_t \mathbf{V}_t$ typically stays within the same orthant. That is to say, in this region the initial firm valuations are already consistent in terms of each firm's (in)solvency.

The figure provides a sanitized view of the real picture—trajectory escapes can occur at the boundary from one region to an adjacent region occasionally. Appendix E contains a discussion of trajectory escape, with Fig. E.4 providing a more nuanced view of the state of affairs illustrated in Fig. 7 adjusted for the possibility of trajectory escape into nearby orthants. In short, a trajectory that begins in the edges of the safety band may remain in the safety band or be attracted by the respective extremal equilibrium.

In conjunction with our concentration result, Theorem 5, we now have reason to believe that, for most instantiations of random, fully diversified cross-shareholding networks: trajectories in the interior of orthants containing equilibria will converge (exponentially) quickly to them in a safe zone of orthants; the basin of attraction of the best equilibrium $\mathbf{V}^{\text{sup}}$ grows as endowment increases, expanding the size of the optimal zone; and, conversely, the basin of attraction of the worst equilibrium $\mathbf{V}^{\text{inf}}$ increases as endowment shrinks, increasing the size of the risky zone. We provide additional empirical evidence in Appendix E. These insights shed light on the response of the system to shocks.

9. SHOCKS

9.1. *Shock to Firm Market Value*

What happens if the system experiences a shock which takes the form of one or more firms experiencing a sudden dip in market valuation and falling below the solvency threshold τ? Concentration of measure suggests that, at least for large networks, the state of affairs captured in Fig. 7 for the regular clique tells the story for most network instantiations.

Reset the time origin to the point of shock and suppose that a market valuation shock is manifested in an initial state $\mathbf{V}_0$ where one or more firms is rendered insolvent. We are interested in the limiting time horizon of market valuations $\mathbf{V}_\infty = \lim_t \mathbf{V}_t$ in the corresponding continuous-time dynamical system (3) with $\mathbf{C} = \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$ and $\mathbf{e} = e\mathbb{1}$. In this setting, two bookend definitions capture the extremes of what a tolerance to shock may entail.

Guaranteeing full recovery: all firms eventually recover to solvency.

Avoiding total collapse: at least one firm is eventually solvent.

Full recovery requires $\mathbf{V}_0$ to be in the basin of attraction of $\mathbf{V}^{\text{sup}}$, strictly above the "safety band" in the schematic state space diagram shown in Fig. 7, while avoiding total collapse requires $\mathbf{V}_0$ to lie either within or above the "safety band," that is to say, *outside* the attraction basin of $\mathbf{V}^{\text{inf}}$.

Suppose that the market valuation shock at time $t = 0$ has resulted in an initial state of displaced firm valuations, $\mathbf{V}_0$, where precisely k firms are solvent. We have now honed in on a precise formulation of tolerance to valuation shock: *what conditions on the minimal endowment guarantee full recovery or, at least, avoid total collapse*? The feasibility condition (8) in Theorem 3 provides a succinct answer:

- *The system recovers fully if*

$$e > \tau(1-c) + \beta - \tfrac{k}{n}\beta c =: \overline{e}.$$

- *The system avoids total collapse if*

$$e > \tau(1-c) + \big(1 - \tfrac{k}{n}\big)\beta c =: \underline{e}.$$

The apparently non-monotonic dependence of the minimal endowments $\overline{e}$ and $\underline{e}$ on exposure c is illusory. As pointed out in Elliott, Golub and Jackson (2014, p. 9, eq. (4)), it is preferable to cast insolvency in terms of *market values* via the market value insolvency threshold τ'. Our book value insolvency threshold τ is related to the *fixed* market value insolvency threshold τ' via the simple scaling relation $\tau = \tau'/(1-c)$ [see comment iii on page 7]: if a firm defaults at some profile of book values in our setting, it will also do so with respect to its market value, and vice versa. Accordingly, the minimal endowments to guarantee full recovery and avoid total collapse, respectively, are given in terms of the market value insolvency threshold τ' by

$$\begin{aligned} \overline{e} &= \tau' + \beta - \tfrac{k}{n}\beta c, \\ \underline{e} &= \tau' + \big(1 - \tfrac{k}{n}\big)\beta c. \end{aligned} \tag{25}$$

With τ', β, k, and n viewed as fixed parameters, the expressions for the minimal endowments needed to tolerate a valuation shock of the given size (measured in terms of the number $n-k$ of firms whose valuations have dropped below the insolvency threshold) both vary monotonically with exposure, with $\overline{e}$ decreasing and $\underline{e}$ increasing as c increases from 0 to 1. The gap between the minimal endowments in the two cases, $\overline{e} - \underline{e} = \beta(1-c)$, is independent of the market value insolvency threshold τ' and shrinks to zero as c increases to one. Fig. 8 shows how the minimal endowments in the two cases behave as a function of c and the number $n-k$ of failures that can be tolerated.

The interplay between endowment, exposure, and shock tolerance may be further clarified by reinterpreting (25) in terms of the basins of attraction arising from a given endowment e and exposure c. As in the previous section, let k_{low} denote the largest index k for which $e \le \tau' + \big(1 - \tfrac{k}{n}\big)\beta c$, and let k_{high} denote the smallest index k for which $e > \tau' + \beta - \tfrac{k}{n}\beta c$. In formalism:

$$\begin{aligned} k_{\text{low}} &:= \big\lfloor (\tau' + \beta c - e)\tfrac{n}{\beta c} \big\rfloor, \\ k_{\text{high}} &:= \big\lfloor (\tau' + \beta - e)\tfrac{n}{\beta c} \big\rfloor + 1. \end{aligned} \tag{26}$$

With $0 < c < 1$, consider the case of intermediate endowments $\tau' < e \le \tau' + \beta$.[5] We may now interpret the three generic zones in Fig. 7 in relation to the response to a valuation shock:

i. *The risky zone*: if $k \le k_{\text{low}}$ then the system collapses.

ii. *The optimal zone*: if $k \ge k_{\text{high}}$ then the system recovers fully.

iii. *The safe zone*: if $k_{\text{low}} < k < k_{\text{high}}$ then the system avoids total collapse but does not recover fully.

[5]The extreme cases $e \le \tau'$ and $e > \tau' + \beta$ are trivial: in the first case the worst equilibrium is the sole attractor and systemic collapse will occur whatever the starting point; in the second case the best equilibrium is the sole attractor and the system recovers fully even if every firm suffers an insolvency shock.

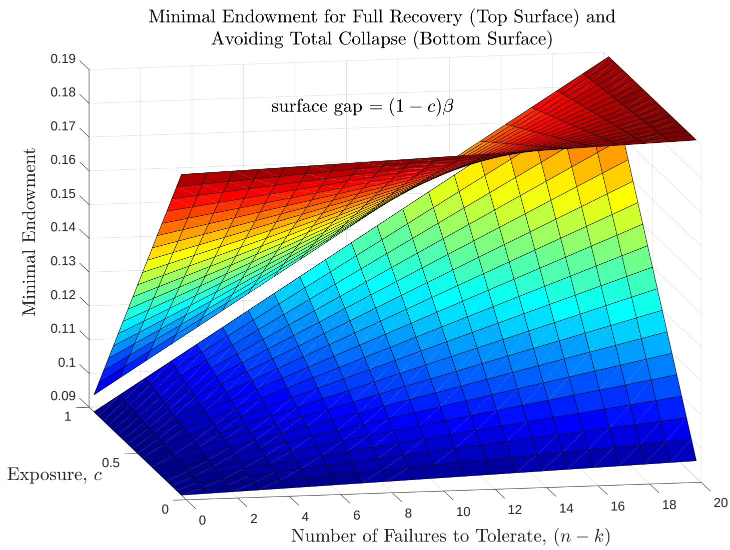

FIGURE 8.—The minimal endowments $\overline{e}$ and $\underline{e}$ requisite for full recovery and avoiding total collapse, respectively, as a function of the exposure c and the number $n-k$ of firm failures.

All three statements have to be hedged by caveats because sporadic (and empirically rare) trajectory escapes do occur at the boundaries of these regions.

There are three lessons that we may now infer from the explicit expressions for k_{low} and k_{high} given in (26). First, k_{high} decreases with both e and c and so *the basin of attraction of the best equilibrium increases monotonically with both endowment and exposure*. Second, k_{low} decreases with e and increases with c, whence *the basin of attraction of the worst equilibrium decreases with endowment but increases with exposure*. And third, ignoring integer round-off, as $k_{\text{high}} - k_{\text{low}} = \frac{n}{c}(1-c)$, *the width of the safety zone is invariant with respect to endowment and decreases with exposure*.

Increasing endowment provides untrammeled benefits to shock tolerance—the minimal endowments needed to provide specified levels of tolerance are given in (25). Increasing exposure also provides benefits but the relation of tolerance to exposure is slightly more nuanced.

When exposure is high (Fig. 9) orthants within the thin safety band are all susceptible to trajectory escapes, making them vulnerable to collapse, *but* the thinning of the safe zone with increasing exposure is compensated for by an increase in the attraction basin of the best equilibrium.

When exposure is low (Fig. 10), the "safety band" in the state space is large and improves resilience against systemic risk (total collapse) by providing a stable buffer between the attraction basins of the extreme equilibria, *but*, at the same time, reduces the likelihood of total recovery for even a moderate number of firm failures. Unsurprisingly, for low exposure, the system is inert, tending to preserve firms' (in)solvency.

9.2. *Shock to Endowment*

To examine how sensitive a firm's valuation is to a change in the endowment of another firm we need a version of Theorem 3 where endowments vary by firm. It is stated below without

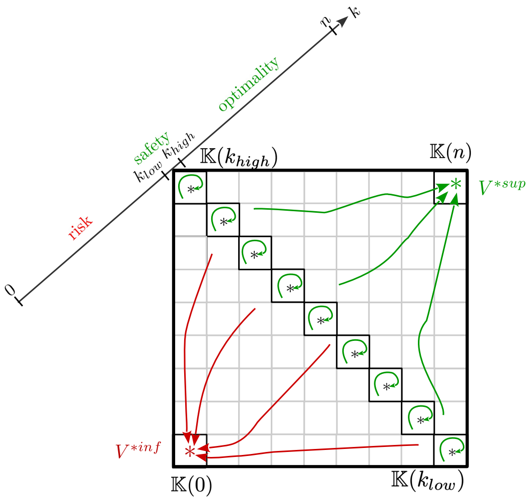

FIGURE 9.—A state space diagram of equilibrium distribution under large exposure c. The "safety band" of equilibria is very narrow, and most of the state space is shared among the massive attraction basins of the extreme equilibria $\mathbf{V}^{*sup}$ and $\mathbf{V}^{*inf}$.

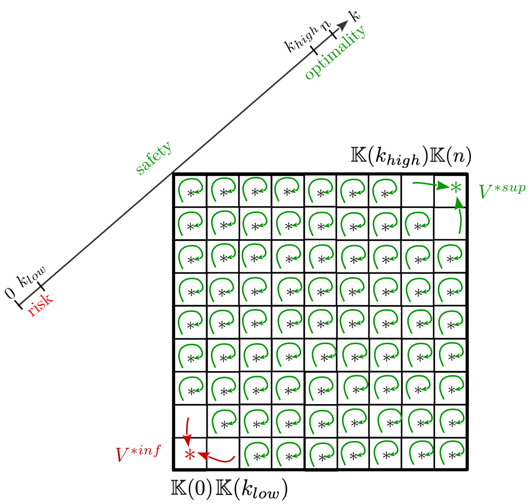

FIGURE 10.—A state space diagram of equilibrium distribution under small exposure c. The "safety band" of equilibria is thick, insulating the states within from the risky region—the attraction basin of $\mathbf{V}^{\text{inf}}$—yet, at the same time, impeding full recovery due to a similar shrinkage of the attraction basin of $\mathbf{V}^{\text{sup}}$.

proof. Let e_i be the endowment of firm i and set

$$v_i(s) := e_i + \frac{c\sum_{i=1}^{n} e_i}{n(1-c)} - \frac{(1-s)\beta c}{1-c} \qquad (0 \leq s \leq 1). \tag{27}$$

THEOREM 8: *For any given* $\mathbf{k} \in \{0,1\}^n$*, the fixed point equation* (7) *has a unique solution* $\mathbf{V}^0(\mathbf{k})$ *with components given by*

$$V_i^0(\mathbf{k}) = \begin{cases} v_i\left(\frac{|\mathbf{k}|}{n}\right) & \textit{if } k_i = 1, \\ v_i\left(\frac{|\mathbf{k}|}{n}\right) - \beta & \textit{if } k_i = 0, \end{cases} \tag{28}$$

Suppose e_1 declines by $\delta > 0$, i.e., firm 1 suffers a negative shock to its endowment. From (27) it is easy to see that the change in $v_i(s)$ for any $i \neq 1$ is $c\delta/n(1-c)$. Hence, the impact on other firms declines with n holding c fixed, and increases with c holding n fixed.

Suppose all firms have the same endowment e, and the only equilibrium is one where all firms are solvent. From (10) we know that this requires $e > \tau' + \beta$ where we recall that the fixed market value solvency threshold τ' is related to the book value solvency threshold τ via the relation $\tau' = \tau(1-c)$. Suppose a negative shock to firm 1's endowment. How large must that shock be to summon into existence an equilibrium with a large number of insolvent firms?

If firm 1 alone suffers a negative shock of δ to its endowment, we determine the smallest δ in order to generate an equilibrium where r firms (including firm 1) are insolvent. Without loss we may suppose the first r firms are insolvent. The relevant putative equilibrium must satisfy $v_i(\frac{r}{n}) \leq \tau$ for $i = 1, \dots, r$, and $v_i(\frac{r}{n}) > \tau$ otherwise. For $i = 1$ we have

$$-\delta + \frac{e}{1-c} - \frac{c\delta}{n(1-c)} - \frac{(1-\frac{r}{n})\beta c}{1-c} \leq \tau.$$

For $i = 2, \dots, r$, we have

$$\frac{e}{1-c} - \frac{c\delta}{n(1-c)} - \frac{(1-\frac{r}{n})\beta c}{1-c} \leq \tau.$$

Clearly, the second inequality gives us the greatest lower bound on δ:

$$\begin{aligned} \delta &\geq \frac{n(1-c)}{c}\left[\frac{e-(1-\frac{r}{n})\beta c}{1-c} - \tau\right] \\ &= \frac{ne-(n-r)\beta c}{c} - \frac{n\tau'}{c} = \frac{n(e-\tau')}{c} - (n-r)\beta. \end{aligned} \tag{29}$$

If r scales with n, as $e > \tau'$, it follows that the size of the shock to firm 1's endowment needed to introduce an equilibrium with a large fraction of defaulting firms scales with n. In other words, the initial shock must be many multiples of what is needed to push a firm's endowment below τ'. Under limited liability this cannot happen.

10. RANDOM GRAPHS & DIVERSIFICATION

A focus of earlier work has been on the role that diversification, measured by the number of firms that own a given firm's shares, plays in spreading contagion. Elliott, Golub and Jackson (2014), for instance, argue that diversification at an "intermediate" level makes networks susceptible to contagion. This non-monotonicity, as we have previously remarked, is absent in our model.

We have hitherto considered networks where the underlying graph is fully connected: the outgoing edge weights are specified by an exchangeable process and, in general, are almost surely all positive. In this model, the size of endowment and the level of exposure are the key parameters governing behavior. As we've seen, in this setting an increase in endowment increases tolerance to shocks. It is natural to ask at this juncture what role diversification can play in the picture.

A natural variant on the theme that we have considered is to impose an Erdös–Rényi graph topology as a superstructure in our model. The idea is as follows. Begin by engendering a random digraph $\mathcal{G}(n,p)$ on n vertices with edge parameter p: for each ordered pair (i,j), insert an edge directed from i to j with probability p, independent of all other ordered pairs; self-loops are permitted. The outdegree of any given vertex j determines its connected neighborhood: these are the firms that are entitled to hold shares in firm j's equity. On average, the number of firms holding a given firm's shares will be np and we can think of p as a diversification parameter. If, for a given exposure c, we allocate j's shares equally to all its neighbors, we obtain the Erdös–Rényi equivalent of the regular clique. If j's shares are allocated to its neighbors via an exchangeable process we obtain the natural analog of our random networks embedded in a stochastic topological superstructure.

Our analysis carries through *in toto* for these model extensions. Our basic conclusions: once n is sufficiently large for a given exposure c and diversification parameter p (roughly, $n > (1-c)^{-4}p^{-1}$), almost all network realizations will behave like the regular clique. Systemic risk is now entirely determined by e and c as in this paper. Simulations provide additional support for the theoretical conclusion that there is no sensitivity to diversification, even for very small p.

These results carry over to the stochastic block cousins of the Erdös–Rényi models though we have to be careful with the asymptotics of block sizes. We leave details of this and the underlying Erdös–Rényi variants to a subsequent paper.

In these random graph extensions to our model we work resolutely in the connected[6] domain—it would be strange to begin an analysis of network systemic risk abeyant connectivity. In the Erdös–Rényi formulation, for instance, this means that $p \gg \log(n)/n$, the critical threshold for connectedness. The concentration of measure phenomenon emerges anew in this domain and almost all network instantiations have similar equilibrium behavior. In this setting we find a monotone relationship between endowment and exposure on the one hand, and tolerance to shock on the other; diversification appears to play no role in systemic risk once we are safely in the connected Erdös–Rényi regime.

As mentioned earlier, the "intermediate level" of integration identified in Elliott, Golub and Jackson (2014) at which non-monotonic behavior is manifested appears to occur when the underlying network is *very* sparse, that is to say, network topologies potentially susceptible to contagion are characterized by bounded, typically small, vertex degrees. If such networks are sampled from, say, the family of d-regular networks where $d = np = \mathcal{O}(1)$ is small and the graph is large, then most network instances are connected but the channels by which contagion can spread are severely limited; if such networks are sampled from an Erdös–Rényi topology with small expected degree this will lead to disconnected networks with many isolated components. This class of networks of low density lies at the boundary of the spectrum of possible network realizations and it is not clear if our techniques are effective in this domain.

[6]A directed graph is connected if for every pair of vertices there is a directed path from one vertex to the other. It is strongly connected if there is a directed cycle between any pair of vertices.

11. CONCLUSION

Rather than rehash our introduction in fewer words, we conclude by emphasizing the methodological contribution of this paper. Specifically, it is to think about network structure as encoded in a probability distribution over networks rather than the features of realized networks. As we have shown here, this allows one to draw conclusions about the behavior of a wide range of networks that by the usual indicia of network structure appear very different.

ACKNOWLEDGEMENTS

We thank Benjamin Brooks, Krishna Dasaratha, George Mailath, Roger Myerson and Eduard Talamas for useful comments.

REFERENCES

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2015. "Systemic risk and stability in financial networks." *American Economic Review*, 105(2): 564–608.

Cortés, Jorge. 2009. "Discontinuous Dynamical Systems: A tutorial on solutions, nonsmooth analysis, and stability." *arXiv:0901.3583 [math.DS].*

Dasaratha, Krishna. 2020. "Distributions of centrality on networks." *Games and Economic Behavior*, 122: 1–27.

Eisenberg, Larry, and Thomas H Noe. 2001. "Systemic risk in financial systems." *Management Science*, 47(2): 236–249.

Elliott, Matthew, Benjamin Golub, and Matthew O Jackson. 2014. "Financial networks and contagion." *American Economic Review*, 104(10): 3115–53.

Filippov, Aleksei F. 1988. *Differential Equations with Discontinuous Righthand Sides.* Vol. 18 of *Mathematics and Its Applications*, Kluwer Academic Publishers.

Gershgorin, Semyon A. 1931. "Uber die Abgrenzung der Eigenwerte einer Matrix." *Bulletin of the USSR Academy of Sciences, Mathematical and Natural Sciences Section*, , (6): 749–754.

Horn, Roger A, and Charles R Johnson. 2012. *Matrix Analysis.* Cambridge University Press.

Jackson, Matthew O, and Agathe Pernoud. 2019. "What Makes Financial Markets Special? Systemic Risk and Its Measurement in Financial Networks." *Systemic Risk and Its Measurement in Financial Networks (January 7, 2019).*

Shi, Yu, Robert Townsend, and Wu Zhu. 2019. "Internal capital markets in business groups and the propagation of credit supply shocks." International Monetary Fund Working Paper WP/19/111.

Tarski, Alfred. 1955. "A lattice-theoretical fixpoint theorem and its applications." *Pacific journal of Mathematics*, 5(2): 285–309.

Tropp, Joel A. 2015. *An Introduction to Matrix Concentration Inequalities.* arXiv:1501.01571v1.

Venkatesh, Santosh S. 2013. *The Theory of Probability: Explorations and Applications.* Cambridge University Press.

APPENDIX A: PROOFS OF THEOREMS 1 AND 2

We begin by verifying that (4) is a bona fide solution of the fixed point equation (1). This follows quickly via the Gershgorin disc theorem (c.f. (Horn and Johnson, 2012)): suppose $\mathbf{A} = [a_{ij}]_{i,j=1}^{n}$ is a complex matrix of order n. For each j, let $r_j = \sum_{i \neq j} |a_{ij}|$ and let the *Gershgorin disc* $D(a_{jj}, r_j)$ be the closed disc of radius r_j centered at the point a_{jj}.

THEOREM [GERSHGORIN (1931)]: *Every eigenvalue of* $\mathbf{A}$ *lies in the union of the Gershgorin discs* $D(a_{jj}, r_j)$.

Recall that the matrix of cross-holdings $\mathbf{C}$ is column sub-stochastic: $0 < \sum_i C_{ij} < 1$ for each j. Setting $\widetilde{c} = \max_j \sum_i C_{ij} = \max_j \big[C_{jj} + \sum_{i \neq j} C_{ij}\big]$, we see that $0 < \widetilde{c} < 1$.

LEMMA 1: *The matrix* $\mathbf{I} - \mathbf{C}$ *is non-singular.*

PROOF: By the Gershgorin disc theorem, each eigenvalue of $\mathbf{C}$ lies within the union of the discs $D\big(C_{jj}, \sum_{i\neq j} C_{ij}\big)$. It follows that the spectrum of $\mathbf{C}$ is constrained within the closed disc $D\big(0, \widetilde{c}\big)$ of radius $\widetilde{c} < 1$ centered at the origin. By symmetry, this is also true for the spectrum of $-\mathbf{C}$. The addition of the identity matrix right shifts the spectrum by 1 and so the spectrum of $\mathbf{I} - \mathbf{C}$ is contained within the disc $D\big(1, \widetilde{c}\big)$. In particular,

$$\min_{\lambda \in \operatorname{spec}(\mathbf{I}-\mathbf{C})} \operatorname{Re}(\lambda) \geq 1 - \widetilde{c} > 0.$$

We conclude that $0 \notin \operatorname{spec}(\mathbf{I} - \mathbf{C})$ or, what is the same thing, $\mathbf{I} - \mathbf{C}$ is non-singular. ▶

The proof of Theorem 1 devolves from the Knaster–Tarski theorem (Tarski, 1955, Theorem 1, (i-iii)). Begin with the setting: a pair $(X, \preccurlyeq)$ consisting of a non-empty set X and a binary relation $\preccurlyeq$ is a *partially-ordered set* (or *poset*) if $\preccurlyeq$ is reflexive, antisymmetric, and transitive on X. An upper bound of a subset $S \subseteq X$ is an element of X such that all members of S are $\preccurlyeq$ it. The least upper bound or *supremum* of S is $\preccurlyeq$ all upper bounds of S. Lower bounds and the greatest lower bound or *infimum* are defined analogously. A *lattice* $(X, \vee, \wedge)$ is a poset $(X, \preccurlyeq)$ any two elements $a, b \in X$ of which have a unique supremum (or *join*) $a \vee b$ and a unique infimum (or *meet*) $a \wedge b$. A lattice is *complete* if each of its subsets has a supremum and an infimum, both belonging to the lattice. A function $u \colon A \to B$ is *order-preserving* with respect to the partial order $\preccurlyeq$ if $f(x) \preccurlyeq f(y)$ for any elements $x, y \in A$ satisfying $x \preccurlyeq y$.

THEOREM [KNASTER–TARSKI (TARSKI, 1955, THEOREM 1, (I-III))]: *Let $(L, \preccurlyeq)$ be a non-empty complete lattice. For every order-preserving function $f \colon L \to L$, the set of its fixed points with the ordering induced by $\preccurlyeq$ forms a non-empty complete lattice.*

THEOREM 1 (Equilibrium Existence): *The fixed point equation (1) has at least one solution.*

PROOF: Consider the function $f \colon \mathcal{D} \to \mathcal{D}$ defined by $f(\mathbf{V}) = \mathbf{e} + \mathbf{C}\mathbf{V} - \beta \mathbb{1}_{\{\mathbf{V} \leq \tau \mathbb{1}\}}$ in a sufficiently large bounded box $\mathcal{D} \subset \mathbb{R}^n$. Begin with the observation that the domain $\mathcal{D}$ together with the elementwise vector operator $\leq$ forms a complete lattice. As $\mathcal{D}$ is a bounded box, the elementwise $\max$ and $\min$ functions deliver subset suprema and infima, respectively, with each infimum and supremum also a part of $\mathcal{D}$. It is now easy to verify that f is order-preserving: suppose $\mathbf{U}$ and $\mathbf{V}$ are elements of $\mathcal{D}$ and $\mathbf{U} \leq \mathbf{V}$. Then $\mathbb{1}_{\{\mathbf{V} \leq \tau \mathbb{1}\}} \leq \mathbb{1}_{\{\mathbf{U} \leq \tau \mathbb{1}\}}$ and, as the elements of $\mathbf{C}$ are all non-negative, $\mathbf{C}\mathbf{U} \leq \mathbf{C}\mathbf{V}$. We conclude that

$$f(\mathbf{U}) = \mathbf{e} + \mathbf{C}\mathbf{U} - \beta \mathbb{1}_{\{\mathbf{U} \leq \tau \mathbb{1}\}} \leq \mathbf{e} + \mathbf{C}\mathbf{V} - \beta \mathbb{1}_{\{\mathbf{V} \leq \tau \mathbb{1}\}} = f(\mathbf{V}).$$

By the Knaster–Tarski theorem, the set of fixed points of f forms a non-empty complete lattice in $\mathcal{D}$. We conclude that (1) has at least one equilibrium point. ▶

THEOREM 2 (Extremal Fixed Points): *Suppose equal cash endowments $\mathbf{e} = e\mathbb{1}$ and that every firm i holds a positive amount of equity shares, that is to say, $C_{ij} > 0$ for at least one j. Then:*

a) *For $\mathbf{V}^{\text{sup}}$ to be an equilibrium it is necessary that $e > \tau \cdot \sigma_{\min}(\mathbf{I} - \mathbf{C})$; for it to be an equilibrium it is sufficient that $e > \tau$.*
b) *For $\mathbf{V}^{\text{inf}}$ to be an equilibrium it is necessary that $e \leq \tau + \beta$; for it to be an equilibrium it is sufficient that $e \leq \tau \cdot \frac{\sigma_{\min}(\mathbf{I}-\mathbf{C})}{\sqrt{n}} + \beta$.*

PROOF: a) *Necessary and sufficient conditions for* $\mathbf{V}^{\text{sup}}$ *to be an equilibrium*. We begin with the sufficient condition. Suppose $e > \tau$. Then

$$\mathbf{V}^{\text{sup}} = e(\mathbf{I}-\mathbf{C})^{-1}\mathbb{1} = e\sum_{k=0}^{\infty}\mathbf{C}^k\mathbb{1} = e\mathbb{1} + e\mathbf{C}\mathbb{1} + \sum_{k=2}^{\infty}\mathbf{C}^k\mathbb{1} \geq e\mathbb{1} + \sum_{k=2}^{\infty}\mathbf{C}^k\mathbb{1} \geq e\mathbb{1} > \tau\mathbb{1},$$

the matrix geometric series converging by virtue of the fact that $\mathbf{C}$ has non-negative elements bounded by one. We conclude that $\mathbf{V}^{\text{sup}} \in \mathbb{K}^n(\mathbb{1})$ and hence that it is feasible.

To establish a necessary condition, observe that $\mathbf{V}^{\text{sup}}$ is feasible if, and only if, $e(\mathbf{I}-\mathbf{C})^{-1}\mathbb{1} > \tau\mathbb{1}$. By the monotonicity of the Euclidean norm, by taking the norm of both sides, we conclude that, if $\mathbf{V}^{\text{sup}}$ is feasible, then

$$e > \tau \cdot \frac{\|\mathbb{1}\|_2}{\left\|(\mathbf{I}-\mathbf{C})^{-1}\mathbb{1}\right\|_2} \overset{(\dagger)}{\geq} \tau \cdot \frac{1}{\left\|(\mathbf{I}-\mathbf{C})^{-1}\right\|_2} = \tau \cdot \frac{1}{\sigma_{\max}\left((\mathbf{I}-\mathbf{C})^{-1}\right)} = \tau \cdot \sigma_{\min}(\mathbf{I}-\mathbf{C})$$

establishing necessity.[7]

b) *Necessary and sufficient conditions for* $\mathbf{V}^{\text{inf}}$ *to be an equilibrium*. Begin with the necessary condition. Suppose $e > \tau + \beta$. The positivity of $\mathbf{C}$ shows again that

$$\begin{aligned}\mathbf{V}^{\text{inf}} = (e-\beta)(\mathbf{I}-\mathbf{C})^{-1}\mathbb{1} = (e-\beta)\sum_{k=0}^{\infty}\mathbf{C}^k\mathbb{1}& \\ = (e-\beta)\mathbb{1} + \sum_{k=1}^{\infty}\mathbf{C}^k\mathbb{1} > \tau\mathbb{1} + \sum_{k=1}^{\infty}\mathbf{C}^k\mathbb{1} > \tau\mathbb{1},&\end{aligned}$$

whence it is necessary that $e \leq \tau + \beta$ if $\mathbf{V}^{\text{inf}}$ is to be feasible.

To establish a sufficient condition, for $\mathbf{V}^{\text{inf}}$ to be feasible in the orthant $\mathbb{K}^n(\mathbb{0})$, we must have $(e-\beta)(\mathbf{I}-\mathbf{C})^{-1}\mathbb{1} \leq \tau\mathbb{1}$. The inequality is tritely satisfied when $e \leq \beta$. Accordingly, suppose $e > \beta$ whence the inequality may be rewritten as

$$(\mathbf{I}-\mathbf{C})^{-1}\mathbb{1} \leq \frac{\tau}{e-\beta}\mathbb{1}.$$

As the vector inequality holds component-wise it will be satisfied *a fortiori* if

$$\left\|(\mathbf{I}-\mathbf{C})^{-1}\mathbb{1}\right\|_\infty \leq \frac{\tau}{e-\beta}. \tag{30}$$

The matrix sup-norm is consistent (by definition) with the inducing vector sup-norm, and as $\frac{1}{\sqrt{n}}\|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_2$ for any square matrix of order n (c.f. Horn and Johnson (2012, Chapter 5)), we may bound

$$\begin{aligned}\left\|(\mathbf{I}-\mathbf{C})^{-1}\mathbb{1}\right\|_\infty \leq \left\|(\mathbf{I}-\mathbf{C})^{-1}\right\|_\infty \cdot \|\mathbb{1}\|_\infty = \left\|(\mathbf{I}-\mathbf{C})^{-1}\right\|_\infty& \\ \leq \sqrt{n}\cdot\left\|(\mathbf{I}-\mathbf{C})^{-1}\right\|_2. = \sqrt{n}\,\sigma_{\max}\left((\mathbf{I}-\mathbf{C})^{-1}\right) = \frac{\sqrt{n}}{\sigma_{\min}(\mathbf{I}-\mathbf{C})}.&\end{aligned} \tag{31}$$

[7]The bound is actually tight as the step marked $(\dagger)$ is satisfied with equality for the regular clique $\mathbf{C} = \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$. This follows from the feasibility constraint (10) and Theorem 3, coupled with Lemma 2d from which we can deduce that $\sigma_{\min}\left(\mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal\right) = 1 - c$.

It hence suffices if the bound on the right in (30) dominates the bound on the right in (31): if $e \le \tau \cdot \frac{\sigma_{\min}(\mathbf{I}-\mathbf{C})}{\sqrt{n}} + \beta$, then

$$\frac{\tau}{e-\beta} \ge \frac{\sqrt{n}}{\sigma_{\min}(\mathbf{I}-\mathbf{C})} \ge \left\|(\mathbf{I}-\mathbf{C})^{-1}\right\|_\infty,$$

completing the proof of sufficiency. ▶

APPENDIX B: PROOF OF THEOREM 3

Suppose n is any fixed positive integer lurking in the background. For each real α, define

$$\mathbf{D}(\alpha) := \mathbf{I} + \alpha \mathbb{1}\mathbb{1}^\intercal. \tag{32}$$

For each α, $\mathbf{D}(\alpha) = \mathbf{D}_n(\alpha)$ is a symmetric matrix of order n whose diagonal elements are equal to $1+\alpha$ and whose off-diagonal elements are equal to α. We say that such a matrix is *almost diagonal*[8]. With the dimensionality parameter n fixed, we will suppress the dependence on n to keep notation uncluttered.

We begin by collecting some preliminary facts about almost diagonal matrices. These will have utility elsewhere as well.

LEMMA 2: *Almost diagonal matrices satisfy the following identities and properties.*

a) *Inverses: if* $n\alpha \ne -1$, *then* $\mathbf{D}(\alpha)^{-1} = \mathbf{D}\left(\frac{-\alpha}{n\alpha+1}\right)$; *the matrix* $\mathbf{D}\left(\frac{-1}{n}\right)$ *is singular.*
b) *Square roots: if* $n\alpha > -1$, *then* $\mathbf{D}(\alpha)^{1/2} = \mathbf{D}\left\{\frac{1}{n}\left(\sqrt{n\alpha+1}-1\right)\right\}$.
c) *Spectrum:* $\operatorname{spec}\mathbf{D}(\alpha) \in \{1, n\alpha+1\}$. *If* $\alpha \ne 0$, *the eigenvalue* 1 *has multiplicity* $n-1$ *and we may select as eigenvectors* $\left(-1, (\mathbf{I}_{n-1})_{1*}\right)^\intercal, \dots, \left(-1, (\mathbf{I}_{n-1})_{n-1,*}\right)^\intercal$; *the eigenvalue* $\alpha n+1$ *has multiplicity* 1 *with corresponding eigenvector* $\mathbb{1}$.
d) *Spectral norm:* $\left\|\mathbf{D}(\alpha)\right\|_2 = \max\{1, |n\alpha+1|\}$.

PROOF: a, b) To verify the given identities for the inverse and the square root, we need to show that

$$\mathbf{D}(\alpha)\cdot\mathbf{D}\left(\frac{-\alpha}{n\alpha+1}\right) = \mathbf{I},$$

$$\mathbf{D}\left\{\tfrac{1}{n}\left(\sqrt{n\alpha+1}-1\right)\right\}\cdot\mathbf{D}\left\{\tfrac{1}{n}\left(\sqrt{n\alpha+1}-1\right)\right\} = \mathbf{D}(\alpha).$$

As $\mathbb{1}^\intercal\mathbb{1} = n$, it only needs algebraic simplification to check for the two cases, respectively, that

$$\alpha + \frac{-\alpha}{n\alpha+1} + \frac{-\alpha^2}{n\alpha+1}\cdot n = 0,$$

$$\frac{\sqrt{n\alpha+1}-1}{n} + \frac{\sqrt{n\alpha+1}-1}{n} + \left(\frac{\sqrt{n\alpha+1}-1}{n}\right)^2\cdot n = \alpha.$$

The verification that $\mathbf{D}\left(\frac{-1}{n}\right)$ is singular is trite as the column sums are zero.

[8] For the curious reader, we will mention that the definition as well as many properties of almost diagonal matrices easily extend to the case of non-identical entries on the main diagonal.

To verify c) we may check by direct substitution that

$$\mathbf{D}(\alpha)\begin{pmatrix}-1\\(\mathbf{I}_{n-1})_{i*}\end{pmatrix} = 1\cdot\begin{pmatrix}-1\\(\mathbf{I}_{n-1})_{i*}\end{pmatrix} \qquad (1\le i\le n-1),$$

$$\mathbf{D}(\alpha)\mathbb{1} = (n\alpha+1)\mathbb{1}.$$

And d) follows from c) because the largest singular value of $\mathbf{D}(\alpha)$ is given by

$$\sigma_{\text{max}}\big(\mathbf{D}(\alpha)\big) = \sqrt{\lambda_{\text{max}}\big(\mathbf{D}(\alpha)^{\intercal}\mathbf{D}(\alpha)\big)} = \sqrt{\lambda_{\text{max}}\big(\mathbf{D}(\alpha)^2\big)} = \sqrt{\max\{1^2,(n\alpha+1)^2\}}$$

as $\mathbf{D}(\alpha)$ is symmetric. ▶

LEMMA 3: *If $n\alpha\neq-1$, the linear system $\mathbf{D}(\alpha)\mathbf{x}=\mathbf{b}$ has the unique solution*

$$\mathbf{x} = \mathbf{D}\left(\frac{-\alpha}{n\alpha+1}\right)\mathbf{b} = \left(\mathbf{I}-\frac{\alpha}{n\alpha+1}\mathbb{1}\mathbb{1}^{\intercal}\right)\mathbf{b}.$$

PROOF: Apply part a) of Lemma 2. ▶

THEOREM 3—Equilibria for Regular Cliques: *For any given $\mathbf{k}\in\{0,1\}^n$, the fixed point equation (7) has a unique solution $\mathbf{V}^0(\mathbf{k})$ with components given by*

$$V_i^0(\mathbf{k}) = \begin{cases} v\big(\frac{|\mathbf{k}|}{n}\big) & \text{if } k_i=1,\\ v\big(\frac{|\mathbf{k}|}{n}\big)-\beta & \text{if } k_i=0,\end{cases} \tag{33}$$

where $v(\cdot)$ is defined in (11). For it to be a feasible equilibrium of the regular clique in orthant $\mathbb{K}^n(\mathbf{k})$ it is necessary and sufficient that the corresponding feasibility constraint in the system of inequalities (8–10) be satisfied. Setting $\mathbf{k}=\mathbb{0}$ and $\mathbb{1}$ in turn, it follows a fortiori that the extremal equilibria are given putatively by

$$\begin{aligned}\mathbf{V}^{\text{inf}} &= \mathbf{V}^0(\mathbb{0}) = \big(v(0)-\beta\big)\mathbb{1} = \frac{e-\beta}{1-c}\mathbb{1},\\ \mathbf{V}^{\text{sup}} &= \mathbf{V}^0(\mathbb{1}) = v(1)\mathbb{1} = \frac{e}{1-c}\mathbb{1}.\end{aligned} \tag{34}$$

They are feasible if the corresponding feasibility constraints (9, 10), respectively, hold.

PROOF: For any $\mathbf{k}\in\{0,1\}^n$, the fixed point equation (7) may be rewritten in the form

$$\mathbf{D}\left(\frac{-c}{n}\right)\mathbf{V}^0(\mathbf{k}) = (e-\beta)\mathbb{1}+\beta\mathbf{k}.$$

We now apply Lemma 3 to obtain the putative equilibrium

$$\mathbf{V}^0(\mathbf{k}) = \left(\mathbf{I}+\frac{c}{n(1-c)}\mathbb{1}\mathbb{1}^{\intercal}\right)\big((e-\beta)\mathbb{1}+\beta\mathbf{k}\big)$$

and recover (33). To test the feasibility of the solution we check to see whether it lies in the orthant $\mathbb{K}^n(\mathbf{k})$ and rediscover the feasibility constraint. ▶

APPENDIX C: PREPARATION FOR THEOREMS 4 AND 5

As a preamble to the proofs of the main theorems, we begin by collecting facts about generic spacings and systems of sub-exponential random variables, and verify the assertions of the behavior of the de Finetti and Haar-induced systems.

C.1. *On the Moments of Spacings*

Suppose $X_1, \ldots, X_n$ are non-negative, exchangeable, and sum to one; in other words, they form an exchangeable system of spacings of the unit interval engendered by any suitable random process. As $X_1 + \cdots + X_n = 1$, by taking expectations of both sides we see that $n\,\mathbf{E}(X_1) = 1$ by an appeal to exchangeability whence the common expectation of the spacings is $1/n$. Write $Z_j = X_j - 1/n$ for the centered spacings. For $\nu_1, \ldots, \nu_n \geq 0$, introduce notation for the mixed centered moment $\mu_n(\nu_1, \ldots, \nu_n) := \mathbf{E}\big(Z_1^{\nu_1} \cdots Z_n^{\nu_n}\big)$. By exchangeability, $\mu_n(\nu_1, \ldots, \nu_n)$ is invariant with respect to permutations of the indices. If $\nu_j = 0$, the corresponding centered spacing Z_j is not represented in the expectation; we accordingly take a slight liberty with notation and also write $\mu_n(\nu_1, \ldots, \nu_k) := \mu_n(\nu_1, \ldots, \nu_k, 0, \ldots, 0) = \mathbf{E}\big(Z_1^{\nu_1} \cdots Z_k^{\nu_k}\big)$. With these notational conventions, $\mu_n(1) = 0$ (as the variables Z_j are properly centered), $\mu_n(2) = \mathbf{E}(Z_1^2) = \mathrm{Var}(X_1)$ is the common variance, and $\mu_n(1,1) = \mathbf{E}(Z_1 Z_2) = \mathrm{Cov}(X_1, X_2)$ is the common covariance.

We extend the notation also to conditional moments. Suppose $\mathcal{Q}$ is any Borel measurable set of positive probability in the sample space of the spacings. We then also write $\mu_n(\nu_1, \ldots, \nu_n \mid \mathcal{Q}) := \mathbf{E}\big(Z_1^{\nu_1} \cdots Z_n^{\nu_n} \mid \mathcal{Q}\big)$ for the conditional mixed moments. If we identify $\mathcal{Q}$ with the entire sample space [which we may identify with the $(n-1)$-dimensional simplex engendered by the spacings] we recover the unconditional moments. We will restrict attention to sets $\mathcal{Q}$ that preserve symmetry in the variables: formally, the centered spacings $Z_1, \ldots, Z_n$ are conditionally exchangeable given $\mathcal{Q}$. In this setting, the conditional mixed moments $\mu_n(\nu_1, \ldots, \nu_n \mid \mathcal{Q})$ will likewise be invariant with respect to permutations of coordinates. As in the unconditional case, we also take a liberty with notation and write $\mu_n(\nu_1, \ldots, \nu_k \mid \mathcal{Q}) := \mu_n(\nu_1, \ldots, \nu_k, 0, \ldots, 0) = \mathbf{E}\big(Z_1^{\nu_1} \cdots Z_k^{\nu_k} \mid \mathcal{Q}\big)$.

An amazing variety of relationships between mixed and pure moments arises out of the sum constraint $Z_1 + \cdots + Z_n = 0$.

LEMMA 4: *Let $\mathcal{Q}$ be any Borel measurable set of positive probability in the $(n-1)$-dimensional simplex engendered by the spacings and suppose further that the centered spacings $Z_1, \ldots, Z_n$ are conditionally exchangeable given $\mathcal{Q}$. If $1 \leq k \leq n-1$ and $\nu_1, \ldots, \nu_k > 0$, then*

$$\begin{aligned}\mu_n(\nu_1, \ldots, \nu_k, 1 \mid \mathcal{Q}) = \frac{-1}{n-k}\big[&\mu_n(\nu_1 + 1, \nu_2, \ldots, \nu_{k-1}, \nu_k \mid \mathcal{Q}) \\ &+ \mu_n(\nu_1, \nu_2 + 1, \ldots, \nu_{k-1}, \nu_k \mid \mathcal{Q}) + \cdots + \mu_n(\nu_1, \nu_2, \ldots, \nu_{k-1}, \nu_k + 1 \mid \mathcal{Q})\big].\end{aligned}$$

PROOF: As $Z_1^{\nu_1} \ldots Z_k^{\nu_k}(Z_1 + \cdots + Z_n) = 0$, by conditioning with respect to $\mathcal{Q}$ and taking expectations of both sides, we conclude that

$$\begin{aligned}0 &= \mathbf{E}\big(Z_1^{\nu_1} Z_2^{\nu_2} \ldots Z_k^{\nu_k}(Z_1 + Z_2 + \cdots + Z_n) \mid \mathcal{Q}\big) \\ &= \mathbf{E}\big(Z_1^{\nu_1+1} Z_2^{\nu_2} \cdots Z_k^{\nu_k} \mid \mathcal{Q}\big) + \mathbf{E}\big(Z_1^{\nu_1} Z_2^{\nu_2+1} \cdots Z_k^{\nu_k} \mid \mathcal{Q}\big) + \cdots \\ &\qquad + \mathbf{E}\big(Z_1^{\nu_1} Z_2^{\nu_2} \cdots Z_k^{\nu_k+1} \mid \mathcal{Q}\big) + \sum_{l=k+1}^{n} \mathbf{E}\big(Z_1^{\nu_1} Z_2^{\nu_2} \cdots Z_k^{\nu_k} Z_l \mid \mathcal{Q}\big)\end{aligned}$$

$$= \mu_n(\nu_1 + 1, \nu_2, \dots, \nu_k \mid \mathcal{Q}) + \mu_n(\nu_1, \nu_2 + 1, \dots, \nu_k \mid \mathcal{Q}) + \cdots$$
$$+ \mu_n(\nu_1, \nu_2, \dots, \nu_k + 1 \mid \mathcal{Q}) + (n-k)\mu_n(\nu_1, \dots, \nu_k, 1 \mid \mathcal{Q})$$

again by virtue of the exchangeability of the system. ▶

Useful special cases are illuminated by setting $k = 1$ and $k = 2$ and it will be convenient to tabulate these two settings. When $k = 1$ we recover the Proposition of Section 7.2 and a little more besides.

COROLLARY 2: *For any $\nu \geq 0$,*

$$\mu_n(\nu, 1 \mid \mathcal{Q}) = \frac{-\mu_n(\nu + 1 \mid \mathcal{Q})}{n-1}.$$

In particular, the spacings are negatively correlated and satisfy

$$\mu_n(1,1 \mid \mathcal{Q}) = \frac{-\mu_n(2 \mid \mathcal{Q})}{n-1},\ \mu_n(2,1 \mid \mathcal{Q}) = \frac{-\mu_n(3 \mid \mathcal{Q})}{n-1},\ \mu_n(3,1 \mid \mathcal{Q}) = \frac{-\mu_n(4 \mid \mathcal{Q})}{n-1}.$$

COROLLARY 3: *For any $\nu_1,\ \nu_2 \geq 0$,*

$$\mu_n(\nu_1, \nu_2, 1 \mid \mathcal{Q}) = \frac{-\bigl(\mu_n(\nu_1 + 1, \nu_2 \mid \mathcal{Q}) + \mu_n(\nu_1, \nu_2 + 1 \mid \mathcal{Q})\bigr)}{n-2}.$$

In particular,

$$\mu_n(1,1,1 \mid \mathcal{Q}) = \frac{-2\mu_n(2,1 \mid \mathcal{Q})}{n-2} = \frac{2\mu_n(3 \mid \mathcal{Q})}{(n-1)(n-2)},$$
$$\mu_n(2,1,1 \mid \mathcal{Q}) = \frac{-\bigl(\mu_n(3,1 \mid \mathcal{Q}) + \mu_n(2,2 \mid \mathcal{Q})\bigr)}{n-2} = \frac{\mu_n(4 \mid \mathcal{Q})}{(n-1)(n-2)} - \frac{\mu_n(2,2 \mid \mathcal{Q})}{n-2}.$$

The inequality of Cauchy–Schwarz leads to other useful relationships between the moments and we tabulate a few that will be of use.

LEMMA 5: *Suppose $\mathcal{Q}$ is any Borel measurable set satisfying the conditions in Lemma 4. Then:*

$$\mu_n(2 \mid \mathcal{Q})^2 \leq \mu_n(4 \mid \mathcal{Q}),$$
$$\frac{\mu_n(2 \mid \mathcal{Q})^2}{(n-1)^2} \leq \mu_n(2,2 \mid \mathcal{Q}) \leq \mu_n(4 \mid \mathcal{Q}).$$

PROOF: The verification is by repeated application of the conditional Cauchy–Schwarz inequality $\bigl|\mathbf{E}(UV \mid \mathcal{Q})\bigr| \leq \sqrt{\mathbf{E}(U^2 \mid \mathcal{Q})\,\mathbf{E}(V^2 \mid \mathcal{Q})}$. For the bound in the first line of the claimed inequalities, identify $U = Z_1^2$ and $V = 1$; for the first of the bounds in the second line, identify $U = Z_1 Z_2$ and $V = 1$ and use Corollary 2; and for the second of the bounds in the second line, identify $U = Z_1^2$ and $V = Z_2^2$. ▶

C.2. *Sub-Exponential Random Variables*

The spacings induced by the de Finetti and Haar distributions in Examples 1 and 2, respectively, are instances of sub-exponential random variables with sub-exponential norm proportional to n^{-1}. A small detour to pick up the terminology helps consolidate the analysis.

We say that a non-negative valued random variable X is *sub-exponential* if there exist positive constants α and β such that

$$\mathbf{P}\{X \geq x\} \leq \alpha e^{-\beta x} \qquad (x \geq 0). \tag{35}$$

LEMMA 6: *If X is sub-exponential then it has moments of all orders and, moreover, there exists an absolute positive constant K such that $\|X\|_p \leq p \cdot K$ for $p \geq 1$. In particular, if X satisfies the sub-exponential inequality (35), then we may take $K = \beta^{-1} e^{\alpha/e}$.*

PROOF: Suppose X is non-negative and satisfies (35). We recall that if Y is a non-negative random variable then $\mathbf{E}(Y) = \int_0^\infty \mathbf{P}\{Y \geq y\}\,dy$. Select any $p \geq 1$ and identify $Y = X^p$. We then obtain

$$\mathbf{E}(X^p) = \int_0^\infty \mathbf{P}\{X^p > y\}\,dy \overset{(x^p \leftarrow y)}{=} \int_0^\infty \mathbf{P}\{X > x\} \cdot p x^{p-1}\,dx$$
$$\leq \alpha p \int_0^\infty e^{-\beta x} x^{p-1}\,dx \overset{(t \leftarrow \beta x)}{=} \frac{\alpha p}{\beta^p} \int_0^\infty e^{-t} t^{p-1}\,dt = \frac{\alpha p}{\beta^p} \cdot \Gamma(p).$$

The simplest bounds suffice: Stirling's upper bound for the gamma function may be further reduced to the simple form

$$\Gamma(p) \leq \sqrt{\frac{2\pi}{p}} \left(\frac{p}{e}\right)^p < p^p$$

as $\sqrt{2\pi}\,p^{-1/2}e^{-p} \leq \sqrt{2\pi}\,e^{-1} < 1$ for $p \geq 1$. We conclude that $\mathbf{E}(X^p) \leq \alpha p \big(\frac{p}{\beta}\big)^p$, whence $\|X\|_p = \mathbf{E}(X^p)^{1/p} = p \cdot \beta^{-1}(\alpha p)^{1/p}$. The function $f(p) = (\alpha p)^{1/p} = \exp\big\{\frac{1}{p}\log(\alpha p)\big\}$ achieves its unique maximum at $p = e/\alpha$ as is easily verified by differentiation. We conclude that $\|X\|_p \leq p \cdot \beta^{-1} e^{\alpha/e}$ for all $p \geq 1$. ▶

The *sub-exponential norm* $\|X\|_{\Psi_1}$ of X is the smallest value of K for which $\|X\|_p \leq p \cdot K$ for all $p \geq 1$. Alternatively, $\|X\|_{\Psi_1} = \sup_{p \geq 1} p^{-1}\|X\|_p$. Lemma 6 shows that if X satisfies (35) then $\|X\|_{\Psi_1} \leq \beta^{-1} \cdot e^{\alpha/e}$. In our applications, β is an asymptotic parameter and we conclude that the sub-exponential norm is $\mathcal{O}(\beta^{-1})$.

LEMMA 7: *Suppose $\mathcal{X} = \big\{ \big(X_1^{(n)}, \dots, X_n^{(n)}\big) : n \geq 1 \big\}$ is a triangular array of row-wise exchangeable, non-negative valued random variables with row sum unit. If there exist absolute positive constants α and κ such that, for each n,*

$$\mathbf{P}\big\{X_1^{(n)} \geq x\big\} \leq \alpha e^{-\kappa n x} \qquad (x \geq 0),$$

then $\mathcal{X}$ is asymptotically diffuse in the sense of Definition 2.

In many settings spacings will exhibit sub-exponential behavior though we will not levy this stronger constraint: the de Finetti and Haar-induced spacings are typical in this regard.

C.3. *The De Finetti Spacings*

Suppose $X_1, \ldots, X_n$ are the spacings engendered by $n-1$ random points in the unit interval. These spacings are governed by the de Finetti distribution (36), reproduced here for convenience of reference:

$$\mathbf{P}\{X_1 > x_1, \ldots, X_n > x_n\} = \bigl[(1 - x_1 - \cdots - x_n)_+\bigr]^{n-1}. \tag{36}$$

By setting $x_1 = x$ and $x_2 = \cdots = x_n = 0$, we obtain the common marginal distribution of the spacings in the simple form

$$\mathbf{P}\{X_1 \geq x\} = \bigl[(1-x)_+\bigr]^{n-1} \qquad (x \geq 0).$$

The elementary inequality $1 - x \leq e^{-x}$ shows that the right-hand side is $\leq e^{-(n-1)x}$ and we've verified that the spacings are of the generic sub-exponential form (35) with parameters $\alpha = 1$ and $\beta = n - 1$.

LEMMA 8: *The de Finetti spacings are sub-exponential with common sub-exponential norm $\|X_1\|_{\Psi_1} \leq (n-1)^{-1} \cdot e^{1/e}$. Equivalently, $0 \leq \mathbf{E}(X_1^p) \leq p^p e^{p/e} \cdot (n-1)^{-p}$, for each $p \geq 1$.*

If p varies over a bounded range $\{1, \ldots, \nu\}$, then $\mathbf{E}(X_1^p) = \mathcal{O}(n^{-p})$ *uniformly* for $1 \leq p \leq \nu$. In the notation introduced in Section C.1, write $\mu_n(\nu) = \mathbf{E}\bigl\{\bigl(X_1 - \frac{1}{n}\bigr)^\nu\bigr\}$.

COROLLARY 4: *The νth centered moment of the de Finetti spacings satisfies $\mu_n(\nu) = \mathcal{O}(n^{-\nu})$ for each fixed $\nu \geq 0$ and, a fortiori, the de Finetti spacings are asymptotically diffuse.*

PROOF: Expanding via the binomial theorem, we have

$$\mu_n(\nu) = \mathbf{E}\bigl\{\bigl(X_1 - \tfrac{1}{n}\bigr)^\nu\bigr\} = \sum_{k=0}^{\nu} \binom{\nu}{k} \frac{(-1)^k}{n^k} \mathbf{E}\bigl(X_1^{\nu-k}\bigr).$$

Each term in the sum on the right is of asymptotic order $n^{-k} \cdot n^{-(\nu-k)} = n^{-\nu}$ and the asserted result follows. ▶

Lemma 4 and its corollaries provide a variety of bounds for the mixed moments of the de Finetti system. More precision is available here in view of the explicit form of the de Finetti distribution. As before, write $Z_j = X_j - 1/n$ for the centered spacings. For $\nu_1, \ldots, \nu_n \geq 0$, as before, write $\mu_n(\nu_1, \ldots, \nu_n) := \mathbf{E}\bigl(Z_1^{\nu_1} \cdots Z_n^{\nu_n}\bigr)$ for the moments of the centered spacings and introduce the nonce notation $m_n(\nu_1, \ldots, \nu_n) := \mathbf{E}\bigl(X_1^{\nu_1} \cdots X_n^{\nu_n}\bigr)$ for the non-centered moments.

LEMMA 9: *For every choice of $\nu_1 \geq 0, \ldots, \nu_n \geq 0$,*

$$m_n(\nu_1, \ldots, \nu_{n-1}, \nu_n) = \frac{\Gamma(n)\Gamma(\nu_1 + 1) \cdots \Gamma(\nu_{n-1} + 1)\Gamma(\nu_n + 1)}{\Gamma(n + \nu_1 + \cdots + \nu_{n-1} + \nu_n)}. \tag{37}$$

PROOF: As $X_1 + \cdots + X_{n-1} + X_n = 0$, we may as well focus on, say, $X_1, \ldots, X_{n-1}$. Setting $x_n = 0$ in (36), we see that

$$\mathbf{P}\{X_1 > x_1, \ldots, X_{n-1} > x_{n-1}\} = \bigl[(1 - x_1 - \cdots - x_{n-1})_+\bigr]^{n-1} \tag{38}$$

for $x_1, \dots, x_{n-1} \geq 0$. It is clear now that $(X_1, \dots, X_{n-1})$ is concentrated in the regular probability simplex $\mathbb{P}^{n-1}$ defined by the inequalities $x_1 \geq 0, \dots, x_{n-1} \geq 0$, and $x_1 + \cdots + x_{n-1} \leq 1$. Indeed, repeated differentiation of (38) shows that $(X_1, \dots, X_{n-1})$ has density

$$\Gamma(n) = (n-1)! = \frac{1}{\mathrm{Vol}(\mathbb{P}^{n-1})} = \frac{1}{\displaystyle\int_{\mathbb{P}^{n-1}} dx_1 \cdots dx_{n-1}}$$

uniformly distributed in $\mathbb{P}^{n-1}$.

Suppose $0 \leq \xi < 1$, $\mu \geq 0$, and $\nu \geq 0$. Begin with the basic identity

$$\int_0^{1-\xi} x^\mu (1-\xi-x)^\nu \, dx = (1-\xi)^{\mu+\nu+1} \cdot \frac{\Gamma(\mu+1)\Gamma(\nu+1)}{\Gamma(\mu+\nu+2)}. \tag{39}$$

(The change of variable $t \leftarrow x/(1-\xi)$ is salutary and reduces the integral to a beta function.) Iterating the basic identity yields the various mixed moments of the spacings. As $X_n = 1 - X_1 - \cdots - X_{n-1}$, the mixed moments of the spacings are given by

$$m_n(\nu_1, \dots, \nu_{n-1}, \nu_n) := \mathbf{E}\left(X_1^{\nu_1} \cdots X_{n-1}^{\nu_{n-1}} (1 - X_1 - \cdots - X_{n-1})^{\nu_n}\right)$$

for $\nu_1 \geq 0, \dots, \nu_n \geq 0$. Integrating with respect to the uniform density over the regular simplex shows that the right-hand side may be written as the iterated integral

$$(n-1)! \int_0^1 dx_1 \, x_1^{\nu_1} \int_0^{1-x_1} dx_2 \, x_2^{\nu_2} \cdots \int_0^{1-x_1-\cdots-x_{n-2}} dx_{n-1} \, x_{n-1}^{\nu_{n-1}} (1 - x_1 - \cdots - x_{n-1})^{\nu_n}.$$

Working outwards from the innermost integral, each integral has the form of the basic identity (39), and accumulating terms yields (37). ▶

If $\nu_1, \dots, \nu_n$ are non-negative integers then the gamma functions in (37) reduce to factorials and the expression for the moments takes a familiar multinomial form. Write $\nu = \nu_1 + \cdots + \nu_n$ to keep the expression compact. Then

$$m_n(\nu_1, \dots, \nu_n) = \frac{(n-1)!\nu_1! \cdots \nu_n!}{(n-1+\nu)!} =: \frac{1}{\binom{n-1+\nu}{n-1, \nu_1, \dots, \nu_n}}. \tag{40}$$

With a view to keeping representations tidy, it will also be convenient to introduce the nonce notation

$$T_n(\nu_1, \dots, \nu_n) := \sum_{i_1=0}^{\nu_1} \cdots \sum_{i_n=0}^{\nu_n} \frac{(-1)^{i_1+\cdots+i_n}}{n^{i_1+\cdots+i_n}} \binom{n-1+\nu}{n-1+\nu-i_1-\cdots-i_n, i_1, \dots, i_n}. \tag{41}$$

LEMMA 10: *If $\nu_1, \dots, \nu_n$ are non-negative integers then the mixed moments of the centered de Finetti spacings are given by*

$$\mu_n(\nu_1, \dots, \nu_n) = m_n(\nu_1, \dots, \nu_n) \cdot T_n(\nu_1, \dots, \nu_n). \tag{42}$$

PROOF: Expanding each $Z_j^{\nu_j} = (X_j - 1/n)^{\nu_j}$ via the binomial theorem and collecting terms, we obtain

$$\mu_n(\nu_1,\dots,\nu_n) = \sum_{i_1=0}^{\nu_1}\cdots\sum_{i_n=0}^{\nu_n}\binom{\nu_1}{i_1}\cdots\binom{\nu_n}{i_n}\frac{(-1)^{i_1+\dots+i_n}}{n^{i_1+\dots+i_n}} m_n(\nu_1-i_1,\dots,\nu_n-i_n)$$

which we can put in a slightly more informative form by leveraging (40). Multiplying and dividing the summands by $m_n(\nu_1,\dots,\nu_n)$ we discover the alternating form (41) on the right once we factor out $m_n(\nu_1,\dots,\nu_n)$. ▶

COROLLARY 5: *The centered moments of orders two and four of the de Finetti spacings are given by:*

$$\begin{aligned}
\mu_n(2) &= \frac{2}{(n+1)n}\cdot\frac{n-1}{2n} \sim \frac{1}{n^2},\\
\mu_n(1,1) &= \frac{-1}{n^2(n+1)} \sim \frac{-1}{n^3},\\
\mu_n(2,2) &= \frac{4}{(n+3)(n+2)(n+1)n}\cdot\frac{n^3-2n^2+15n-18}{4n^3} \sim \frac{1}{n^4},\\
\mu_n(4) &= \frac{24}{(n+3)(n+2)(n+1)n}\cdot\frac{(n-1)(3n^2-7n+6)}{8n^3} \sim \frac{9}{n^4}.
\end{aligned} \tag{43}$$

The form of the covariance $\mu_n(1,1)$ obtained from Lemma 10 may be verified directly via the negative correlation relation $\mu_n(1,1) = -\mu_n(2)/(n-1)$ obtained in Corollary 2.

The explicit form (42) for the de Finetti moments invites an asymptotic analysis leading to a precise form of Corollary 4. The result illuminates precisely how the negatively correlated dependency structure of the spacings creates additional rapid decay in the mixed moments.

LEMMA 11: *Let ν be any fixed positive integer and $\nu_1, \dots, \nu_n$ any collection of non-negative integers summing to ν. Suppose further that ρ of the ν_j are equal to one where $0 \le \rho \le \nu$. Then the centered de Finetti moment $\mu_n(\nu_1,\dots,\nu_n)$ has sign $(-1)^{\lceil \frac{\rho}{2}\rceil}$ and asymptotic order $n^{-(\nu+\lceil\frac{\rho}{2}\rceil)}$.*

PROOF: The alternating sum (41) looks formidable to be sure but repeated evaluations of the innermost sum show that it takes on a generic form asymptotically with n. We take in a small analytic detour in anticipation.

As the alternating sum is invariant with respect to permutations of coordinates, in an exuberant extension of notation we also write

$$\begin{aligned}
T_n(\nu_1,\dots,\nu_{n-\tau},\underbrace{0,\dots,0}_{\tau\text{ terms}}) &= \sum_{i_1=0}^{\nu_1}\cdots\sum_{i_{n-\tau}=0}^{\nu_{n-\tau}}\frac{(-1)^{i_1+\dots+i_{n-\tau}}}{n^{i_1+\dots+i_{n-\tau}}}\\
&\quad\times\binom{n-1+\nu}{n-1+\nu-i_1-\dots-i_{n-\tau}, i_1,\dots,i_{n-\tau}} =: T_n(\nu_1,\dots,\nu_{n-\tau})
\end{aligned} \tag{44}$$

for a version of the alternating sum when $\nu_{i_j} = 0$ for τ of the indices j (which, by invariance with respect to permutations of indices in the sum, we may as well take to be the final τ indices,

$j = n - \tau + 1, n - \tau + 2, \dots, n$). For each fixed non-negative integer I and integer $\xi \leq I$, now consider the basic alternating sum

$$S(\xi; n, I) := \sum_{i=0}^{\xi} \frac{(-1)^i}{n^i} \binom{n+I}{i} = \sum_{i=0}^{\xi} \frac{(-1)^i}{i!} \prod_{j=0}^{i-1} \left(1 + \frac{I-j}{n}\right). \tag{45}$$

In particular, direct calculation shows that for all $I \geq 0$,

$$S(0; n, I) = 1,$$
$$S(1; n, I) = 1 - \frac{n+I}{n} = -\frac{I}{n}.$$

More generally, asymptotically with n, the product on the right in (45) approaches one. Indeed, by expanding out the product, we see that

$$S(\xi; n, I) = \sum_{i=0}^{\xi} \frac{(-1)^i}{i!} \left[1 + \frac{1}{n} \sum_{j=0}^{i-1} (I - j) + \mathcal{O}\left(\frac{1}{n^2}\right)\right]$$
$$= \sum_{i=0}^{\xi} \frac{(-1)^i}{i!} \left[1 + \frac{i\left(I - \frac{i-1}{2}\right)}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)\right].$$

The term inside square brackets on the right approaches one asymptotically and we are led to introduce notation for the truncated alternating exponential series,

$$E_\xi := \sum_{i=0}^{\xi} \frac{(-1)^i}{i!} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^\xi}{\xi!} \qquad (\xi \geq 0).$$

Excepting only the case $\xi = 1$, when $E_1 = 0$ because of accidental cancellation, the term E_ξ is strictly positive and converges rapidly to e^{-1} as ξ increases. In particular, $E_0 = 1$, $E_1 = 0$, and, for any $\xi \geq 0$, $\left|E_\xi - e^{-1}\right| \leq \frac{1}{(\xi+1)!}$. We conclude that

$$S(\xi; n, I) = E_\xi + \mathcal{O}\left(\frac{1}{n}\right) = e^{-1} + \mathcal{O}\left(\frac{1}{(\xi+1)!} + \frac{1}{n}\right)$$

is strictly positive and bounded away from zero for $\xi \neq 1$. It is only the case $\xi = 1$ when accidental cancellations result in $S(1; n, I)$ being strictly negative [unless $I = 0$, in which case $S(1; n, 0) = 0$] and of order $1/n$: $S(1; n, I) = -I/n$.

Returning to the alternating sum (41), suppose τ of the ν_j are equal to zero, ρ of the ν_j are equal to one, and the remaining $\sigma = n - \rho - \tau$ of the ν_j exceed one. We may suppose that we so arrange the indices that

$$\begin{aligned} \nu_1 = \nu_2 = \cdots = \nu_\rho &= 1, \\ \nu_{\rho+1},\ \nu_{\rho+2},\ \dots,\ \nu_{\rho+\sigma} &\geq 2, \\ \nu_{\rho+\sigma+1} = \nu_{\rho+\sigma+2} = \cdots = \nu_{\rho+\sigma+\tau} &= 0. \end{aligned} \tag{46}$$

Recursively specify the variables $I_0 = \nu - 1$ and $I_j = I_{j-1} - i_j$ for $j \geq 1$. Writing $1^{(\rho)}$ for a ρ-term sequence of 1s, (44) reduces to the form

$$\begin{aligned} T_n\bigl(1^{(\rho)}, \nu_{\rho+1}, \dots, \nu_{\rho+\sigma}\bigr) &= \sum_{i_1=0}^{1} \frac{(-1)^{i_1}}{n^{i_1}} \binom{n+I_0}{i_1} \cdots \sum_{i_\rho=0}^{1} \frac{(-1)^{i_\rho}}{n^{i_\rho}} \binom{n+I_{\rho-1}}{i_\rho} \\ &\quad \times \sum_{i_{\rho+1}=0}^{\nu_{\rho+1}} \frac{(-1)^{i_{\rho+1}}}{n^{i_{\rho+1}}} \binom{n+I_\rho}{i_{\rho+1}} \cdots \sum_{i_{\rho+\sigma}=0}^{\nu_{\rho+\sigma}} \frac{(-1)^{i_{\rho+\sigma}}}{n^{i_{\rho+\sigma}}} \binom{n+I_{\rho+\sigma-1}}{i_{\rho+\sigma}}. \end{aligned}$$

We identify the innermost sum on the right with

$$S(\nu_{\rho+\sigma}; n, I_{\rho+\sigma-1}) = E_{\nu_{\rho+\sigma}} + \mathcal{O}\bigl(n^{-1}\bigr)$$

where $\nu_{\rho+\sigma} \geq 2$, whence $E_{\nu_{\rho+\sigma}}$ is strictly positive. We conclude that

$$T_n\bigl(1^{(\rho)}, \nu_{\rho+1}, \dots, \nu_{\rho+\sigma}\bigr) = T_n\bigl(1^{(\rho)}, \nu_{\rho+1}, \dots, \nu_{\rho+\sigma-1}\bigr) \cdot \bigl[E_{\nu_{\rho+\sigma}} + \mathcal{O}\bigl(n^{-1}\bigr)\bigr].$$

As $\nu_{\rho+1}, \dots, \nu_{\rho+\sigma} \geq 2$, each of $E_{\nu_{\rho+1}}, \dots, E_{\nu_{\rho+\sigma}}$ is positive and bounded away from zero. Churning through the induction machinery hence results in

$$\begin{aligned} T_n\bigl(1^{(\rho)}, \nu_{\rho+1}, \dots, \nu_{\rho+\sigma}\bigr) &= T_n\bigl(1^{(\rho)}\bigr) \times \bigl[E_{\nu_{\rho+1}} + \mathcal{O}\bigl(n^{-1}\bigr)\bigr] \times \cdots \times \bigl[E_{\nu_{\rho+\sigma}} + \mathcal{O}\bigl(n^{-1}\bigr)\bigr] \\ &= T_n\bigl(1^{(\rho)}\bigr) \cdot E_{\nu_{\rho+1}} \cdots E_{\nu_{\rho+\sigma}} \bigl[1 + \mathcal{O}\bigl(n^{-1}\bigr)\bigr]. \end{aligned} \tag{47}$$

To estimate $T_n\bigl(1^{(\rho)}\bigr)$, begin by streamlining our expressions by introducing the nonce notation

$$h_j = \frac{(-1)^{i_j}}{n^{i_j}} \binom{n+I_{j-1}}{i_j} \qquad (1 \leq j \leq \rho).$$

With base $R_0 = 1$, now recursively define the iterated sums

$$R_j = \sum_{i_{\rho-j+1}=0}^{1} h_{\rho-j+1} R_{j-1} \qquad (1 \leq j \leq \rho),$$

whence $T_n\bigl(1^{(\rho)}\bigr) = R_\rho$. Computing the recurrence through the first step shows

$$R_1 = \sum_{i_\rho=0}^{1} h_\rho R_0 = \sum_{i_\rho=0}^{1} \frac{(-1)^{i_\rho}}{n^{i_\rho}} \binom{n+I_{\rho-1}}{i_\rho} = 1 - \frac{n+I_{\rho-1}}{n} = -\frac{I_{\rho-1}}{n}.$$

Observe the accidental cancellation of the order one term; this is a feature of this iterated sum. Evaluating the recurrence through one more step extracts the repeating pattern. As $I_{\rho-1} = I_{\rho-2} - i_{\rho-1}$, we obtain

$$\begin{aligned} R_2 = \sum_{i_{\rho-1}=0}^{1} h_{\rho-1} R_1 &= \sum_{i_{\rho-1}=0}^{1} \frac{(-1)^{i_{\rho-1}}}{n^{i_{\rho-1}}} \binom{n+I_{\rho-2}}{i_{\rho-1}} \left(-\frac{I_{\rho-2} - i_{\rho-1}}{n}\right) \\ &= -\frac{1}{n}\left[I_{\rho-2} - \frac{n+I_{\rho-2}}{n}(I_{\rho-2} - 1)\right] \\ &= -\frac{1}{n} + \frac{I_{\rho-2}(I_{\rho-2} - 1)}{n^2} = -\frac{1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right). \end{aligned}$$

By induction, we conclude that,

$$R_{2k-1} = \frac{(-1)^k I_{\rho-2k+1}}{n^k} + \mathcal{O}\left(\frac{1}{n^{k+1}}\right),$$
$$R_{2k} = \frac{(-1)^k}{n^k} + \mathcal{O}\left(\frac{1}{n^{k+1}}\right),$$

for $k \geq 1$, and, *a fortiori*,

$$T_n\bigl(1^{(\rho)}\bigr) = \mathcal{O}\bigl(n^{-\lceil \frac{\rho}{2} \rceil}\bigr) \tag{48}$$

and has sign $(-1)^{\lceil \frac{\rho}{2} \rceil}$.

It only remains to estimate the first term on the right in (42). Under the arrangement (46), one last appeal to (40) shows that

$$m_n(\nu_1, \dots, \nu_n) = m_n\bigl(1^{(\rho)}, \nu_{\rho+1}, \dots, \nu_{\rho+\sigma}\bigr) = \frac{\nu_{\rho+1}! \cdots \nu_{\rho+\sigma}!}{n^\nu}\left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right]. \tag{49}$$

With the asymptotic estimates (47,48,49) securely in hand, we return victorious to (42) and conclude that

$$\begin{aligned}\mu_n\bigl(1^{(\rho)}, \nu_{\rho+1}, \dots, \nu_{\rho+\sigma}\bigr) &= m_n\bigl(1^{(\rho)}, \nu_{\rho+1}, \dots, \nu_{\rho+\sigma}\bigr) \cdot T_n\bigl(1^{(\rho)}, \nu_{\rho+1}, \dots, \nu_{\rho+\sigma}\bigr) \\ &= \frac{\nu_{\rho+1}! \cdots \nu_{\rho+\sigma}!}{n^\nu} \cdot \Theta_n(\rho) \\ &\qquad\qquad \times E_{\nu_{\rho+1}} \cdots E_{\nu_{\rho+\sigma}} \cdot \left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right]\end{aligned}$$

where $\Theta_n(\rho) = \mathcal{O}\bigl(n^{-\lceil \frac{\rho}{2} \rceil}\bigr)$ and has sign $(-1)^{\lceil \frac{\rho}{2} \rceil}$. ▶

Specializing Lemma 11 to the case of a pure moment yields a refinement of Corollary 4.

COROLLARY 6: *The pure centered de Finetti moments have asymptotic order $\mu_n(\nu) \asymp n^{-\nu}$. More expansively, for every integer $\nu \geq 0$, there exist constants A and B such that $An^{-\nu} \leq \mu_n(\nu) \leq Bn^{-\nu}$.*

C.4. *Spacings Induced by the Haar System*

The distribution of spacings inherited from the Haar system as in Example 2 is related to the area of caps on the unit sphere. Begin with some terminology and notation.

Suppose $\mathbf{t}$ is any point on the unit sphere $\mathbb{S}^{n-1}$. Each such point determines a ray from the origin passing through that point. For each $0 \leq \rho \leq 1$, identify the set $\mathbb{C}_n(\rho; \mathbf{t})$ as the collection of points $\mathbf{z}$ on the unit sphere satisfying $\langle \mathbf{z}, \mathbf{t} \rangle \geq \rho$. Geometrically, these are the points on the sphere whose projection onto the ray passing through $\mathbf{t}$ is at least as large as ρ. In a geometrically vivid terminology, we call each such collection a *cap* on the sphere centered at $\mathbf{t}$ and with height $1 - \rho$.

For each n, it will also be convenient to introduce the nonce notation $\mathbb{S}^{n-1}(r)$ for the sphere (nominally centered at the origin) of radius r in n dimensions, with $\mathbb{S}^{n-1} = \mathbb{S}^{n-1}(1)$ representing the unit sphere. Writing $d\Omega$ for a differential element with respect to geodesic distance on the unit sphere, a simple scale of variable shows that

$$\operatorname{Area}\bigl(\mathbb{S}^{n-1}(r)\bigr) = \int_{\mathbb{S}^{n-1}(r)} d\Omega \overset{(\Omega \leftarrow \Omega/r)}{=} r^{n-1} \int_{\mathbb{S}^{n-1}} d\Omega = r^{n-1} A_n$$

where, rewriting (17) for ease of reference,

$$A_n := \mathrm{Area}(\mathbb{S}^{n-1}) = \int_{\mathbb{S}^{n-1}} d\Omega = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)} \tag{50}$$

is the area of the unit sphere in n dimensions.

LEMMA 12: *Suppose* $\mathbf{t} \in \mathbb{S}^{n-1}$ *and* $0 \leq \rho \leq 1$. *Then:*

$$\mathrm{Area}\big(\mathbb{C}_n(\rho;\mathbf{t})\big) = \frac{2\pi^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \int_\rho^1 \left(1-z^2\right)^{\frac{n-3}{2}} dz. \tag{51}$$

PROOF: The area of caps of given height parameter ρ is invariant with respect to choice of center $\mathbf{t}$ as the Haar measure on the sphere is rotation invariant and we may as well identify the cap center $\mathbf{t}$ with the unit first coordinate vector $\mathbf{e}_1 := (1,0,\dots,0)$. The cap $\mathbb{C}_n(\rho;\mathbf{e}_1)$ consists of the points $\mathbf{z} = (z_1,\dots,z_n)$ on the unit sphere $\mathbb{S}^{n-1}$ for which $z_1 \geq \rho$. Moving to spherical coordinates, identify the first coordinate with an azimuthal projection $z_1 = \sin\theta$ for $0 \leq \theta \leq \pi/2$. For each $z \geq 0$, the intersection of the hyperplane $z_1 = z$ with the unit sphere consists of the points on $\mathbb{S}^{n-1}$ for which

$$z_1 = z \text{ and } \sqrt{z_2^2 + \cdots + z_n^2} = \sqrt{1-z^2} = \sqrt{1-\sin^2\theta} = \cos\theta.$$

That is to say, for $\rho \leq z \leq 1$, the intersection of the hyperplane $z_1 = z = \sin\theta$ with $\mathbb{S}^{n-1}$ is isomorphic to $\mathbb{S}^{n-2}(\cos\theta)$, the sphere of radius $\cos\theta$ in $(n-1)$ dimensions. Integrating out over θ shows that

$$\mathrm{Area}\big(\mathbb{C}_n(\rho;\mathbf{e}_1)\big) = \int_{\sin^{-1}\rho}^{\pi/2} \mathrm{Area}\big(\mathbb{S}^{n-2}(\cos\theta)\big)\, d\theta = A_{n-1} \int_{\sin^{-1}\rho}^{\pi/2} \cos(\theta)^{n-2}\, d\theta.$$

Use (50) with n replaced by $n-1$ and introduce the natural change of variable $z = \sin\theta$ to complete the proof. ▶

As in Example 2, suppose $(Z_1,\dots,Z_n)$ is a random point on the unit sphere $\mathbb{S}^{n-1}$ obtained by sampling from the Haar measure on the sphere. In view of (50), the identity (51) is equivalent to the statement

$$\mathbf{P}\{Z_1 \geq t\} = \frac{\mathrm{Area}\big(\mathbb{C}_n(t;\mathbf{e}_1)\big)}{\mathrm{Area}(\mathbb{S}^{n-1})} = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)} \int_t^1 \left(1-z^2\right)^{\frac{n-3}{2}} dz \tag{52}$$

which establishes the common marginal distribution of the variables Z_1, ..., Z_n. As an easy consequence, we recover the well-known property that the area of a cap is exponentially small in relation to the area of the sphere.

LEMMA 13: *Suppose* $n \geq 3$. *Then* $\mathbf{P}\{Z_1 \geq t\} \leq \frac{3}{2} e^{-(n-3)t^2/2}$ *for every* $t \geq 0$.[9]

[9] The constant "$\frac{3}{2}$" can be improved but is not germane for our purposes.

PROOF: When $n = 3$ the result is trite; suppose accordingly that $n > 3$. Starting from (52) we then have

$$\begin{aligned}
\mathbf{P}\{Z_1 \geq t\} &\overset{\text{(i)}}{\leq} \frac{\Gamma\big(\frac{n}{2}\big)}{\sqrt{\pi}\,\Gamma\big(\frac{n-1}{2}\big)} \int_t^1 e^{-(n-3)z^2/2}\,dz \\
&\leq \frac{\Gamma\big(\frac{n}{2}\big)}{\sqrt{\pi}\,\Gamma\big(\frac{n-1}{2}\big)} \int_t^\infty e^{-(n-3)z^2/2}\,dz \\
&\overset{\text{(ii)}}{=} \frac{\Gamma\big(\frac{n}{2}\big)}{\sqrt{\pi}\,\Gamma\big(\frac{n-1}{2}\big)} \cdot \frac{\sqrt{2\pi}}{\sqrt{n-3}} \int_{\sqrt{n-3}\,t}^\infty \phi(y)\,dy \\
&\overset{\text{(iii)}}{\leq} \frac{\Gamma\big(\frac{n}{2}\big)}{\sqrt{\pi}\,\Gamma\big(\frac{n-1}{2}\big)} \cdot \frac{\sqrt{2\pi}}{\sqrt{n-3}} \cdot \frac{e^{-(n-3)t^2/2}}{2}.
\end{aligned}$$

Step (i) follows from the exponential inequality $1 - x \leq e^{-x}$; (ii) is via the natural change of variable $y \leftarrow \sqrt{n-3}\,z$ with $\phi(y) = (2\pi)^{-1/2}e^{-y^2/2}$; and (iii) follows from the normal tail bound $\int_\tau^\infty \phi(x)\,dx \leq \frac{1}{2}e^{-\tau^2/2}$ for $\tau \geq 0$ [Venkatesh (2013, Lemma VI.1.3, p. 165)]. Numerical computation starting from Stirling's bounds for the gamma function shows that the right-hand side is $\leq 1.3e^{-(n-3)t^2/2}$ for $n \geq 4$ and the claimed result follows where we have worsened the constant a little in the interests of keeping the bounds neat. ▶

The non-negative, exchangeable values $X_1 = Z_1^2, \dots, X_n = Z_n^2$ are the Haar-induced spacings of the unit interval. Lemma 13 shows that for $n \geq 4$,

$$\mathbf{P}\{X_1 \geq x\} = \mathbf{P}\big\{|Z_1| \geq \sqrt{x}\,\big\} = 2\,\mathbf{P}\big\{Z_1 \geq \sqrt{x}\,\big\} \leq 3e^{-(n-3)x/2} \qquad (x \geq 0),$$

whence the marginal distribution is sub-exponential with parameters $\alpha = 3$ and $\beta = (n-3)/2$.

LEMMA 14: *The Haar-induced spacings are sub-exponential with common sub-exponential norm* $\|X_1\|_{\Psi_1} \leq 2e^{3/e}(n-3)^{-1}$.

COROLLARY 7: *The νth centered moment of the Haar-induced spacings satisfies* $\mu_n(\nu) := \mathbf{E}\big((X_1 - 1/n)^\nu\big) = \mathcal{O}(n^{-\nu})$ *for each fixed* $\nu \geq 0$ *and, a fortiori, the Haar-induced spacings are asymptotically diffuse.*

Explicit expressions for the distribution and the moments of the Haar-induced spacings are readily inferred from (52). As before, write $m_n(\nu) = \mathbf{E}(X_1^\nu)$ and $\mu_n(\nu) = \mathbf{E}\big((X_1 - 1/n)^\nu\big)$.

LEMMA 15: *The Haar-induced spacings share a common marginal distribution with density*

$$f(x) = \frac{\Gamma\big(\frac{n}{2}\big)}{\sqrt{\pi}\,\Gamma\big(\frac{n-1}{2}\big)} \cdot x^{-\frac{1}{2}}(1-x)^{\frac{n-3}{2}} \qquad (0 < x < 1). \tag{53}$$

This is the beta density with parameters $1/2$ *and* $(n-1)/2$.

PROOF: For $x \geq 0$, courtesy (52), we have

$$\mathbf{P}\{X_1 \geq x\} = 2\,\mathbf{P}\big\{Z_1 \geq \sqrt{x}\,\big\} = \frac{2\Gamma\big(\frac{n}{2}\big)}{\sqrt{\pi}\,\Gamma\big(\frac{n-1}{2}\big)} \int_{\sqrt{x}}^1 \big(1 - z^2\big)^{\frac{n-3}{2}}\,dz$$

$$= \frac{2\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)} \int_x^1 \left(1-\xi\right)^{\frac{n-3}{2}} \cdot \frac{1}{2\sqrt{\xi}}\,d\xi$$

via the natural change of variable $\xi \leftarrow z^2$. Differentiation with respect to x yields the common density via the fundamental theorem of the calculus. ▶

COROLLARY 8: *For each non-negative integer ν, the moments of the spacings induced by the Haar distribution on the unit sphere $\mathbb{S}^{n-1}$ satisfy*

$$m_n(\nu) = \frac{\Gamma\left(\nu+\frac{1}{2}\right)}{\sqrt{\pi}} \cdot \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n}{2}+\nu\right)} \sim \frac{2^\nu\Gamma\left(\nu+\frac{1}{2}\right)}{\sqrt{\pi}} \cdot \frac{1}{n^\nu} \qquad (n \to \infty).$$

The corresponding centered moments $\mu_n(\nu)$ are hence also of asymptotic order $n^{-\nu}$.

PROOF: Integrating out with respect to the density (53), we have

$$m_n(\nu) = \int_0^1 x^\nu f(x)\,dx = \frac{\Gamma(\dfrac{n}{2})}{\sqrt{\pi}\,\Gamma(\dfrac{n-1}{2})} \int_0^1 x^{\left(\nu+\frac{1}{2}\right)-1}(1-x)^{\frac{n-1}{2}-1}\,dx,$$

The integral on the right is in the form of the beta function

$$B(a,b) := \int_0^1 x^{a-1}(1-x)^{b-1}\,dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

where we identify $a = \nu + \frac{1}{2}$ and $b = \frac{n-1}{2}$. Stirling's formula wraps up the asymptotics. ▶

APPENDIX D: PROOFS OF THEOREMS 4 AND 5

Theorem 4 is subsumed by Theorem 5 and it hence suffices to prove the latter. Our approach exploits two main ideas: (1) the putative equilibria for the random matrix of cross-holdings represent perturbations of the putative equilibria for regular cliques; and (2) the individual spacings cannot get pathologically large. The properties of the random matrix of cross-holdings are naturally central to the analysis and we begin by a characterization of its structural properties.

D.1. *The Centered Cross-Holdings Matrix*

Recall that, for each $n \geq 1$, the matrix of cross-holdings of a regular clique is given by $\mathbf{C}^0 = \mathbf{C}^{(0,n)} = \frac{c}{n}\mathbb{1}_n\mathbb{1}_n^\intercal$ and that of a random clique is given by $\mathbf{C} = \mathbf{C}^{(n)} = c\left[\mathbf{X}_1^{(n)} \cdots \mathbf{X}_n^{(n)}\right]$ where the columns are drawn by independent sampling from a common asymptotically diffuse distribution.

Hide the dependence on n for the nonce. In view of the exchangeability of the variables, $\mathbf{E}(\mathbf{C}) = \mathbf{C}^0 = \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$ (see Section C.1). Then

$$\mathbf{\Lambda} = [\Lambda_{ij}]_{i,j=1}^n := \mathbf{C} - \tfrac{1}{n}\mathbb{1}\mathbb{1}^\intercal = c\left[\mathbf{X}_1 - \tfrac{1}{n}\mathbb{1}\ \mathbf{X}_2 - \tfrac{1}{n}\mathbb{1} \cdots \mathbf{X}_n - \tfrac{1}{n}\mathbb{1}\right],$$

where $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ are independent and asymptotically diffuse, represents the proper centering of the cross-holdings matrix at its mean value. The centered spacings are patently bounded and it is worth cataloguing it for future reference.

LEMMA 16: *The centered spacings are uniformly bounded and satisfy* $|\Lambda_{ij}| < c < 1$.

PROOF: As each spacing X_{ij} lies in the unit interval, $|\Lambda_{ij}| \leq c \max\{1 - \frac{1}{n}, \frac{1}{n}\} < c$. ▶

The fact that the column sum of the spacings is unit leads to another simple observation.

LEMMA 17: *The matrix* $\mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$ *is invertible and satisfies* $\big(\mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)^{-1}\mathbf{\Lambda} = \mathbf{\Lambda}$.

PROOF: The matrix $\mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$ is almost diagonal and non-singular as $c < 1$. Applying Lemma 2a with $\alpha = -c/n$, we obtain

$$\left(\mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal\right)^{-1}\mathbf{\Lambda} = \left(\mathbf{I} + \frac{c}{n(1-c)}\mathbb{1}\mathbb{1}^\intercal\right)\mathbf{\Lambda} = \mathbf{\Lambda} + \frac{c}{n(1-c)}\mathbb{1}\mathbb{1}^\intercal\mathbf{\Lambda}.$$

Our centering means that the column sums of $\mathbf{\Lambda}$ are identically zero: for each j,

$$(\mathbb{1}^\intercal\mathbf{\Lambda})_j = \Lambda_{1j} + \cdots + \Lambda_{nj} = \big(C_{1j} - \tfrac{c}{n}\big) + \cdots + \big(C_{nj} - \tfrac{c}{n}\big) = 0.$$

We conclude that $\mathbb{1}^\intercal\mathbf{\Lambda} = \mathbb{0}^\intercal$ and hence that $\big(\mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)^{-1}\mathbf{\Lambda} = \mathbf{\Lambda}$. ▶

LEMMA 18: *The matrix* $\mathbf{I} - \mathbf{\Lambda}$ *is non-singular and, furthermore, its inverse has a convergent Neumann series*

$$(\mathbf{I} - \mathbf{\Lambda})^{-1} = \sum_{l=0}^{\infty} \mathbf{\Lambda}^l.$$

PROOF: Gershgorin discs are still the driving force but the influence is a little more subtle: a direct attempt at proof via a Gershgorin argument along the lines of Lemma 1 is a bit too crude as it only allows us to deduce that the spectrum of $\mathbf{I} - \mathbf{\Lambda}$ lies in the disc $D(1, 2c)$ in the complex plane and, for $1/2 < c < 1$, this disc includes the origin. However, by Lemma 17,

$$\mathbf{I} - \mathbf{C} = \big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big) - \mathbf{\Lambda} = \big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)\big[\mathbf{I} - \big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)^{-1}\mathbf{\Lambda}\big] = \big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)(\mathbf{I} - \mathbf{\Lambda}).$$

But $\mathbf{I} - \mathbf{C}$ is invertible in view of Lemma 1. Multiplying both sides by $(\mathbf{I} - \mathbf{C})^{-1}$ leads hence to the identity

$$\mathbf{I} = (\mathbf{I} - \mathbf{C})^{-1}\big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)(\mathbf{I} - \mathbf{\Lambda}).$$

We conclude that $\mathbf{I} - \mathbf{\Lambda}$ is invertible and identify

$$(\mathbf{I} - \mathbf{\Lambda})^{-1} = (\mathbf{I} - \mathbf{C})^{-1}\big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big).$$

Now write $\rho(\mathbf{\Lambda})$ for the spectral radius of $\mathbf{\Lambda}$. Following Carl Neumann, the Neumann series $\sum_{l\geq 0}\mathbf{\Lambda}^l$ converges if, and only if, $\rho(\mathbf{\Lambda}) < 1$, and, in this case, the series converges to $(\mathbf{I} - \mathbf{\Lambda})^{-1}$. Write $\mathbf{X} = \big[\mathbf{X}_1\ \mathbf{X}_2 \cdots \mathbf{X}_n\big]$ for the matrix of spacings and identify $\mathbf{\Lambda} = c\big(\mathbf{X} - \frac{1}{n}\mathbb{1}\mathbb{1}^\intercal\big)$. Then

$$\mathbf{\Lambda}^2 = c^2\big(\mathbf{X}^2 - \mathbf{X}\cdot\tfrac{1}{n}\mathbb{1}\mathbb{1}^\intercal - \tfrac{1}{n}\mathbb{1}\mathbb{1}^\intercal\mathbf{X} + \tfrac{1}{n^2}\mathbb{1}\mathbb{1}^\intercal\mathbb{1}\mathbb{1}^\intercal\big) = c^2\big(\mathbf{X}^2 - \mathbf{X}\cdot\tfrac{1}{n}\mathbb{1}\mathbb{1}^\intercal\big)$$

as $\mathbb{1}^\intercal\mathbf{X} = \mathbb{1}^\intercal$ because $\mathbf{X}$ is column stochastic and $\mathbb{1}^\intercal\mathbb{1} = n$. It follows by induction that

$$\mathbf{\Lambda}^l = c^l\big(\mathbf{X}^l - \mathbf{X}^{l-1}\cdot\tfrac{1}{n}\mathbb{1}\mathbb{1}^\intercal\big) \qquad (l \geq 1). \tag{54}$$

Again invoking the column stochasticity of the spacings matrix, the sub-multiplicativity of the operator norm induced by the vector l^1-norm shows that

$$\left\|\mathbf{X}^l\right\|_1 \leq \|\mathbf{X}\|_1^l = 1 \qquad (l \geq 0)$$

so that each of the matrices $\mathbf{X}^l$ is column sub-stochastic for each l. (This also follows by appeal to the Chapman–Kolmogorov equations by identifying the (row) stochastic matrix $\mathbf{X}^\intercal$ with the transition matrix of a Markov chain.) Likewise, by appeal to sub-multiplicativity of the operator norm once more,

$$\left\|\mathbf{X}^{l-1} \cdot \tfrac{1}{n}\mathbb{1}\mathbb{1}^\intercal\right\|_1 \leq \left\|\mathbf{X}^{l-1}\right\|_1 \cdot \left\|\tfrac{1}{n}\mathbb{1}\mathbb{1}^\intercal\right\|_1 \leq 1 \qquad (l \geq 1)$$

as $\frac{1}{n}\mathbb{1}\mathbb{1}^\intercal$ is trivially doubly stochastic. We conclude *a fortiori* that the components of $\mathbf{X}^l - \mathbf{X}^{l-1} \cdot \frac{1}{n}\mathbb{1}\mathbb{1}^\intercal$ are uniformly bounded and take values in $[-1,1]$. As $0 < c < 1$, the representation (54) allows us to conclude that $\mathbf{\Lambda}^l \to \mathbb{0}$ as $l \to \infty$ (and, indeed, component-wise convergence is exponentially fast and uniform).

Now suppose r is any eigenvalue of $\mathbf{\Lambda}$ in the complex plane, $\mathbf{v}$ any associated eigenvector: $\mathbf{\Lambda v} = r\mathbf{v}$. By iteration, $\mathbf{\Lambda}^l\mathbf{v} = r^l\mathbf{v}$. But the left-hand side converges to $\mathbb{0}$ which can only happen if $|r| < 1$. All the eigenvalues of $\mathbf{\Lambda}$ are hence confined to the *open* unit disc centered at the origin. Or, what is the same thing, $\rho(\mathbf{\Lambda}) < 1$. ▶

D.2. *A Perturbation Analysis of Putative Equilibria*

Fix any solvency indicator $\mathbf{k} \in \{0,1\}^n$ and let $\mathbf{V}^0(\mathbf{k})$ and $\mathbf{V}(\mathbf{k})$ be the corresponding putative equilibria for the regular clique and the random clique, respectively. These putative equilibria exist in view of Lemma 1. Since $\mathbb{1}_{\{\mathbf{V} \leq \tau\mathbb{1}\}} = \mathbb{1} - \mathbf{k}$ if $\mathbf{V} \in \mathbb{K}^n(\mathbf{k})$, these putative equilibria satisfy the fixed point equations

$$\mathbf{V}^0(\mathbf{k}) = \big((e-\beta)\mathbb{1} + \beta\mathbf{k}\big) + \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\mathbf{V}^0(\mathbf{k}),$$
$$\mathbf{V}(\mathbf{k}) = \big((e-\beta)\mathbb{1} + \beta\mathbf{k}\big) + \mathbf{C}\mathbf{V}(\mathbf{k}).$$

Subtracting the two equations and rearranging terms leads to the identity

$$(\mathbf{I} - \mathbf{C})\mathbf{V}(\mathbf{k}) - \big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)\mathbf{V}^0(\mathbf{k}) = \mathbb{0}. \tag{55}$$

We prepare for the proofs of Theorems 4 and 5 by a perturbation analysis of the separation between the putative equilibria.

Introduce notation for the *gap* between the two putative equilibria: write

$$\mathbf{\Delta} = \mathbf{\Delta}(\mathbf{k}) := \mathbf{V}(\mathbf{k}) - \mathbf{V}^0(\mathbf{k}). \tag{56}$$

In terms of the gap, the left-hand side of (55) becomes

$$\big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal - \mathbf{\Lambda}\big)\big(\mathbf{V}^0(\mathbf{k}) + \mathbf{\Delta}(\mathbf{k})\big) - \big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)\mathbf{V}^0(\mathbf{k}) = -\mathbf{\Lambda}\mathbf{V}^0(\mathbf{k}) + \big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal - \mathbf{\Lambda}\big)\mathbf{\Delta}(\mathbf{k}).$$

It follows that

$$\big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal - \mathbf{\Lambda}\big)\mathbf{\Delta}(\mathbf{k}) = \mathbf{\Lambda}\mathbf{V}^0(\mathbf{k}). \tag{57}$$

An analytical characterization of the gap between the solutions is now within reach.

LEMMA 19: *The gap may be represented in terms of a convergent Neumann series:*

$$\boldsymbol{\Delta}(\mathbf{k}) = \sum_{l=0}^{\infty} \boldsymbol{\Lambda}^l \cdot \boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k}). \tag{58}$$

PROOF: The matrix $\mathbf{I} - \mathbf{C} = \mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal - \boldsymbol{\Lambda}$ is invertible in view of Lemma 1. We may hence solve (57) explicitly for the gap and obtain

$$\begin{aligned}
\boldsymbol{\Delta}(\mathbf{k}) &= \left(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal - \boldsymbol{\Lambda}\right)^{-1}\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k}) \\
&= \left[\left(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\right)\left(\mathbf{I} - \left(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\right)^{-1}\boldsymbol{\Lambda}\right)\right]^{-1}\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k}) \\
&= \left(\mathbf{I} - \left(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\right)^{-1}\boldsymbol{\Lambda}\right)^{-1}\left(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\right)^{-1}\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k}) \\
&= \left[\mathbf{I} - \boldsymbol{\Lambda}\right]^{-1}\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})
\end{aligned}$$

by two applications of Lemma 17. Lemma 18 finishes up as $(\mathbf{I} - \boldsymbol{\Lambda})^{-1}$ has a convergent Neumann series. ▶

In our representation of the gap (58) we have resisted the temptation to absorb the dangling $\boldsymbol{\Lambda}$ in the term $\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})$ on the right into the Neumann series: a too early consolidation of matrix products loses sight of the key idea that $\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})$ is already highly concentrated near zero: indeed,

$$\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k}) = \sum_{j=1}^{n} \boldsymbol{\Lambda}_{*j} V_j^0(\mathbf{k}),$$

and, as the columns of $\boldsymbol{\Lambda}$ are independent and properly centered, $\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})$ is a centered, vector random walk. Keeping this term separate from the Neumann series allows us to exploit this concentration.

We wish to characterize the component-wise behavior of the gap $\boldsymbol{\Delta} = (\Delta_i)_{1 \le i \le n}$. The final piece in our preparation finesses difficulties with the ℓ^∞-norm by replacing it with the ℓ^2-norm to better exploit the inner product structure of the Hilbert space.

LEMMA 20: *For each solvency indicator* $\mathbf{k} \in \{0,1\}^n$,

$$\left\|\boldsymbol{\Delta}(\mathbf{k})\right\|_\infty \le \left\|\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_\infty + \sum_{l=1}^{\infty} \|\boldsymbol{\Lambda}\|_2^l \cdot \left\|\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_2. \tag{59}$$

PROOF: Beginning with (58), by separating the first term in the Neumann series from the rest of the terms, the triangle inequality shows that

$$\left\|\boldsymbol{\Delta}(\mathbf{k})\right\|_\infty \le \left\|\boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_\infty + \left\|\sum_{l=1}^{\infty} \boldsymbol{\Lambda}^l \cdot \boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_\infty.$$

The following string of inequalities for the second term on the right is now almost self-explanatory:

$$\left\|\sum_{l=1}^{\infty} \boldsymbol{\Lambda}^l \cdot \boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_\infty \overset{\text{(i)}}{\le} \left\|\sum_{l=1}^{\infty} \boldsymbol{\Lambda}^l \cdot \boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_2 \overset{\text{(ii)}}{\le} \sum_{l=1}^{\infty} \left\|\boldsymbol{\Lambda}^l \cdot \boldsymbol{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_2$$

$$\stackrel{\text{(iii)}}{\le} \sum_{l=1}^{\infty} \left\|\mathbf{\Lambda}^{l}\right\|_2 \cdot \left\|\mathbf{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_2$$

$$\stackrel{\text{(iv)}}{\le} \sum_{l=1}^{\infty} \|\mathbf{\Lambda}\|_2^{l} \cdot \left\|\mathbf{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_2.$$

The step marked: (i) follows by the monotonicity of the ℓ^p-norms, $\|\mathbf{x}\|_\infty \le \|\mathbf{x}\|_2$; (ii) follows by sub-additivity of matrix norms; (iii) follows as the operator spectral norm is consistent with the inducing ℓ^2-vector norm; and (iv) follows by sub-multiplicativity of matrix norms. ▶

D.3. *Concentration of Measure in Random Cliques*

We have thus far not yet exploited the probabilistic skeleton of the spacings: the negatively correlated structure of the spacings illuminated in Lemma 4 provides the framework for an assault on the gap; the asymptotically diffuse nature of the spacings as outlined in Definition 2 punctuates the approach by providing concentration.

Our workhorse is a concentration of measure inequality for sums of independent, centered, bounded norm random matrices: the formulation is a matrix version of Bernstein's classical inequality [c.f. Tropp (2015, Theorem 6.6.1)].

THEOREM [BERNSTEIN'S INEQUALITY FOR MATRICES]: *Suppose $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ is a sequence of independent Hermitian matrices of order d satisfying*

$$\mathbf{E}(\mathbf{Z}_i) = \mathbb{0} \textit{ and } \|\mathbf{Z}_i\|_2 \le \mathcal{L} \qquad (1 \le i \le n).$$

Write $\mathbf{S} = \sum_i \mathbf{Z}_i$ and let $\mathcal{V}(\mathbf{S})$ denote the matrix variance statistic of the sum:

$$\mathcal{V}(\mathbf{S}) = \left\|\mathbf{E}(\mathbf{S}^2)\right\|_2 = \left\|\sum\nolimits_i \mathbf{E}(\mathbf{Z}_i^2)\right\|_2.$$

Then, for all $t \ge 0$,

$$\mathbf{P}\left\{\|\mathbf{S}\|_2 \ge t\right\} \le 2d\exp\left(\frac{-t^2/2}{\mathcal{V}(\mathbf{S}) + \mathcal{L}t/3}\right) \le 2d\exp\left(-\frac{1}{4}\min\left\{\frac{t^2}{\mathcal{V}(\mathbf{S})}, \frac{t}{\mathcal{L}}\right\}\right). \tag{60}$$

The first of the inequalities in (60) is the classical Bernstein formulation; the second simplifies algebra at the cost of a slight loosening of the bound via the simple observation that, for positive a and b,

$$\frac{1}{a+b} \ge \min\left\{\frac{1}{2a}, \frac{1}{2b}\right\}.$$

When $d = 1$ we recover the usual scalar version of Bernstein's inequality. When $d = n$ we have our setting of a triangular array.

The expression on the right in (60) is particularly illuminating. It shows that the tail bound for the norm of a random sum of centered, independent matrices with bounded spectrum consists of a mixture of a sub-Gaussian term manifested in the exponent $t^2/\mathcal{V}(\mathbf{S})$ and a sub-exponential term manifested in the exponent $t/\mathcal{L}$.

D.3.0.1. A HIGH PROBABILITY SET Recall that the centered cross-holdings matrix is given by

$$\mathbf{\Lambda} = [\Lambda_{ji}]_{1\le j,i\le n} = \big[\mathbf{\Lambda}_{*i}\big]_{1\le i\le n} = c\big[\mathbf{X}_i - \tfrac{1}{n}\mathbb{1}\big]_{1\le i\le n} \tag{61}$$

where the columns are independent and asymptotically diffuse [Definition 2] and each element is appropriately centered with $\Lambda_{ji} = c\big(X_{ji} - \frac{1}{n}\big)$.

A naïve approach to bounding the gap (58) by attempting to exploit the fact that the centered spacings are bounded, hence tritely sub-exponential, falters because the sub-exponential bound is much too loose. A more subtle approach is to exploit the fact that the spacings cannot get pathologically large. We have selected the asymptotic rates *ex post facto* to optimize the bounds.

With $\mathfrak{a}$ a positive constant to be selected later, identify the high probability measurable set[10]

$$\mathcal{Q} = \mathcal{Q}_n(\mathfrak{a}) := \bigcap_{j,i}\left\{\left|X_{ji} - \frac{1}{n}\right| < \frac{1}{\mathfrak{a}\sqrt{n}\log(n)}\right\}. \tag{62}$$

The reason for the nomenclature becomes clear by an application of Chebyshev's moment inequality.

LEMMA 21: *For every choice of* $\mathfrak{a} > 0$, *there exists a constant* K_1 *determined by* $\mathfrak{a}$ *such that*

$$\mathbf{P}\big(\mathcal{Q}^{\complement}\big) = \mathbf{P}\big(\mathcal{Q}_n(\mathfrak{a})^{\complement}\big) \le \frac{K_1 \log(n)^8}{n^2}$$

for all $n \ge 2$.

PROOF: As before, write $\mu_n(\nu) = \mathbf{E}\big\{\big(X_{ji} - \frac{1}{n}\big)^{\nu}\big\}$ for the νth central moment of the spacings. Then

$$\mathbf{P}\big(\mathcal{Q}^{\complement}\big) \overset{\text{(i)}}{\le} n^2 \max_{j,i} \mathbf{P}\left\{\left|X_{ji} - \frac{1}{n}\right| \ge \frac{1}{\mathfrak{a}\sqrt{n}\log(n)}\right\} \overset{\text{(ii)}}{\le} n^2 \cdot \frac{\mu_n(8)}{\Big(\dfrac{1}{\mathfrak{a}\sqrt{n}\log(n)}\Big)^8}$$
$$\overset{\text{(iii)}}{\le} (A\mathfrak{a})^8 \cdot \frac{\log(n)^8}{n^2}.$$

Step (i) follows by Boole's inequality; (ii) is via Chebyshev's inequality in its eighth power incarnation (c.f. Venkatesh (2013, Theorem XVI.1.1 and Problem XVI.13.11)); and (iii) is a consequence of the asymptotically diffuse nature of the spacings [Definition 2]. We may identify $K_1 = (A\mathfrak{a})^8$ with $\mathfrak{a}$ to be specified. ▶

Faster concentration obtains for spacings with a sub-exponential character satisfying Lemma 7. This is the case for the de Finetti and Haar-induced settings. In such cases, step (ii) in the proof is replaced by the much stronger exponential bound

$$\alpha n^2 \exp\left(\frac{-\kappa n^{1/2}}{\mathfrak{a}\log(n)}\right)$$

and concentration is very rapid.

[10] The result outlined in Section 10 for an Erdös–Rényi topology can be obtained by repeating the argument here by first conditioning on the event that every firm has out-degree $n(p \pm \varepsilon)$ with high probability.

As the variance of the spacings for an asymptotically diffuse distribution is $\mathcal{O}(n^{-2})$, it is clear that each spacing has a deviation approximately of order $\mathcal{O}(n^{-1})$ from its mean n^{-1}; however, with the large number of spacings in view, the chance of a significant number of spacings deviating substantially from the mean is high. The eighth moment structure of an asymptotically diffuse distribution allows us to conclude that *all* n^2 spacings are loosely concentrated around their means: "looseness" here means of order $n^{-1/2}/\log(n)$ which, of course, is much more collectively dispersive than what the individual standard deviation of order n^{-1} would have told us. This, as we shall see, is the critical range that we have to control: the asymptotics are most delicately balanced.

D.3.0.2. CONCENTRATION OF THE RANDOM WALK $\mathbf{\Lambda V}^0(\mathbf{k})$ We now turn our attention to bounding the ℓ^∞-norm of the random walk $\mathbf{\Lambda V}^0(\mathbf{k})$ on the right in (58). This will provide a platform for estimating the ℓ^∞-norm of the gap using (59). The choice of constants and rates has again been fine-tuned *ex post facto* so, without further apology, we will introduce another positive constant $\mathfrak{b}$ to be selected appropriately later and seek to estimate

$$\mathbf{P}\left\{\left\|\mathbf{\Lambda V}^0(\mathbf{k})\right\|_\infty < \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}}\right\},$$

the probability that the components of the random walk do not stray too far from the origin. The notion of "far" from the origin is captured in the mysterious deviation term $\mathfrak{b}\sqrt{\log(n)}/\sqrt{n}$.

Unburden notation by writing $\mathbf{V}^0 = \mathbf{V}^0(\mathbf{k})$ and keeping the dependence on the solvency indicator $\mathbf{k}$ implicit. We break the analysis down into stages.

1° *Conditioning on the high probability set.* Most of the action takes place in the high probability set $\mathcal{Q} = \mathcal{Q}_n(\mathfrak{a})$. We will again keep notation sane by hiding its dependence on n and $\mathfrak{a}$. Accordingly, by conditioning on $\mathcal{Q}$, we obtain

$$\begin{aligned}\mathbf{P}\left\{\left\|\mathbf{\Lambda V}^0\right\|_\infty \geq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}}\right\} &\leq \mathbf{P}\left\{\left\|\mathbf{\Lambda V}^0\right\|_\infty \geq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} \,\middle|\, \mathcal{Q}\right\} + \mathbf{P}(\mathcal{Q}^\complement)\\ &\leq \mathbf{P}\left\{\left\|\mathbf{\Lambda V}^0\right\|_\infty \geq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} \,\middle|\, \mathcal{Q}\right\} + \frac{K_1\log(n)^8}{n^2} \qquad (63)\end{aligned}$$

by Lemma 21. Now, by Boole's inequality,

$$\begin{aligned}\mathbf{P}\left\{\left\|\mathbf{\Lambda V}^0\right\|_\infty \geq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} \,\middle|\, \mathcal{Q}\right\} &= \mathbf{P}\left\{\max_{1\leq j\leq n}\left|\sum_{i=1}^n \Lambda_{ji}V_i^0\right| \geq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} \,\middle|\, \mathcal{Q}\right\}\\ &\leq n\max_j \mathbf{P}\left\{\left|\sum_{i=1}^n \Lambda_{ji}V_i^0\right| \geq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} \,\middle|\, \mathcal{Q}\right\}. \qquad (64)\end{aligned}$$

Our game plan is to use Bernstein's inequality to estimate the conditional probability on the right. We will need to prepare the ground by re-centering the summands.

Fix any index j and, promptly hiding it from view to further unburden notation, introduce the nonce notation

$$Z_i = \big(\Lambda_{ji} - \mathbf{E}(\Lambda_{ji} \mid \mathcal{Q})\big)V_i^0 \qquad (1 \leq i \leq n). \qquad (65)$$

As the columns of the matrix of spacings $\mathbf{X} = \bigl[\mathbf{X}_i\bigr]_{1\le i\le n}$ are independent, the sets

$$\mathcal{Q}_i := \bigcap_j \biggl\{ \biggl| X_{ji} - \frac{1}{n} \biggr| < \frac{1}{\mathfrak{a}\sqrt{n}\log(n)} \biggr\} \qquad (1 \le i \le n)$$

are independent. As $\mathcal{Q} = \bigcap_i \mathcal{Q}_i$, it follows that the spacings X_{ji} $(1 \le i \le n)$, hence also the centered cross-holdings Λ_{ji} $(1 \le i \le n)$, are conditionally independent given $\mathcal{Q}$, the symmetry in the specification of the component sets $\mathcal{Q}_i$ ensuring that they also share a common conditional distribution. We conclude that, conditioned on $\mathcal{Q}$, the variables $Z_1, \dots, Z_n$ are conditionally independent and centered, and share a common conditional distribution. Now form the conditionally centered sum

$$\sum_{i=1}^n Z_i = \sum_{i=1}^n \Lambda_{ji} V_i^0 - \mathbf{E}\biggl(\sum_{i=1}^n \Lambda_{ji} V_i^0 \Bigm| \mathcal{Q} \biggr).$$

By the triangle inequality,

$$\biggl| \sum_{i=1}^n \Lambda_{ji} V_i^0 \biggr| \le \biggl| \sum_{i=1}^n Z_i \biggr| + \biggl| \mathbf{E}\biggl(\sum_{i=1}^n \Lambda_{ji} V_i^0 \Bigm| \mathcal{Q} \biggr) \biggr| \le \biggl| \sum_{i=1}^n Z_i \biggr| + \sum_{i=1}^n \bigl| \mathbf{E}(\Lambda_{ji} \mid \mathcal{Q}) \bigr| \cdot \bigl| V_i^0 \bigr|.$$

Introducing one more temporary piece of notation to aid visual clarity, write

$$\tau_n := \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} - \sum_{i=1}^n \bigl| \mathbf{E}(\Lambda_{ji} \mid \mathcal{Q}) \bigr| \cdot \bigl| V_i^0 \bigr| \tag{66}$$

and, after a proper conditional centering of the sum, bound the conditional probability on the right in (64) by

$$\mathbf{P}\biggl\{ \biggl| \sum_{i=1}^n \Lambda_{ji} V_i^0 \biggr| \ge \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} \Bigm| \mathcal{Q} \biggr\} \le \mathbf{P}\biggl\{ \biggl| \sum_{i=1}^n Z_i \biggr| \ge \tau_n \Bigm| \mathcal{Q} \biggr\}. \tag{67}$$

The table is set for Bernstein's inequality (60) specialized to the scalar case $d = 1$. In order to deploy it we will need to estimate the variance statistic and the bounded range which contribute to the sub-Gaussian and sub-exponential components of the inequality, respectively. The conditional expectation is the first step.

2° *Estimating the conditional expectation and the deviation from the mean.* In the usual notation, we write $1_{\mathcal{Q}}$ and $1_{\mathcal{Q}^\complement}$ for the indicator variables for the sets $\mathcal{Q}$ and $\mathcal{Q}^\complement$, respectively. As $\mathcal{Q}$ and $\mathcal{Q}^\complement$ partition the space, we may write $\Lambda_{ji} = \Lambda_{ji} 1_{\mathcal{Q}} + \Lambda_{ji} 1_{\mathcal{Q}^\complement}$. By virtue of the centering of the spacings, it follows by additivity that

$$0 = \mathbf{E}(\Lambda_{ji}) = \mathbf{E}(\Lambda_{ji} 1_{\mathcal{Q}}) + \mathbf{E}(\Lambda_{ji} 1_{\mathcal{Q}^\complement}), \text{ or, equivalently, } \mathbf{E}(\Lambda_{ji} 1_{\mathcal{Q}}) = -\mathbf{E}(\Lambda_{ji} 1_{\mathcal{Q}^\complement}).$$

As $\mathbf{E}(\Lambda_{ji} \mid \mathcal{Q}) = \frac{1}{\mathbf{P}(\mathcal{Q})} \mathbf{E}(\Lambda_{ji} 1_{\mathcal{Q}})$, we obtain

$$\bigl| \mathbf{E}(\Lambda_{ji} \mid \mathcal{Q}) \bigr| = \frac{\bigl| \mathbf{E}(\Lambda_{ji} 1_{\mathcal{Q}}) \bigr|}{\mathbf{P}(\mathcal{Q})} = \frac{\bigl| -\mathbf{E}(\Lambda_{ji} 1_{\mathcal{Q}^\complement}) \bigr|}{1 - \mathbf{P}(\mathcal{Q}^\complement)} \le \frac{\mathbf{E}\bigl(|\Lambda_{ji}| 1_{\mathcal{Q}^\complement} \bigr)}{1 - \mathbf{P}(\mathcal{Q}^\complement)} \le \frac{\mathbf{E}(1_{\mathcal{Q}^\complement})}{1 - \mathbf{P}(\mathcal{Q}^\complement)} = \frac{\mathbf{P}(\mathcal{Q}^\complement)}{1 - \mathbf{P}(\mathcal{Q}^\complement)},$$

the penultimate step a consequence of Lemma 16 in view of the crude bound $|\Lambda_{ji}| < 1$. The tail bound probability from Lemma 21 allows us to conclude that

$$\bigl|\mathbf{E}(\Lambda_{ji} \mid \mathcal{Q})\bigr| \le \frac{\dfrac{K_1 \log(n)^8}{n^2}}{1 - \dfrac{K_1 \log(n)^8}{n^2}} = \frac{K_1 \log(n)^8}{n^2}\left[1 + \mathcal{O}\left(\frac{\log(n)^8}{n^2}\right)\right]. \tag{68}$$

The putative equilibria for a regular clique are given in (33) and in view of (11) they may be uniformly absolutely bounded by

$$\bigl|V_i^0\bigr| \le \max_s |v(s)| \le \frac{\max\{e, \beta\}}{1-c}. \tag{69}$$

Putting these observations together, we conclude that

$$\sum_{i=1}^{n} \bigl|\mathbf{E}(\Lambda_{ji} \mid \mathcal{Q})\bigr| \cdot \bigl|V_i^0\bigr| \le \frac{\max\{e, \beta\}}{1-c} \cdot n \cdot \frac{K_1 \log(n)^8}{n^2}\left[1 + \mathcal{O}\left(\frac{\log(n)^8}{n^2}\right)\right] = \mathcal{O}\left(\frac{\log(n)^8}{n}\right).$$

The contribution of the conditional expectation is hence sub-dominant compared to the term $\mathfrak{b}\sqrt{\log(n)}/\sqrt{n}$ on the right in (66) and so the conditional deviation from the mean on the right in (67) is of asymptotic order

$$\tau_n = \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}}\left[1 - \mathcal{O}\left(\frac{\log(n)^{15/2}}{\sqrt{n}}\right)\right] \sim \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}}. \tag{70}$$

3° *Estimating the conditional variance.* The variables Z_i given in (65) are properly centered conditioned on $\mathcal{Q}$. Recalling that $\Lambda_{ji} = c\bigl(X_{ji} - \frac{1}{n}\bigr)$, we have

$$\begin{aligned}
\mathbf{E}\bigl(Z_i^2 \mid \mathcal{Q}\bigr) = \bigl(V_i^0\bigr)^2 \cdot \mathbf{E}\Bigl[\bigl(\Lambda_{ji} - \mathbf{E}(\Lambda_{ji} \mid \mathcal{Q})\bigr)^2 \Bigm| \mathcal{Q}\Bigr] &\le \bigl(V_i^0\bigr)^2 \cdot \mathbf{E}\bigl(\Lambda_{ji}^2 \mid \mathcal{Q}\bigr) \\
&= \bigl(V_i^0\bigr)^2 \cdot \frac{\mathbf{E}(\Lambda_{ji}^2 1_{\mathcal{Q}})}{\mathbf{P}(\mathcal{Q})} \\
&\le \bigl(V_i^0\bigr)^2 \cdot \frac{\mathbf{E}(\Lambda_{ji}^2)}{\mathbf{P}(\mathcal{Q})} \\
&= \bigl(cV_i^0\bigr)^2 \cdot \frac{\mathbf{E}\Bigl[\bigl(X_{ji} - \frac{1}{n}\bigr)^2\Bigr]}{1 - \mathbf{P}\bigl(\mathcal{Q}^\complement\bigr)}.
\end{aligned}$$

Leveraging the asymptotically diffuse condition (20), Lemma 21, and (69), we obtain

$$\mathbf{E}\bigl(Z_i^2 \mid \mathcal{Q}\bigr) \le \left(\frac{c\max\{e, \beta\}}{1-c}\right)^2 \cdot \frac{\dfrac{A^2}{n^2}}{1 - \dfrac{K_1 \log(n)^8}{n^2}} = \left(\frac{c\max\{e, \beta\}A}{1-c}\right)^2 \cdot \frac{1}{n^2} \cdot \left[1 + \mathcal{O}\left(\frac{\log(n)^8}{n^2}\right)\right].$$

Conditioned on $\mathcal{Q}$, the variables $Z_1, \dots, Z_n$ are conditionally independent—whence the conditional variance is additive—and share a common distribution—*a fortiori* are exchangeable. The conditional variance statistic that contributes to the sub-Gaussian term in Bernstein's inequality may hence be bounded by

$$\begin{aligned}\mathcal{V}\Bigl(\sum_{i=1}^{n} Z_i \Bigm| \mathcal{Q}\Bigr) := \mathbf{E}\Bigl[\Bigl(\sum_{i=1}^{n} Z_i\Bigr)^2 \Bigm| \mathcal{Q}\Bigr] &= \sum_{i=1}^{n} \mathbf{E}\bigl(Z_i^2 \mid \mathcal{Q}\bigr)\\ &= n \cdot \mathbf{E}(Z_1^2 \mid \mathcal{Q})\\ &\leq \left(\frac{c\max\{e,\beta\}A}{1-c}\right)^2 \cdot \frac{1}{n} \cdot \left[1 + \mathcal{O}\left(\frac{\log(n)^8}{n^2}\right)\right].\end{aligned} \tag{71}$$

4° *Estimating the bounded range over $\mathcal{Q}$.* By the triangle inequality,

$$|Z_i| \leq \bigl|V_i^0\bigr| \cdot \bigl(c\bigl|X_{ji} - \tfrac{1}{n}\bigr| + \bigl|\mathbf{E}(\Lambda_{ji} \mid \mathcal{Q})\bigr|\bigr).$$

The first term on the right is uniformly bounded by (69). Of the two terms inside the round brackets, the absolute value of the centered spacing $\bigl|X_{ji} - \frac{1}{n}\bigr|$ is uniformly controlled over the high probability set $\mathcal{Q}$ by (62) while the conditional expectation $\mathbf{E}(\Lambda_{ji} \mid \mathcal{Q})$ is sub-dominant with asymptotic character given by (68). Stitching the pieces together, on the high probability set $\mathcal{Q}$,

$$\begin{aligned}\max_i |Z_i| &\leq \frac{\max\{e,\beta\}}{1-c} \cdot \left[\frac{c}{\mathfrak{a}\sqrt{n}\log(n)} + \frac{K_1 \log(n)^8}{n^2}\left\{1 + \mathcal{O}\left(\frac{\log(n)^8}{n^2}\right)\right\}\right]\\ &= \frac{c\max\{e,\beta\}}{(1-c)\mathfrak{a}} \cdot \frac{1}{\sqrt{n}\log(n)}\left[1 + \mathcal{O}\left(\frac{\log(n)^9}{n^{3/2}}\right)\right] =: \mathcal{L}.\end{aligned} \tag{72}$$

5° *Bernstein's inequality.* All the pieces are in hand: we apply the (conditional) scalar version of Bernstein's inequality (60) to the probability on the right in (67) to obtain

$$\mathbf{P}\left\{\left|\sum_{i=1}^{n} Z_i\right| \geq \tau_n \Bigm| \mathcal{Q}\right\} \leq 2\exp\left(-\frac{1}{4}\min\left\{\frac{\tau_n^2}{\mathcal{V}\bigl(\sum_{i=1}^{n} Z_i \mid \mathcal{Q}\bigr)}, \frac{\tau_n}{\mathcal{L}}\right\}\right). \tag{73}$$

The sub-Gaussian term: applying the estimates (70) and (71),

$$\begin{aligned}\frac{\tau_n^2}{\mathcal{V}\bigl(\sum_{i=1}^{n} Z_i \mid \mathcal{Q}\bigr)} &\geq \frac{\left[\frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}}\left\{1 - \mathcal{O}\left(\frac{\log(n)^{15/2}}{\sqrt{n}}\right)\right\}\right]^2}{\left(\frac{c\max\{e,\beta\}A}{1-c}\right)^2 \cdot \frac{1}{n}\left\{1 + \mathcal{O}\left(\frac{\log(n)^8}{n^2}\right)\right\}}\\ &= \left(\frac{\mathfrak{b}(1-c)}{c\max\{e,\beta\}A}\right)^2 \cdot \log(n) \cdot \left[1 - \mathcal{O}\left(\frac{\log(n)^{15/2}}{\sqrt{n}}\right)\right].\end{aligned}$$

The term in square brackets on the right $\to 1$ as $n \to \infty$, and so the right-hand side has a logarithmic growth rate in n. *It now becomes apparent how we should select the constant $\mathfrak{b}$. Looking ahead with a view to obtaining a roughly quadratic decay in the probability in* (63)*,* *select* $\mathfrak{b}$ *to be any positive constant satisfying*

$$\mathfrak{b} > \frac{\sqrt{12}\,c\max\{e,\beta\}A}{1-c}. \tag{74}$$

With this choice we conclude that the sub-Gaussian exponent is bounded below by

$$\frac{\tau_n^2}{\mathcal{V}\Big(\sum_{i=1}^{n} Z_i \mid \mathcal{Q}\Big)} \geq 12\log(n) \qquad \text{(for all sufficiently large } n\text{).}$$

The smallest value $n = n_0$ for which this is true depends on the choices of both $\mathfrak{a}$ and $\mathfrak{b}$ but we are interested in the asymptotics and the assertion holds for all $n \geq n_0$.

The sub-exponential term: now applying the estimates (70) and (72),

$$\begin{aligned}
\frac{\tau_n}{\mathcal{L}} &\geq \frac{\dfrac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}}\left\{1-\mathcal{O}\Big(\dfrac{\log(n)^{15/2}}{\sqrt{n}}\Big)\right\}}{\dfrac{c\max\{e,\beta\}}{(1-c)\mathfrak{a}}\cdot\dfrac{1}{\sqrt{n}\,\log(n)}\left\{1+\mathcal{O}\Big(\dfrac{\log(n)^9}{n^{3/2}}\Big)\right\}}\\
&= \frac{\mathfrak{a}\mathfrak{b}(1-c)}{c\max\{e,\beta\}}\cdot\log(n)^{3/2}\cdot\left[1-\mathcal{O}\left(\frac{\log(n)^{15/2}}{\sqrt{n}}\right)\right]\\
&\geq \sqrt{12}\,A\mathfrak{a}\cdot\log(n)^{3/2}\cdot\left[1-\mathcal{O}\left(\frac{\log(n)^{15/2}}{\sqrt{n}}\right)\right]
\end{aligned}$$

in view of the selection (74). Again, the term in square brackets $\to 1$ as $n \to \infty$ so that the right-hand side has a growth rate of order $\log(n)^{3/2}$. We can afford to be cavalier with our bounds and conclude that

$$\frac{\tau_n}{\mathcal{L}} \geq A\mathfrak{a}\log(n)^{3/2} \qquad \text{(for all sufficiently large } n\text{).}$$

Again, the asymptotics kick in for all $n \geq n_1$ for some n_1 depending on the choices of $\mathfrak{a}$ and $\mathfrak{b}$.

As we may have anticipated because we are dealing with the tail of a random walk, we conclude that the tail behavior is sub-Gaussian—the sub-exponential exponent growth rate eventually outstrips the sub-Gaussian exponent growth rate. Let n_2 be the smallest value of n so that $A\mathfrak{a}\log(n)^{3/2} \geq 12\log(n)$. It follows that, for all $n \geq \max\{n_0, n_1, n_2\}$, the inequality (73) simplifies to

$$\mathbf{P}\left\{\left|\sum_{i=1}^{n} Z_i\right| \geq \tau_n \,\middle|\, \mathcal{Q}\right\} \leq 2\exp\left(-\frac{1}{4}\min\big\{12\log(n), A\mathfrak{a}\log(n)^{3/2}\big\}\right) = 2e^{-3\log(n)} = \frac{2}{n^3} \tag{75}$$

eventually, the estimate uniform in the hidden index j. Tracing the thread back via (75, 67, 64) to (63), we conclude that

$$\mathbf{P}\left\{\|\mathbf{\Lambda}\mathbf{V}^0\|_\infty \geq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}}\right\} \leq n\cdot\frac{2}{n^3} + \frac{K_1\log(n)^8}{n^2} = \frac{2}{n^2} + \frac{K_1\log(n)^8}{n^2}$$

eventually for all sufficiently large n. We encapsulate our findings.

LEMMA 22: *Suppose* $\mathfrak{a}$ *is positive and* $\mathfrak{b}$ *satisfies* (74). *Then, there exists a constant* K_2 *determined by* $\mathfrak{a}$ *and* $\mathfrak{b}$ *such that*

$$\mathbf{P}\left\{\left\|\mathbf{\Lambda}\mathbf{V}^0(\mathbf{k})\right\|_\infty \geq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}}\right\} \leq \frac{K_2\log(n)^8}{n^2}$$

eventually, for all sufficiently large values of n.

The asymptotics kick in beyond a horizon determined by the selections of $\mathfrak{a}$ and $\mathfrak{b}$, while our development shows that we may generously take $K_2 = K_1 + 2 = (A\mathfrak{a})^8 + 2$. We will resist the temptation to attempt to optimize the results over $\mathfrak{a}$ at this stage.

D.3.0.3. CONCENTRATION OF THE LARGEST SINGULAR VALUE OF $\mathbf{\Lambda}$ We now turn our attention to bounding the spectral norm of the centered matrix of cross-holdings. This will be the final step in the estimation of the ℓ^∞-norm of the gap using (59).

We have a few more technical complexities to contend with but our main line of attack follows the line of the proof of Lemma 22 in that we reduce the problem to that of concentration of a random matrix walk. The key is to exploit the fact that the centered cross-holdings matrix

$$\mathbf{\Lambda} = \left[\mathbf{\Lambda}_{*1}\ \mathbf{\Lambda}_{*2} \cdots \mathbf{\Lambda}_{*n}\right] = c\left[\mathbf{X}_1 - \tfrac{1}{n}\mathbb{1}\ \mathbf{X}_2 - \tfrac{1}{n}\mathbb{1} \cdots \mathbf{X}_n - \tfrac{1}{n}\mathbb{1}\right]$$

has independent columns via the useful observation that the spectral norm is invariant under transposition. Accordingly,

$$\|\mathbf{\Lambda}\|_2 = \|\mathbf{\Lambda}^\intercal\|_2 = \sqrt{\|\mathbf{\Lambda}\mathbf{\Lambda}^\intercal\|_2} = \sqrt{\left\|\sum_{i=1}^n \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal\right\|_2}\text{; equivalently, } \|\mathbf{\Lambda}\|_2^2 = \left\|\sum_{i=1}^n \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal\right\|_2. \tag{76}$$

With $\mathfrak{b}$ selected in accordance with (74) and ε a tiny positive quantity which may be selected arbitrarily small, we seek now to estimate,

$$\mathbf{P}\left\{\|\mathbf{\Lambda}\|_2 < \frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}}\right\},$$

the probability that the spectral norm of $\mathbf{\Lambda}$ is concentrated near the origin. The identification of what "near" means has been carefully calibrated but is admittedly mysterious—the rates are glacially slow compared to our usual ideas of concentration; the selection becomes a little more transparent if Lemma 22 is kept in mind. We again proceed in stages.

1° *Conditioning on the high probability set.* A naïve bounding of the centered cross-holdings via Lemma 16 fails as the bound does not adequately capture the picture of the loose concentration of cross-holdings conjured up by Lemma 21. We accordingly follow our well-worn path by conditioning on the high probability set $\mathcal{Q}$. Invoking additivity, again,

$$\begin{aligned}\mathbf{P}\left\{\|\mathbf{\Lambda}\|_2 \geq \frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}}\right\} &\leq \mathbf{P}\left\{\|\mathbf{\Lambda}\|_2 \geq \frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}} \;\middle|\; \mathcal{Q}\right\} + \mathbf{P}\left(\mathcal{Q}^\complement\right) \qquad (77)\\ &\leq \mathbf{P}\left\{\|\mathbf{\Lambda}\|_2 \geq \frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}} \;\middle|\; \mathcal{Q}\right\} + \frac{K_1\log(n)^8}{n^2}\end{aligned}$$

by Lemma 21. In view of (76), the conditional probability on the right in (77) can be re-expressed in the form

$$\mathbf{P}\left\{\|\mathbf{\Lambda}\|_2 \geq \frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}} \;\middle|\; \mathcal{Q}\right\} = \mathbf{P}\left\{\sqrt{\left\|\sum_{i=1}^{n} \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^{\intercal}\right\|_2} \geq \frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}} \;\middle|\; \mathcal{Q}\right\} \tag{78}$$
$$= \mathbf{P}\left\{\left\|\sum_{i=1}^{n} \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^{\intercal}\right\|_2 \geq \frac{\varepsilon^2}{\mathfrak{b}^2\log(n)} \;\middle|\; \mathcal{Q}\right\}.$$

On the right-hand side we have a sum of independent matrices with a common distribution and the conditioning suggests that we center them appropriately. Introduce notation for the common matrix of means

$$\mathbf{M}_{\mathcal{Q}} := \mathbf{E}\big(\mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^{\intercal} \mid \mathcal{Q}\big), \tag{79}$$

and the corresponding conditionally centered matrices

$$\mathbf{Z}_i := \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^{\intercal} - \mathbf{M}_{\mathcal{Q}} \qquad (1 \leq i \leq n). \tag{80}$$

Recall that the sets

$$\mathcal{Q}_i = \bigcap_j \left\{\left|X_{ji} - \frac{1}{n}\right| < \frac{1}{\mathfrak{a}\sqrt{n}\log(n)}\right\} \qquad (1 \leq i \leq n)$$

are independent and $\mathcal{Q} = \bigcap_i \mathcal{Q}_i$. For each i, the (column) vector $\mathbf{\Lambda}_{*i} = c\big[\mathbf{X}_i - \frac{1}{n}\mathbb{1}\big]$ is independent of $\mathcal{Q}_k$ for $k \neq i$ and so the matrices $\mathbf{Z}_1, \dots, \mathbf{Z}_n$ are conditionally independent given $\mathcal{Q}$ and, by the symmetry inherent in the specification of the sets $\mathcal{Q}_i$, have the same conditional distribution. Centering the sum on the right in (78) results in

$$\sum_{i=1}^{n} \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^{\intercal} = \sum_{i=1}^{n} \big(\mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^{\intercal} - \mathbf{M}_{\mathcal{Q}}\big) + n\mathbf{M}_{\mathcal{Q}} = \sum_{i=1}^{n} \mathbf{Z}_i + n\mathbf{M}_{\mathcal{Q}},$$

whence, by the triangle inequality, we have

$$\left\|\sum_{i=1}^{n} \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^{\intercal}\right\|_2 \leq \left\|\sum_{i=1}^{n} \mathbf{Z}_i\right\|_2 + n\big\|\mathbf{M}_{\mathcal{Q}}\big\|_2.$$

Recycle notation and write

$$\tau_n := \frac{\varepsilon^2}{\mathfrak{b}^2\log(n)} - n\big\|\mathbf{M}_{\mathcal{Q}}\big\|_2. \tag{81}$$

We may now conservatively bound the expression on the right in (78) by

$$\mathbf{P}\left\{\left\|\sum_{i=1}^{n} \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^{\intercal}\right\|_2 \geq \frac{\varepsilon^2}{\mathfrak{b}^2\log(n)} \;\middle|\; \mathcal{Q}\right\} \leq \mathbf{P}\left\{\left\|\sum_{i=1}^{n} \mathbf{Z}_i\right\|_2 \geq \tau_n \;\middle|\; \mathcal{Q}\right\} \tag{82}$$

and we are primed for another use of Bernstein's inequality—in its matrix incarnation this time.

2° *Estimating the norm of the conditional mean matrix.* In the language of Section C.1, write

$$\mu_n(\nu_1,\dots,\nu_n) := \mathbf{E}\big\{\big(X_{1i}-\tfrac{1}{n}\big)^{\nu_1}\times\cdots\times\big(X_{ni}-\tfrac{1}{n}\big)^{\nu_n}\big\},$$
$$\mu_n(\nu_1,\dots,\nu_n\mid\mathcal{Q}) := \mathbf{E}\big\{\big(X_{1i}-\tfrac{1}{n}\big)^{\nu_1}\times\cdots\times\big(X_{ni}-\tfrac{1}{n}\big)^{\nu_n}\mid\mathcal{Q}\big\}$$

for the common mixed moments and the conditional mixed moments, respectively, of the columns of the centered matrix of asymptotically diffuse spacings $\big[\mathbf{X}_i-\frac{1}{n}\mathbb{1}\big]_{1\le i\le n}$. We now identify $\mathcal{Q}=\bigcap_i\mathcal{Q}_i$ with the high probability set (62). The symmetries in the specifications of the sets $\mathcal{Q}_i$ ensure that exchangeability is preserved for the spacings under conditioning with respect to $\mathcal{Q}$ and so both $\mu_n(\nu_1,\dots,\nu_n)$ and $\mu_n(\nu_1,\dots,\nu_n\mid\mathcal{Q})$ are invariant with respect to permutations of coordinates. As before, we take a liberty with notation and, for $1\le k\le n$, also write

$$\mu_n(\nu_1,\dots,\nu_k) = \mathbf{E}\big\{\big(X_{1i}-\tfrac{1}{n}\big)^{\nu_1}\times\cdots\times\big(X_{ki}-\tfrac{1}{n}\big)^{\nu_k}\big\},$$
$$\mu_n(\nu_1,\dots,\nu_k\mid\mathcal{Q}) = \mathbf{E}\big\{\big(X_{1i}-\tfrac{1}{n}\big)^{\nu_1}\times\cdots\times\big(X_{ki}-\tfrac{1}{n}\big)^{\nu_k}\mid\mathcal{Q}\big\}.$$

With this for preparation, per the definition (79), the components of the conditional mean matrix $\mathbf{M}_{\mathcal{Q}}=[M_{jk}]_{1\le j,k\le n}$ are given by

$$M_{jk}=\mathbf{E}(\Lambda_{ji}\Lambda_{ki}\mid\mathcal{Q})=\begin{cases}c^2\mu_n(2\mid\mathcal{Q}) & \text{if } j=k,\\ c^2\mu_n(1,1\mid\mathcal{Q}) & \text{if } j\ne k.\end{cases}$$

Exploiting the fact that the spacings are negatively correlated, by Corollary 2 in Section C.1, $\mu_n(1,1\mid\mathcal{Q})=-\mu_n(2\mid\mathcal{Q})/(n-1)$. In the terminology of Appendix B we conclude that $\mathbf{M}_{\mathcal{Q}}$ is a scaled almost diagonal matrix,

$$\begin{aligned}\mathbf{M}_{\mathcal{Q}}=c^2\mu_n(2\mid\mathcal{Q})\cdot\mathbf{I}-\frac{c^2\mu_n(2\mid\mathcal{Q})}{n-1}\cdot\big(\mathbb{1}\mathbb{1}^\intercal-\mathbf{I}\big)&=\frac{c^2n}{n-1}\cdot\mu_n(2\mid\mathcal{Q})\cdot\Big(\mathbf{I}-\frac{1}{n}\mathbb{1}\mathbb{1}^\intercal\Big)\\&=\frac{c^2n}{n-1}\cdot\mu_n(2\mid\mathcal{Q})\cdot\mathbf{D}\Big(\frac{-1}{n}\Big).\end{aligned}\tag{83}$$

By identifying $\alpha=-1/n$ in Lemma 2d, we see that $\big\|\mathbf{D}\big(-n^{-1}\big)\big\|_2=1$ and so

$$\big\|\mathbf{M}_{\mathcal{Q}}\big\|_2=\frac{c^2n}{n-1}\cdot\mu_n(2\mid\mathcal{Q})\cdot\Big\|\mathbf{D}\Big(\frac{-1}{n}\Big)\Big\|_2=\frac{c^2n}{n-1}\cdot\mu_n(2\mid\mathcal{Q}).\tag{84}$$

With $1_{\mathcal{Q}}$ again representing the indicator for the set $\mathcal{Q}$, we may estimate the conditional second moment via

$$\begin{aligned}\mu_n(2\mid\mathcal{Q})=\frac{\mathbf{E}\big[\big(X_{1i}-\frac{1}{n}\big)^21_{\mathcal{Q}}\big]}{\mathbf{P}(\mathcal{Q})}\overset{\text{(i)}}{\le}\frac{\mathbf{E}\big[\big(X_{1i}-\frac{1}{n}\big)^2\big]}{\mathbf{P}(\mathcal{Q})}&=\frac{\mu_n(2)}{1-\mathbf{P}\big(\mathcal{Q}^\complement\big)}\\&\overset{\text{(ii)}}{\le}\frac{\dfrac{A^2}{n^2}}{1-\dfrac{K_1\log(n)^8}{n^2}}\\&=\frac{A^2}{n^2}\Big[1+\mathcal{O}\Big(\frac{\log(n)^8}{n^2}\Big)\Big];\end{aligned}\tag{85}$$

the step labelled (i) follows by monotonicity of expectation, while (ii) follows by appeal to the asymptotically diffuse condition (20) to bound the numerator and Lemma 21 to bound the denominator. It follows that

$$n\|\mathbf{M}_{\mathcal{Q}}\| = n \cdot \frac{c^2 n}{n-1} \cdot \mu_n(2 \mid \mathcal{Q}) = \frac{c^2 A^2}{n}\left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right]$$

and we conclude that the impact of conditional centering is sub-dominant in (81). We are hence interested in the right tail of asymptotic order

$$\tau_n = \frac{\varepsilon^2}{\mathfrak{b}^2 \log(n)} - \frac{c^2 A^2}{n}\left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right] = \frac{\varepsilon^2}{\mathfrak{b}^2 \log(n)}\left[1 - \mathcal{O}\left(\frac{\log(n)}{n}\right)\right] \tag{86}$$

of the distribution of $\|\sum_i \mathbf{Z}_i\|$.

3° *Estimating the matrix variance statistic.* The conditional covariance matrix $\mathbf{K}_{\mathcal{Q}}$ of the (conditionally centered) Hermitian form $\mathbf{Z}_i = \mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal - \mathbf{M}_{\mathcal{Q}}$ is given by

$$\mathbf{K}_{\mathcal{Q}} := \mathbf{E}\big(\mathbf{Z}_i^2 \mid \mathcal{Q}\big) = \mathbf{E}\big(\mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal\mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal \mid \mathcal{Q}\big) - \mathbf{M}_{\mathcal{Q}}^2.$$

The representation (83) for the mean matrix $\mathbf{M}_{\mathcal{Q}}$ shows that

$$\mathbf{M}_{\mathcal{Q}}^2 = \left(\frac{c^2 n}{n-1}\right)^2 \cdot \mu_n(2 \mid \mathcal{Q})^2 \cdot \left(\mathbf{I} - \frac{1}{n}\mathbb{1}\mathbb{1}^\intercal\right)^2 = \left(\frac{c^2 n}{n-1}\right)^2 \cdot \mu_n(2 \mid \mathcal{Q})^2 \cdot \left(\mathbf{I} - \frac{1}{n}\mathbb{1}\mathbb{1}^\intercal\right)$$

in view of the identity $\mathbf{D}\big(-\frac{1}{n}\big)^2 = \mathbf{D}\big(-\frac{1}{n}\big)$ [a consequence of the elementary observation $\mathbb{1}^\intercal\mathbb{1} = n$]. As the spacings $X_{1i}, \dots, X_{ni}$ are exchangeable conditioned on $\mathcal{Q}$, the conditional covariance matrix $\mathbf{K}_{\mathcal{Q}}$ is a scaled almost diagonal matrix. Its common diagonal terms are given by

$$\begin{aligned}
K_{jj} &= \mathbf{E}\big(\mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal\mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal\big)_{jj} - \big(\mathbf{M}_{\mathcal{Q}}^2\big)_{jj} \\
&= \sum_{l=1}^{n} \mathbf{E}\big(\Lambda_{ji}^2\Lambda_{li}^2\big) - \left(\frac{c^2 n}{n-1}\right)^2 \cdot \mu_n(2 \mid \mathcal{Q})^2 \cdot \left(1 - \frac{1}{n}\right) \\
&= \mathbf{E}\big(\Lambda_{ji}^4 \mid \mathcal{Q}\big) + \sum_{l \neq j} \mathbf{E}\big(\Lambda_{ji}^2\Lambda_{li}^2\big) - \frac{c^4 n}{n-1} \cdot \mu_n(2 \mid \mathcal{Q})^2 \\
&= c^4\left[\mu_n(4 \mid \mathcal{Q}) + (n-1)\mu_n(2,2 \mid \mathcal{Q}) - \frac{n}{n-1} \cdot \mu_n(2 \mid \mathcal{Q})^2\right] =: \gamma_n.
\end{aligned} \tag{87}$$

Likewise, for $j \neq k$, the common off-diagonal terms of $\mathbf{K}_{\mathcal{Q}}$ are given by

$$\begin{aligned}
K_{jk} &= \mathbf{E}\big(\mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal\mathbf{\Lambda}_{*i}\mathbf{\Lambda}_{*i}^\intercal\big)_{jk} - \big(\mathbf{M}_{\mathcal{Q}}^2\big)_{jk} \\
&= \sum_{l=1}^{n} \mathbf{E}\big(\Lambda_{ji}\Lambda_{ki}\Lambda_{li}^2\big) - \left(\frac{c^2 n}{n-1}\right)^2 \cdot \mu_n(2 \mid \mathcal{Q})^2 \cdot \left(\frac{-1}{n}\right) \\
&= \mathbf{E}\big(\Lambda_{ji}^3\Lambda_{ki}\big) + \mathbf{E}\big(\Lambda_{ji}\Lambda_{ki}^3\big) + \sum_{l \notin \{j,k\}} \mathbf{E}\big(\Lambda_{ji}\Lambda_{ki}\Lambda_{li}^2\big) + \frac{1}{n-1} \cdot \frac{c^4 n}{n-1} \cdot \mu_n(2 \mid \mathcal{Q})^2
\end{aligned}$$

$$
\begin{aligned}
&= c^4\left[2\mu_n(3,1 \mid \mathcal{Q}) + (n-2)\mu_n(2,1,1 \mid \mathcal{Q}) + \frac{1}{n-1}\cdot\frac{n}{n-1}\cdot\mu_n(2 \mid \mathcal{Q})^2\right] \\
&= \frac{-c^4}{n-1}\left[\mu_n(4 \mid \mathcal{Q}) + (n-1)\mu_n(2,2 \mid \mathcal{Q}) - \frac{n}{n-1}\cdot\mu_n(2 \mid \mathcal{Q})^2\right] \\
&= \frac{-\gamma_n}{n-1}.
\end{aligned}
$$

In the penultimate step we've appealed to the correlation identities in Corollaries 2 and 3 in Appendix C.1 and consolidated terms; in the final step we've identified the expression within the square brackets with the diagonal term γ_n given in (87).

We conclude that $\mathbf{K}_{\mathcal{Q}}$ is almost diagonal up to a constant scale and is given by

$$
\mathbf{K}_{\mathcal{Q}} = \gamma_n\cdot\mathbf{I} - \frac{\gamma_n}{n-1}\cdot\left(\mathbb{1}\mathbb{1}^\intercal - \mathbf{I}\right) = \frac{n}{n-1}\cdot\gamma_n\cdot\left(\mathbf{I} - \frac{1}{n}\mathbb{1}\mathbb{1}^\intercal\right) = \frac{n}{n-1}\cdot\gamma_n\cdot\mathbf{D}\left(\frac{-1}{n}\right).
$$

We again identify $\alpha = -1/n$ in Lemma 2d and exploit the identity $\left\|\mathbf{D}\left(-n^{-1}\right)\right\|_2 = 1$ to obtain

$$
\left\|\mathbf{K}_{\mathcal{Q}}\right\|_2 = \frac{n}{n-1}\cdot\gamma_n = \frac{c^4 n}{n-1}\left[\mu_n(4 \mid \mathcal{Q}) + (n-1)\mu_n(2,2 \mid \mathcal{Q}) - \frac{n}{n-1}\cdot\mu_n(2 \mid \mathcal{Q})^2\right].
$$

We may rewrite the expression in square brackets on the right in the form

$$
\left[\mu_n(4 \mid \mathcal{Q}) - \mu_n(2 \mid \mathcal{Q})^2\right] + (n-1)\left[\mu_n(2,2 \mid \mathcal{Q}) - \frac{\mu_n(2 \mid \mathcal{Q})^2}{(n-1)^2}\right],
$$

and verify by two applications of Lemma 5 that it is non-negative. The simplest bounds will suffice here. As $\mu_n(2,2 \mid \mathcal{Q}) \le \mu_n(4 \mid \mathcal{Q})$ by another appeal to Lemma 5, we have

$$
\left\|\mathbf{K}_{\mathcal{Q}}\right\|_2 \le \frac{c^4 n}{n-1}\left[\mu_n(4 \mid \mathcal{Q}) + (n-1)\mu_n(2,2 \mid \mathcal{Q})\right] \le \frac{c^4 n}{n-1}\cdot n\cdot\mu_n(4 \mid \mathcal{Q}). \tag{88}
$$

The asymptotic estimate of the fourth conditional moment on the right follows exactly the same pattern as the estimate of the conditional second moment (85) and we merely note down the result:

$$
\mu_n(4 \mid \mathcal{Q}) \le \frac{\mu_n(4)}{1 - \mathbf{P}\left(\mathcal{Q}^\complement\right)} \le \frac{\dfrac{A^4}{n^4}}{1 - \dfrac{K_1 \log(n)^8}{n^2}} = \frac{A^4}{n^4}\left[1 + \mathcal{O}\left(\frac{\log(n)^8}{n^2}\right)\right].
$$

By substitution into (88), we see that

$$
\left\|\mathbf{K}_{\mathcal{Q}}\right\| \le \frac{c^4 A^4}{n^3}\left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right].
$$

The conditional variance statistic of the sum $\sum_i \mathbf{Z}_i$ may hence be bounded by

$$
\mathcal{V}\left(\sum\nolimits_i \mathbf{Z}_i \,\middle|\, \mathcal{Q}\right) := \left\|\sum_{i=1}^n \mathbf{E}\left(\mathbf{Z}_i^2 \mid \mathcal{Q}\right)\right\|_2 = \left\|n\mathbf{K}_{\mathcal{Q}}\right\|_2 = n\cdot\left\|\mathbf{K}_{\mathcal{Q}}\right\|_2 \le \frac{c^4 A^4}{n^2}\left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right]. \tag{89}
$$

4° *Estimating the bounded range over $\mathcal{Q}$.* Applying the triangle inequality to the right hand side of (80) we obtain

$$\begin{aligned}\left\|\mathbf{Z}_i\right\|_2 = \left\|\boldsymbol{\Lambda}_{*i}\boldsymbol{\Lambda}_{*i}^{\intercal} - \mathbf{M}_{\mathcal{Q}}\right\|_2 &\leq \left\|\boldsymbol{\Lambda}_{*i}\boldsymbol{\Lambda}_{*i}^{\intercal}\right\|_2 + \left\|\mathbf{M}_{\mathcal{Q}}\right\|_2 \\ &= \|\boldsymbol{\Lambda}_{*i}\|_2^2 + \left\|\mathbf{M}_{\mathcal{Q}}\right\|_2 \\ &= \sum_{j=1}^{n} \Lambda_{ji}^2 + \left\|\mathbf{M}_{\mathcal{Q}}\right\|_2 \\ &= c^2 \sum_{j=1}^{n} \left(X_{ji} - \frac{1}{n}\right)^2 + \left\|\mathbf{M}_{\mathcal{Q}}\right\|_2.\end{aligned}$$

Over $\mathcal{Q}$ the first term on the right is uniformly bounded by (62); the second term is sub-dominant via the estimates (84) and (85). Reusing notation, over the high probability set $\mathcal{Q}$, we may hence bound

$$\begin{aligned}\max_i \left\|\mathbf{Z}_i\right\|_2 &\leq c^2 \cdot n \cdot \frac{1}{\mathfrak{a}^2 n \log(n)^2} + \frac{c^2 A^2}{n^2}\left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right] \\ &= \frac{c^2}{\mathfrak{a}^2 \log(n)^2}\left[1 + \mathcal{O}\left(\frac{\log(n)^2}{n^2}\right)\right] =: \mathcal{L}_{\mathcal{Q}}.\end{aligned} \tag{90}$$

5° *Bernstein's inequality.* We apply (60) with $d = n$ to the probability on the right in (82) to obtain

$$\mathbf{P}\left\{\left\|\sum_{i=1}^{n} \mathbf{Z}_i\right\|_2 \geq \tau_n \;\middle|\; \mathcal{Q}\right\} \leq 2n \exp\left(-\frac{1}{4}\min\left\{\frac{\tau_n^2}{\mathcal{V}\left(\sum_{i=1}^{n} \mathbf{Z}_i \mid \mathcal{Q}\right)}, \frac{\tau_n}{\mathcal{L}_{\mathcal{Q}}}\right\}\right). \tag{91}$$

The sub-Gaussian term: applying the estimates (86) and (89),

$$\frac{\tau_n^2}{\mathcal{V}\left(\sum_{i=1}^{n} \mathbf{Z}_i \mid \mathcal{Q}\right)} \geq \frac{\left(\dfrac{\varepsilon^2}{\mathfrak{b}^2 \log(n)}\left[1 - \mathcal{O}\left(\dfrac{\log(n)}{n}\right)\right]\right)^2}{\dfrac{c^4 A^4}{n^2}\left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right]} = \left(\frac{\varepsilon}{\mathfrak{b}cA}\right)^4 \cdot \frac{n^2}{\log(n)^2}\left[1 - \mathcal{O}\left(\frac{\log(n)}{n}\right)\right].$$

The sub-Gaussian exponent hence has a growth rate of at least $n^2/\log(n)^2$. We will see shortly that the sub-Gaussian contribution to the tail probability is sub-dominant and we can afford to be cavalier with the bound: we may select n_4 determined by $\mathfrak{a}$, $\mathfrak{b}$, and ε, such that

$$\frac{\tau_n^2}{\mathcal{V}\left(\sum_{i=1}^{n} \mathbf{Z}_i \mid \mathcal{Q}\right)} \geq \frac{1}{2} \cdot \left(\frac{\varepsilon}{\mathfrak{b}cA}\right)^4 \cdot \frac{n^2}{\log(n)^2}$$

eventually, for all $n \geq n_4$. The constant "$\frac{1}{2}$" may, of course, have been replaced by any other constant < 1.

The sub-exponential term: applying the estimates (86) and (90),

$$\frac{\tau_n}{\mathcal{L}_{\mathcal{Q}}} \geq \frac{\dfrac{\varepsilon^2}{\mathfrak{b}^2 \log(n)} \left[1 - \mathcal{O}\left(\dfrac{\log(n)}{n}\right)\right]}{\dfrac{c^2}{\mathfrak{a}^2 \log(n)^2} \left[1 + \mathcal{O}\left(\dfrac{\log(n)^2}{n^2}\right)\right]} = \left(\frac{\varepsilon \mathfrak{a}}{\mathfrak{b} c}\right)^2 \cdot \log(n) \cdot \left[1 - \mathcal{O}\left(\frac{\log(n)}{n}\right)\right].$$

The expression in square brackets on the right tends to 1 as $n \to \infty$ and so the right-hand side has a logarithmic growth rate in n.

It is now clear that the order $n^2/\log(n)^2$ growth of the sub-Gaussian exponent rapidly outstrips the logarithmic growth rate of the sub-exponential exponent. With the benefit of hindsight we may feel that this is intuitive: the distribution of the norm has a sub-exponential tail. *The proper choice of the constant* $\mathfrak{a}$ *now becomes evident: we select* $\mathfrak{a}$ *with a view to obtaining an approximately quadratic rate of decay in the probability* (77)*. With* $\mathfrak{b}$ *any constant satisfying* (74)*, select* $\mathfrak{a}$ *to satisfy*

$$\mathfrak{a} > \frac{\sqrt{12}\, \mathfrak{b} c}{\varepsilon}. \tag{92}$$

With these selections for $\mathfrak{a}$ and $\mathfrak{b}$ we conclude that the sub-exponential exponent is bounded below by

$$\frac{\tau_n}{\mathcal{L}_{\mathcal{Q}}} \geq 12 \log(n) \qquad \text{(eventually)},$$

the caveat to be taken to mean for all $n \geq n_4$ for some n_4 determined by $\mathfrak{a}$, $\mathfrak{b}$, and ε.

With our discretionary choices now all established, let n_5 determined by $\mathfrak{a}$, $\mathfrak{b}$, and ε be the smallest value of n so that

$$\frac{1}{2} \cdot \left(\frac{\varepsilon}{\mathfrak{b} c A}\right)^4 \cdot \frac{n^2}{\log(n)^2} \geq 12 \log(n).$$

Then the inequality (91) simplifies to

$$\begin{aligned} \mathbf{P}\left\{ \left\| \sum_{i=1}^n \mathbf{Z}_i \right\|_2 \geq \tau_n \;\middle|\; \mathcal{Q} \right\} &\leq 2n \exp\left(-\frac{1}{4} \min\left\{ \frac{1}{2} \cdot \left(\frac{\varepsilon}{\mathfrak{b} c A}\right)^4 \cdot \frac{n^2}{\log(n)^2}, 12 \log(n) \right\} \right) \\ &= 2n \exp\left\{ -\frac{1}{4} \cdot 12 \log(n) \right\} = \frac{2}{n^2} \end{aligned} \tag{93}$$

eventually, for all $n \geq \max\{n_3, n_4, n_5\}$. Tracing the sequence back via (93, 82, 78) to (77), we conclude that

$$\mathbf{P}\left\{ \|\mathbf{\Lambda}\|_2 \geq \frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}} \right\} \leq \frac{2}{n^2} + \frac{K_1 \log(n)^8}{n^2}$$

eventually, for all sufficiently large n. Here is the capsule summary.

LEMMA 23: *Fix any* $\varepsilon > 0$ *and suppose* $\mathfrak{b}$ *and* $\mathfrak{a}$ *are constants satisfying* (74) *and* (92)*, respectively. Then there exists a constant* K_3 *determined by* $\mathfrak{a}$*,* $\mathfrak{b}$*, and* ε *such that*

$$\mathbf{P}\left\{ \|\mathbf{\Lambda}\|_2 \geq \frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}} \right\} \leq \frac{K_3 \log(n)^8}{n^2}$$

eventually, for all sufficiently large values of n.

Again, the asymptotics kick in beyond a horizon specified by the selections of ε, $\mathfrak{a}$, and $\mathfrak{b}$. As before, we may take $K_3 = (A\mathfrak{a})^8 + 2$.

All is in readiness for Theorem 5. Reintroduce superscripts $^{(n)}$ to keep track of dimensionality.

THEOREM 5—Equilibria in Large Random Networks: Full Version: *For any sequence of index vectors,* $\{\mathbf{k}^{(n)} \in \{0,1\}^n, n \geq 1\}$, *we have*

$$\sup_{1\leq i\leq n}\left|V_i^n\big(\mathbf{k}^{(n)}\big) - V_i^{0,n}\big(\mathbf{k}^{(n)}\big)\right| \to 0$$

almost surely as $n \to \infty$.

PROOF: Fix any $\varepsilon > 0$. Write $\boldsymbol{\Delta} = \boldsymbol{\Delta}^{(n)}\big(\mathbf{k}^{(n)}\big) = \mathbf{V}^{(n)}\big(\mathbf{k}^{(n)}\big) - \mathbf{V}^{(0,n)}\big(\mathbf{k}^{(n)}\big)$ for the gap sequence. By Lemma 20,

$$\mathbf{P}\Big\{\big\|\boldsymbol{\Delta}^{(n)}\big(\mathbf{k}^{(n)}\big)\big\|_\infty \geq \varepsilon\Big\} \leq \mathbf{P}\Bigg\{\big\|\boldsymbol{\Lambda}^{(n)}\mathbf{V}^{(0,n)}\big(\mathbf{k}^{(n)}\big)\big\|_\infty + \sum_{l=1}^{\infty}\big\|\boldsymbol{\Lambda}^{(n)}\big\|_2^l \cdot \big\|\boldsymbol{\Lambda}^{(n)}\mathbf{V}^{(0,n)}\big(\mathbf{k}^{(n)}\big)\big\|_2 \geq \varepsilon\Bigg\}.$$

But Lemmas 22 and 23 together assert that, with probability at least $1 - 2K_3\log(n)^8/n^2$,

$$\begin{aligned}
\big\|\boldsymbol{\Lambda}^{(n)}\mathbf{V}^{(0,n)}\big(\mathbf{k}^{(n)}\big)\big\|_\infty &+ \sum_{l=1}^{\infty}\big\|\boldsymbol{\Lambda}^{(n)}\big\|_2^l \cdot \big\|\boldsymbol{\Lambda}^{(n)}\mathbf{V}^{(0,n)}\big(\mathbf{k}^{(n)}\big)\big\|_2 \\
&\leq \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} + \sum_{l=1}^{\infty}\left(\frac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}}\right)^l \cdot \mathfrak{b}\sqrt{\log(n)} \\
&= \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} + \varepsilon + \frac{\varepsilon^2}{\mathfrak{b}\sqrt{\log(n)}} + \frac{\varepsilon^3}{\mathfrak{b}^2\log(n)} + \cdots \\
&= \frac{\mathfrak{b}\sqrt{\log(n)}}{\sqrt{n}} + \frac{\varepsilon}{1 - \dfrac{\varepsilon}{\mathfrak{b}\sqrt{\log(n)}}} \\
&= \varepsilon\left[1 + \mathcal{O}\left(\frac{1}{\sqrt{\log(n)}}\right)\right].
\end{aligned}$$

It follows that there exists a constant K_4 determined by the fixed selections of ε, $\mathfrak{b}$, and $\mathfrak{a}$ such that

$$\mathbf{P}\Big\{\big\|\boldsymbol{\Delta}^{(n)}\big(\mathbf{k}^{(n)}\big)\big\|_\infty \geq \varepsilon\Big\} \leq \frac{K_4\log(n)^8}{n^2} \qquad \text{(for all sufficiently large } n\text{)}.$$

The series $\sum_n n^{-2}\log(n)^8$ converges, and so

$$\mathbf{P}\Big\{\big\|\boldsymbol{\Delta}^{(n)}\big(\mathbf{k}^{(n)}\big)\big\|_\infty \geq \varepsilon \text{ i.o.}\Big\} = 0$$

by the first Borel–Cantelli lemma (c.f. Venkatesh (2013, Section IV.4)).[11] As the tiny positive ε may be taken arbitrarily small, we conclude that $\left\|\mathbf{\Delta}^{(n)}\left(\mathbf{k}^{(n)}\right)\right\|_\infty \to 0$ almost surely. ▶

We include a word on the choices of the discretionary constants $\mathfrak{a}$ and $\mathfrak{b}$ that appeared somewhat mysteriously in the proof and have had a magical effect on the outcome.

Bearing in mind that the variance of the spacings is of order n^{-2}, individual components $\left(\mathbf{\Lambda}\mathbf{V}^0(\mathbf{k})\right)_i$ of the random walk exhibit excursions of order $n^{-1/2}$ from the origin by virtue of the central limit theorem. The need to control the *maximal* excursion of the n components of the walk leads to the considered deviation of $\mathfrak{b}\sqrt{\log(n)}\big/\sqrt{n}$. The constant $\mathfrak{b}$ is now chosen just large enough—in other words, the maximal excursion is pegged to a value just big enough—to guarantee that the sub-Gaussian tail goes to zero quickly enough for the strong law to kick in via the Borel–Cantelli lemma.

With the collective excursions of the components of the random walk confined to this order, the ℓ^2-norm of the random walk grows at the rate of $\mathfrak{b}\sqrt{\log(n)}$ which increases slowly but inexorably. The need to rein in this growth compels us to try to control the spectral norm of the centered cross-holdings matrix $\mathbf{\Lambda}$ to within $\varepsilon\big/\left(\mathfrak{b}\sqrt{\log(n)}\,\right)$ where the tiny ε is chosen as small as desired.

The key to the analysis is the fact that the negative correlative structure of the spacings in the asymptotically diffuse setting prohibits pathologically large deviations from the mean. While tighter concentrations could have been achieved at the expense of more smoothness constraints, for example, by levying yet higher moment constraints or a sub-exponential condition, the choice of an eighth moment condition imposes the least onerous constraints consistent with our goal. But we now have to step delicately. The eighth moment constraint concentrates the spacings collectively within a deviation of $1\big/\sqrt{n}$ from their means but this is just a smidgin too loose. The key to managing the sub-exponential tail of the spectral norm of $\mathbf{\Lambda}$ is to buy ourselves a little wiggle room by considering the set where the n^2 spacings collectively are constrained to excursions from the mean within a *slightly* more corseted box of size $1\big/\left(\mathfrak{a}\sqrt{n}\,\log(n)\right)$. The logarithmic tightening of the concentration box gives us that little extra wiggle room: the logarithmic term opens up a tiny window of opportunity through which the square root logarithmic control desired for the spectral norm of $\mathbf{\Lambda}$ can now opportunistically squeeze in. And finally, when the dust has settled, the constant $\mathfrak{a}$ provides the coup de grâce by fine tuning the logarithmic rate so that the sub-exponential tail decays quickly enough for the Borel–Cantelli lemma to come into force.

APPENDIX E: THE NO-ESCAPE ASSUMPTION

We remark upon the repeatedly invoked assumption that the system's trajectory $\mathbf{V}_t$ does not escape the orthant(s) under consideration. Such an escape is possible even in regular cliques, as shown in Fig. E.1. Orthant boundaries are linear, while the solution to the system—expressed through exponential functions per Theorem 7—is not, and, since orthant boundaries do not explicitly repel or attract trajectories, and the only force driving the dynamics is that of attraction by a (putative or feasible) equilibrium, a trajectory passing close to an orthant's boundary may escape into a nearby orthant.

Fig. E.2 overlays the trajectories of the two-dimensional model instance from Fig. E.1 on top of the basins of attraction of the four equilibria.

An empirical observation that may be made from these figures is that, within a bounding box whose boundary is not too far from the equilibria, the boundaries of the attraction basins stay

[11] In the terminology introduced by Kai Lai Chung, the event $\limsup_n A_n = \bigcap_m \bigcup_{n\geq m} A_n$ is more vividly captured in the language "A_n *occurs infinitely often*" and denoted A_n i. o..

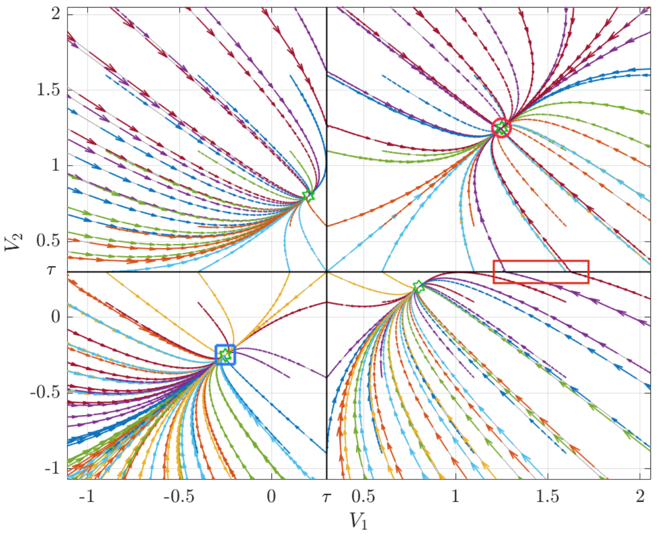

FIGURE E.1.—State trajectories $\mathbf{V}_t$ satisfying (3) for a regular clique of two firms. Equilibria are shown as stars, with $\mathbf{V}^{\text{inf}}$ and $\mathbf{V}^{\text{sup}}$ wrapped in a square and a circle, respectively. Two trajectories beginning in the orthant $\mathbb{K}^2([1,0]^\intercal)$ and passing close to the orthant $\mathbb{K}^2([1,1]^\intercal)$ escape into it. The location of escape is shown in a red frame.

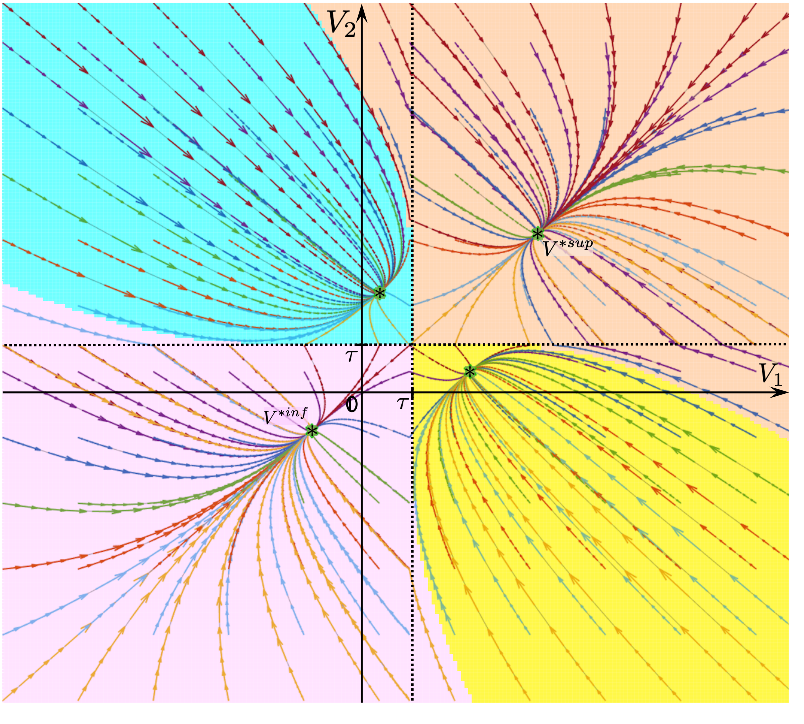

FIGURE E.2.—Four equilibria, their attraction basins, and several representative trajectories $\mathbf{V}_t$ for the model with two firms: $\beta = 0.6$, $c = 0.6$, $\tau = 0.4$, and $e = 0.5$.

close to the boundaries of the orthants. This suggests that, at least in low dimensions, trajectory escape to nearby orthants is a rare event.

But what about higher dimensions and (locally) highly irregular network instances? Are such escapes rare in such networks? And, if present, are they localized, occurring only in some orthants? In Fig. E.3, we display the aggregated results of a large number of model simulations for a network of $n = 32$ firms for the cases of a *regular clique*, a *single instance of an irregular clique*, and a *random clique* (with each simulation run using a new network instance). For each network, we track the orthant locations of the initial firm valuations $\mathbf{V}_0$ and the final firm valuations $\mathbf{V}_\infty = \lim_t \mathbf{V}_t$. In regular cliques, trajectory dynamics largely follow the idealized dynamics diagram of Fig. 7 (though, as mentioned earlier, trajectory escapes *are* possible). In the case of a single instance of an irregular clique, there is an increased frequency of sporadic trajectory escape from orthants $\mathbb{K}(k)$ where the number k of solvent firms is close to k_{low} or k_{high} into the extremal basins of attraction. The picture from a single instance of an irregular clique does not change significantly when we aggregate results over a large number of random clique instantiations in the third figure though we now also see the sporadic emergence of equilibria in orthants adjacent to the two extremal orthants in rare random clique instances. To summarize: the simulations suggest that trajectory escapes are largely confined to orthants $\mathbb{K}(k)$ where the number k of solvent firms is close to k_{low} or k_{high}, that is to say, in the orthants at the boundaries of the "safety band" in the state space.

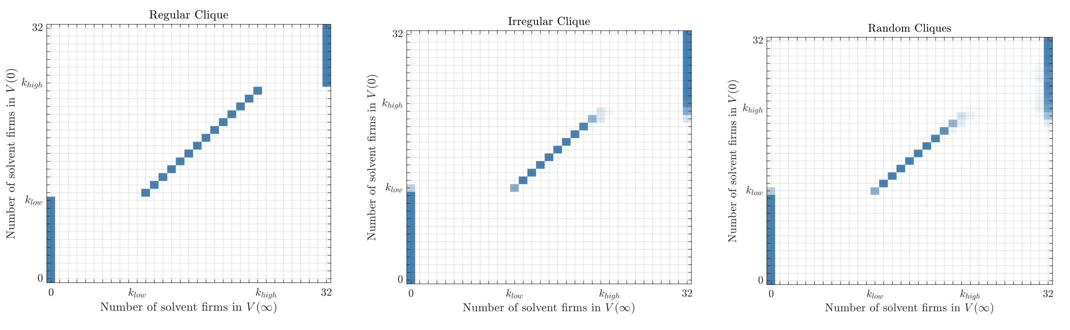

FIGURE E.3.—Tracking trajectory escape for a network of $n = 32$ firms for the cases of a regular clique, a single instantiation of an irregular clique, and a random clique. Each cell represents the number of solvent firms in a given orthant $\mathbb{K}(k)$. A fixed number of runs is made for each value of k by, for each of the three cases, selecting a random initial point $\mathbf{V}_0$ in that orthant and tracking the orthant location of the corresponding limit point $\mathbf{V}_\infty$. The cell color intensity reflects the frequency of the associated transition aggregated over the individual runs. The diagonal cells correspond to orthant preservation by a trajectory (no escape). In all three cases, trajectories originating in orthants with $k \le k_{\text{low}}$ are typically absorbed by $\mathbf{V}^{\text{inf}}$, while trajectories originating in orthants with $k \ge k_{\text{high}}$ are typically absorbed by $\mathbf{V}^{\text{sup}}$. Trajectories originating in orthants with $k_{\text{low}} < k < k_{\text{high}}$ are generally confined to the orthant, though there is a greater frequency of escape from boundary orthants in the cases of the irregular and random cliques.

It should be noted that the similarity of the dynamics when considering a large number of random cliques drawn from a distribution (the diagram on the right in Fig. E.3) to that of the regular clique (the diagram on the left in Fig. E.3) is *not* due to any sort of "averaging" over random network instances; even a *single* arbitrarily chosen, sufficiently large, *irregular* network instance (the middle diagram in Fig. E.3) with high probability manifests the "typical" dynamics of the regular clique, and the slightly "diffused" nature of two latter diagrams in Fig. E.3 stems from local irregularities in the network rather than randomness per se.

Our findings for trajectories for random cliques are summarized in the revised firm valuation dynamics diagram shown schematically in Fig. E.4.

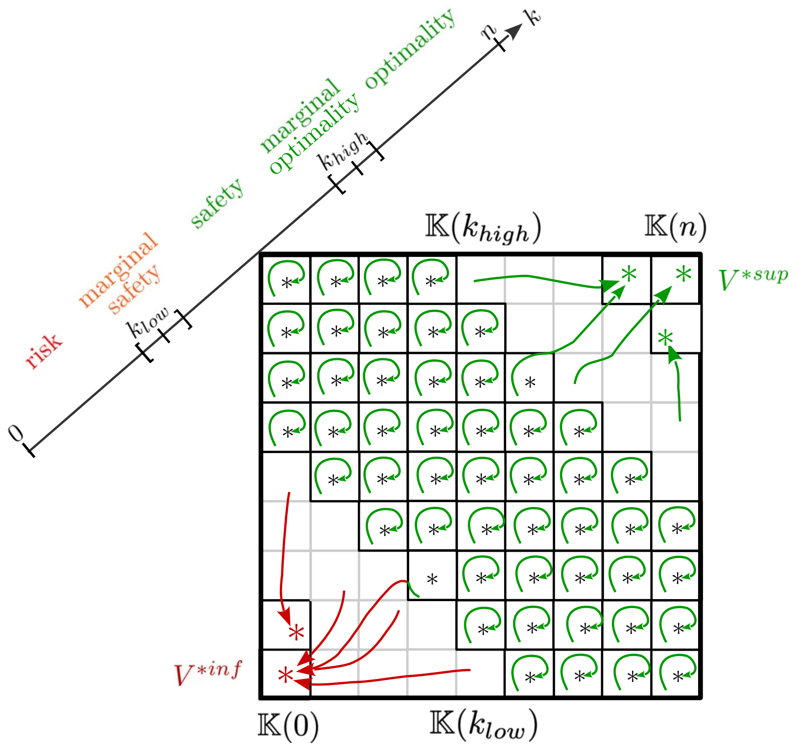

FIGURE E.4.—Diagram of firm valuation dynamics in random cliques adjusted to account for possible trajectory escapes from orthants at the boundaries of the "safety band" in the state space to the adjacent extremal orthant.

APPENDIX F: PROOFS OF THEOREMS 6 AND 7

THEOREM 6—Dynamics in Orthants Containing an Equilibrium: *If the trajectory $\mathbf{V}_t$ starts in orthant $\mathbb{K}^n(\mathbf{k})$, then, for all $t > 0$ until the trajectory escapes $\mathbb{K}^n(\mathbf{k})$ or, if such an escape never happens, for all $t > 0$,*

$$\mathbf{V}_t = \exp\big(-(\mathbf{I} - \mathbf{C})t\big)\big(\mathbf{V}_0 - \mathbf{V}(\mathbf{k})\big) + \mathbf{V}(\mathbf{k}), \tag{22}$$

and, a fortiori, if the trajectory does not escape $\mathbb{K}^n(\mathbf{k})$, $\lim_{t\to\infty} \mathbf{V}_t = \mathbf{V}(\mathbf{k})$.

PROOF: The assumption that the state $\mathbf{V}_t$ never leaves the orthant $\mathbb{K}^n(\mathbf{k})$ for all considered times t allows us to treat the system (3) as an affine linear system. We begin by a change of coordinates, moving the origin to $\mathbf{V}(\mathbf{k})$: set $\mathbf{x}_t = \mathbf{V}_t - \mathbf{V}(\mathbf{k})$. In these new coordinates, the original system (3) takes the form

$$\dot{\mathbf{x}}_t = -(\mathbf{I} - \mathbf{C})\mathbf{x}_t \quad \big[\text{with initial condition } \mathbf{x}_0 = \mathbf{V}_0 - \mathbf{V}(\mathbf{k})\big].$$

It is well known from the theory of linear systems, that this system has the unique solution

$$\mathbf{x}_t = \exp\big(-(\mathbf{I} - \mathbf{C})t\big)\mathbf{x}_0.$$

Returning from $\mathbf{x}$ to $\mathbf{V}$, we obtain (22).

To show convergence to the limit point $\mathbf{V}(\mathbf{k})$, begin with the trite observation $(\mathbf{C}^\intercal \mathbb{1})_i = c$. As $0 < c < 1$, by the Gershgorin disc theorem, the spectral abscissa of the matrix $-(\mathbf{I} - \mathbf{C})$ is negative (see the proof of Lemma 1),

$$\mu\big(-(\mathbf{I}-\mathbf{C})\big) := \max\big\{\mathrm{Re}\,(\lambda) : \lambda \in \mathrm{spec}\big(-(\mathbf{I}-\mathbf{C})\big)\big\} \leq c - 1 < 0.$$

It follows that the matrix $-(\mathbf{I} - \mathbf{C})$ is Hurwitz (or convergent), whence, for its matrix exponential, $\lim_{t\to\infty} \exp\big(-(\mathbf{I}-\mathbf{C})t\big) = \mathbb{0}_{n\times n}$. We conclude that $\lim_{t\to\infty} \mathbf{V}_t = \mathbf{V}(\mathbf{k})$. ▶

THEOREM 7—Dynamics in Orthants Containing an Equilibrium for Regular Cliques: *If the trajectory $\mathbf{V}_t = (V_{1,t}, \dots, V_{n,t})^\intercal$ starts in orthant $\mathbb{K}^n(\mathbf{k})$, then, for all $t > 0$ until the trajectory escapes $\mathbb{K}^n(\mathbf{k})$ or, if such an escape never happens, for all $t > 0$,*

$$\begin{aligned} V_{i,t} = {} & \big[\mathrm{avg}(\mathbf{V}_0) + \beta\big(1 - \tfrac{|\mathbf{k}|}{n}\big) - V_i^0(\mathbf{k})\big] \exp\big(-(1-c)t\big) \\ & - \big[\mathrm{avg}(\mathbf{V}_0) + \beta\big(1 - \tfrac{|\mathbf{k}|}{n}\big) - V_{i,0}\big] \exp(-t) + V_i^0(\mathbf{k}), \end{aligned} \tag{24}$$

and, a fortiori, the trajectory approaches the equilibrium exponentially fast.

PROOF: The trajectory $\mathbf{V}_t$ is given by (22) with $\mathbf{C} = \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal$ and $\mathbf{V}(\mathbf{k}) = \mathbf{V}^0(\mathbf{k})$. Recall two elementary facts about matrix exponentials: (1) if $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ is diagonal then its matrix exponential is given by $\exp(\mathbf{\Lambda}) = \mathrm{diag}\big(e^{\lambda_1}, \dots, e^{\lambda_n}\big)$; and (2) if $\mathbf{A}$ is a real, symmetric matrix with eigendecomposition $\mathbf{A} = \mathbf{U\Lambda U}^{-1}$, its matrix exponential is given by $\exp(\mathbf{A}) = \mathbf{U}\exp(\mathbf{\Lambda})\mathbf{U}^{-1}$. For the case of the regular clique we may identify $\mathbf{A} = -(\mathbf{I} - \mathbf{C}) = -\big(\mathbf{I} - \frac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)$ as almost diagonal. Routine computation now shows (see Lemma 2)

$$\mathbf{\Lambda} = \left[\begin{array}{c|c} \mathbb{0}_{(n-1)\times(n-1)} & \mathbb{0} \\ \hline \mathbb{0} & c \end{array}\right],$$

$$\mathbf{U} = \left[\begin{array}{c|c} -\mathbb{1}_{n-1}^\intercal & 1 \\ \hline \mathbf{I}_{(n-1)\times(n-1)} & \mathbb{1}_{n-1} \end{array}\right],$$

$$\mathbf{U}^{-1} = \frac{1}{n}\left(\left[\begin{array}{c} -\mathbb{1}_{(n-1)\times n} \\ \hline \mathbb{1}_{1\times n} \end{array}\right] + (n-1)\mathbf{S}_n\right),$$

where $\mathbf{S}_n$ is the n-by-n upper shift matrix with $n-1$ ones above the main diagonal, with all other elements equal to zero. After simplification, we now get

$$\exp\big\{-\big(\mathbf{I} - \tfrac{c}{n}\mathbb{1}\mathbb{1}^\intercal\big)t\big\} = \tfrac{1}{n}\big\{\exp\big(-(1-c)t\big) - \exp(-t)\big\}\mathbb{1}\mathbb{1}^\intercal + \exp(-t)\mathbf{I}.$$

Substituting into the expression (22) for $\mathbf{V}_t$, after simplification, we recover the claimed result (24). ▶