Proceedings of the 2019 Winter Simulation Conference
N. Mustafee, K.-H.G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, eds.

LIMITING DISTRIBUTIONAL FIXED POINTS IN SYSTEMIC RISK GRAPH MODELS

Anand Deo
Sandeep Juneja

School of Technology and Computer Science
Tata Institute of Fundamental Research
Mumbai, 400005, INDIA

ABSTRACT

We analyse the equilibrium behaviour of a large network of banks in presence of incomplete information, where inter-bank borrowing and lending is allowed, and banks suffer shocks to assets. In a two time period graphical model, we show that the equilibrium wealth distribution is the unique fixed point of a complex, high dimensional distribution-valued map. Fortunately, there is a dimension collapse in the limit as the network size increases, where the equilibriated system converges to the unique fixed point involving a simple, one dimensional distribution-valued operator, which, we show, is amenable to simulation. Specifically, we develop a Monte-Carlo algorithm that computes the fixed point of a general distribution-valued map and derive sample complexity guarantees for it. We numerically show that this limiting one-dimensional regime can be used to obtain useful structural insights and approximations for networks with as low as a few hundred banks.

1 INTRODUCTION

We analyse a large network of financial institutions (henceforth referred to as banks) where inter-bank borrowing and lending of capital is allowed, and arrive at structural simplifications as the network size increases to infinity. Specifically, we consider a graphical model of the banking network where vertices denote banks and edges spell out the inter-bank liability structure. A two time period framework is considered: In period one, banks borrow and lend capital to each other, and inter-bank linkages in the network are realised. In the next period, they receive income from their external assets modulo random shocks, and use it to clear external and inter-bank liabilities. The equilibrium wealth of the system is the resultant fixed point solution that balances the incoming and the outgoing wealth at each bank. Such graphical models have been extensively studied in literature, see, e.g., Allen and Gale (2000), Eisenberg and Noe (2001), Haldane and May (2011) and Glasserman and Young (2016). It can be shown that the wealth of banks in equilibrium is the unique fixed point of a vector valued map. See, e.g., Eisenberg and Noe (2001) or Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015). However, one shortcoming in much of the existing literature is the assumption that the entire network wealth and liability structure is known to the modeller. This is often unrealistic in practice (see Anand, Craig, and Von Peter (2015)): a modeller typically only gets access to balance sheets of banks in the network and only knows the aggregate assets and liabilities, i.e., how much each bank lends or borrows, but not from whom. A modeller may also know the average number of creditors and debtors of a typical bank, without access to further information granularity.

To address this lack of complete information, we model the banking network as a random, weighted graph, where the randomness captures the modeller's incomplete information on exact assets and liabilities, as well as unobserved (both idiosyncratic and systemic) shocks to the bank's assets. We allow the inter-bank connectivities, liabilities and external assets to be random variables, whose distribution matches the observed statistical properties of the network. Other works also consider random graphical models of banking networks (see, e.g., Amini, Cont, and Minca (2016)). However in their models, the recoveries of a bank from a defaulting debtor are independent of the actual wealth possessed by the debtor. Our model departs from this by allowing recoveries on default to be state dependent, resulting in a more realistic view of the banking network.

We consider a sparse graph regime where each bank has only a few counter-parties and show that the wealth possessed by banks in equilibrium is a random vector whose distribution is a fixed point of a distribution-valued map (specified in Section 3). A similar approach is also followed by Kavitha, Saha, and Juneja (2018), however, they model the network as a dense graph, where the expected number of counter-parties of a bank go to infinity. The sparse graph model which we consider differs from theirs both in terms of the analysis and conclusions. Further, it has been observed empirically, (see Cont and Moussa (2010)) that banking networks are sparse, and hence we expect our regime to more accurately capture real banking networks.

To infer the statistical properties of the network, it is essential to sample from this fixed point. However, typical banking networks are quite large, often consisting of hundreds or thousands of banks. Sampling thus is computationally prohibitive due to the underlying high dimensionality. We show that as the size of the banking network grows, due to conditional independence amongst underlying random variables, there is a dimension collapse: the distributional fixed point of a large banking system converges to the product of a one-dimensional distributions. Hence, the statistical properties of a large network are well approximated by the limiting fixed point distribution. This provides a great deal of structural insights and may be of use in conducting many *what if* analysis on a large network. As we discuss below, it is easy to generate approximate samples from the limiting fixed point distribution, providing a large amount of computational reduction for the network. Proving the existence and uniqueness of fixed points as well as distributional convergence for the growing random banking network is typically not straightforward. One of our contributions is to leverage the theory of optimal transport, which enables us to metrize the infinite dimensional space of distributions in which the fixed points and their limits lie. This provides a relatively simple approach to prove limit theorems.

The second part of this paper focuses on simulation of the limiting fixed point and related computational issues. Let $T(\cdot)$ be a map on the space of probability measures on the line. Suppose $T(\cdot)$ is contractive (and hence has a unique fixed point), and satisfies certain moment bounds. Then, assuming that for any distribution P one can sample from $T(P)$, we develop a Monte Carlo algorithm which gives a distribution which is close to the fixed point of $T(\cdot)$, and derive probabilistic guarantees for it. To the best of our knowledge, the algorithm and its analysis are new. Applying this to the banking network, we show that to approximate a realistic system a small number of computations are required. Testing this algorithm on a simulated example where there are a few hundred banks in the network, we find that the limiting fixed point distribution approximates the global properties of the large network well.

To summarise, our contributions are two-fold: We provide a framework to capture the modelling uncertainty in a graphical model of a banking system, and provide asymptotic analysis and approximations. Secondly, we provide a method to simulate these approximations in an efficient way, and provide some structural insights into the behaviour of a large banking system.

In Section 2, we provide the technical background and notation used throughout this paper. In Section 3 we introduce our model and derive the asymptotic behaviour of the large banking system. In Sections 4 and 5 we discuss the simulation of distributional fixed points, and related computational issues. Section 6 concludes with simulation experiments, which validate our theoretical results. Due to space limitations, only proof sketches are provided. A more elaborate version of this paper will consider richer models and refined mathematical analysis.

2 BACKGROUND

In our asymptotic analysis, we encounter measures on the vertices of an increasing sequence of graphs. We embed these in a common space of measures on $\Re^\infty$ (by setting all marginals of components larger than the system size to 0). To study convergence of properties of the large graph to its limit, as well as to easily establish existence and uniqueness of fixed points, it is essential to metrize this space. An elegant and natural way of doing this is through Wasserstein distances, which we describe next. Let $\mathscr{X}$ be a Polish space equipped with a metric $d(\cdot,\cdot)$, and let $\mathscr{M}(\mathscr{X})$ denote the set of probability measures on $\mathscr{X}$. For $\mu,\nu \in \mathscr{M}(\mathscr{X})$, we define the Wasserstein distance of order p under d as

$$W_p(\mu,\nu) = \inf_{X\sim\mu, Y\sim\nu} (E(d(X,Y)^p)^{\frac{1}{p}}. \tag{1}$$

We call the joint distribution of μ and ν which achieves this infimum as the optimal coupling between μ and ν. It can be shown that for all $p<\infty$, W_p makes the space of probability measures a Polish space. Further, weak convergence together with convergence of pth moments is equivalent to convergence in W_p (see Villani (2003), Theorem 7.12). The following duality for Wasserstein distance of order 1 is also useful:

$$W_1(\mu,\nu) = \sup_{f\in \mathrm{Lip}(1)} \left| \int f d\mu - f d\nu \right|, \tag{2}$$

where $\mathrm{Lip}(1)$ is the set of all 1-Lipschitz function with respect to the metric d. For any distribution $P \in \mathscr{M}(\mathscr{X})$, and any $(i_1,\dots,i_k)$, $k<\infty$, we denote its finite dimensional projections on the components $(i_1,\dots,i_k)$ by $\pi^{-1}_{i_1,\dots,i_k}P$. Let, $\mathbf{X} = \{X_i\}$ denote a random vector with marginals X_i. We use $\xrightarrow{\mathscr{D}}$ to denote convergence in distribution. Let $\{Y_i\}$ be an indexed set of random variables, and $\{a_i\}$ be a sequence of real numbers. Then, $Y_n = O_{\mathscr{P}}(a_n)$ if for any $\varepsilon>0$, there exists an $M>0$, such that for all n large enough, $P\left(\left|\frac{Y_n}{a_n}\right| > M\right) < \varepsilon$. For any random element Z (scalar or vector), let $\mathscr{L}(Z)$ denote its law. For two random elements Z_1 and Z_2, we say that $Z_1 \overset{\mathscr{D}}{=} Z_2$ if Z_1 and Z_2 have the same distribution. We denote a Poisson variable with rate α by $\mathrm{Poi}(\alpha)$, and use $\mathrm{Bin}(N,p)$ to denote a binomial random variable with size N and success probability p. Lastly, for any $\mu \in \mathscr{M}(\Re)$, suppose $\{X_i : i\geq 1\}$ are drawn from μ independently. Letting $\delta_{(\cdot)}$ be the Dirac point measure, denote the the empirical distribution of μ, of size n is given by $\frac{1}{n}\sum_{i=1}^n \delta_{X_i}$.

3 LARGE NETWORK UNDER MODELLING UNCERTAINTY

Consider an economy where there are N banks of identical seniority, indexed by $\{1,2,\dots,N\}$ and represented as the vertices of a directed graph $\mathscr{G}_N$. For a bank j, let A_j, B_j denote the total external assets and inter-bank liabilities, respectively. We additionally assume that j has an external liability of V_j (e.g. customer deposits and operational costs). Note that these are modelled as random variables to capture the uncertainty in the network, as well as unobserved idiosyncratic shocks to assets. We make the following structural assumption on the network:

Assumption 1 The graph $\mathscr{G}_N$ is directed Erdos-Renyi. Letting p_N denote the probability of there being an edge between a pair of vertices, $Np_N \to \alpha > 0$. Further, there is a common shock $\mathbf{Z}\in\Re^d$, conditioned on which (A_i,B_i,V_i) is i.i.d. over i such that A_i, B_i and V_i each have a finite moment of order $\gamma>2$.

Assumption 1 captures the fact that banking networks are sparse (see Cont and Moussa (2010)): inter-bank relations are costly, and typical banks only have a small number of counter-parties. In this regime, the number of creditors and debtors of any bank are distributed as $\mathrm{Bin}(N-1,p_N)$. The shock $\mathbf{Z}$ can be interpreted as a loss due a common risk for each bank: it could consist of portfolios invested by the banks in risky assets, as well as the general state of the economy. For instance, the wealth possessed by each bank may be a random function of a common risky asset $\mathbf{Z}$. A crisis in the stock market, or general economic depression then reduces the external assets. Since all the results we state and prove are be conditioned on a fixed realisation of the common shock $\mathbf{Z}$, we henceforth hide $\mathbf{Z}$ and assume that (A_i, B_i, V_i) are i.i.d. across i. Suppose the economy operates over two time periods, $t = \{0, 1\}$ as follows:

- At t=0, the banks borrow capital from others. An edge from bank j to bank i (which we denote by $(j \to i)$) is present if bank j is liable to i. The total liability inter-bank of i is given by B_i.
- At time t=1, each bank i receives random external asset income A_i which may be small if it suffers an idiosyncratic shock. Bank i first clears its liabilities outside the system, V_i, and then pays back the amounts it has borrowed from others within the system. Bank i defaults if the wealth available to it post clearing, denoted by X_i^N is less than $B_i + V_i$. Since we have assumed equal seniority, payments made within the system are pro-rata, that is, if a bank defaults, it clears its dues to its creditors in proportion to what it owes them.

The total wealth held by bank i is the sum of its external assets and inter-bank claims. In equilibrium, this holds simultaneously for all banks in the network. Further, any bank pays the system the minimum of its wealth net senior liabilities and what it owes, i.e., bank i pays the system $(X_i^N - V_i)^+ \wedge B_i$. Since for each i, only the aggregate liability B_i is available to the modeller, it is reasonable to assume that a bank borrows equally from each of its creditors (see Gai and Kapadia (2010)), although this is easily relaxed (see Remark 2). The joint distribution of wealths of the banks hence satisfies

$$\{X_i^N\}_{i=1,\dots,N} \stackrel{\mathscr{D}}{=} \left\{ A_i + \sum_{j=1}^{N} \frac{(X_j^N - V_j)^+ \wedge B_j}{P_j^N} \mathbf{I}(j \to i) \right\}_{i=1,\dots,N}, \tag{3}$$

where recall that X_j^N is the wealth of bank j post clearing, $P_j^N \stackrel{\mathscr{D}}{=} \mathrm{Bin}(N-1, p_N)$ is the out-degree of bank j (this represents the number of banks j is liable to). Recall that $\mathscr{M}(\mathfrak{R}^\infty)$ is the space of probability measures on $\mathfrak{R}^\infty$. Now define the map $T_N : \mathscr{M}(\mathfrak{R}^\infty) \to \mathscr{M}(\mathfrak{R}^\infty)$,

$$T_N(\lambda) = \mathscr{L}\left(\left\{ A_i + \sum_{j=1}^{N} \frac{(X_j - V_j)^+ \wedge B_j}{P_j^N} \mathbf{I}(j \to i) \right\}_{i=1}^{N} \right), \tag{4}$$

where $\mathbf{X} = \{X_i\}$ is drawn jointly from $\lambda \in \mathscr{M}(\mathfrak{R}^\infty)$. Then, the wealths of the banks in an economy of size N are a fixed point of $T_N(\cdot)$. Observe that for all $\lambda \in \mathscr{M}(\mathfrak{R}^\infty)$ if $k > N$, then $\pi_k^{-1} T_N(\lambda) = 0$. The following lemma shows that the system of distributional equations in (4) is consistent. Let $\mathscr{C}_i^N$ be the set of debtors of bank i. Define the distance on $\mathfrak{R}^\infty$: $d_N(x,y) = \sum_{i=1}^{N}(|x_i - y_i|) \triangleq \|x-y\|_{1,N}$. Then,

Lemma 1 For every N, $T_N(\cdot)$ is a contraction in the Wasserstein-1 metric, and thus has a unique fixed point, call it λ_N^*. Further, $\{\lambda_N^*\}$ is a relatively compact sequence in $\mathscr{M}(\mathfrak{R}^\infty)$.

Proof sketch of Lemma 1 For every fixed N,

$$W_1(T_N(\mu), T_N(\nu)) = \inf_{\mathbf{U} \sim T_N(\mu), \mathbf{V} \sim T_N(\nu)} \mathbf{E}\|\mathbf{U} - \mathbf{V}\|_{1,N} \text{ , taking Wasserstein distances with respect to } d_N(\cdot,\cdot)$$

$$\leq \mathbf{E}\left\| \left\{ \sum_{j \in \mathscr{C}_i^N} \frac{((X_j - V_j)^+ \wedge B_j) - ((Y_j - V_j)^+ \wedge B_j))}{P_j^N} \right\}_i \right\|_{1,N} \text{ where } \mathbf{X} \sim \mu, \mathbf{Y} \sim \nu \tag{5}$$

$$\leq \mathbf{E}\left(\sum_{i=1}^{N} \sum_{j \in \mathscr{C}_i^N} \frac{|((X_j - V_j)^+ \wedge B_j) - ((Y_j - V_j)^+ \wedge B_j)|}{P_j^N} \right) \tag{6}$$

$$\leq \sum_{i=1}^{N} \mathbf{E}\left(\sum_{j \in \mathscr{C}_i^N} \frac{|X_j - Y_j|}{P_j^N} \right) \leq \alpha_N \mathbf{E}\left(\sum_{j=1}^{N} |X_j - Y_j| \right) = \alpha_N \mathbf{E}(\|\mathbf{X} - \mathbf{Y}\|_{1,N}), \tag{7}$$

where the measures μ, ν in $\mathscr{M}(\Re^\infty)$. In (5), we have selected all random variables except $\mathbf{X}$ and $\mathbf{Y}$ to be identical. It is easy to see that this gives the correct marginals, and hence is a valid coupling of $T_N(\mu)$ and $T_N(\nu)$ (joint distribution with the correct marginals). The inequality is then a consequence of (1). To see (6) use the definition of the distance $d(\cdot,\cdot)$, and the triangle inequality. To go from (6) to (7), note that for all x, y, a and b, $|(x\vee a)\wedge b-(y\vee a)\wedge(b)| \leq |x-y|$. Since the bound in (7) is independent of the joint distribution of $\mathbf{X}$ and $\mathbf{Y}$, it also holds under the optimal coupling between of $\mathbf{X}$ and $\mathbf{Y}$ (this exists by Theorem 4.1 from Villani (2003)). Then, one can replace the expectation in (7) by $W_1(\mu,\nu)$. α_N in (7) is $\mathbf{E}\left(\sum_{j\in\mathscr{C}_i^N}\frac{1}{P_j^N}\right)$. For $j\in\mathscr{C}_i^N$, P_j^N are distributed as $1+\mathrm{Bin}(N-2,p_N)$, i.i.d. It can be shown that for all N, $\alpha_N<1$, and hence $T_N(\cdot)$ is a contraction and has a unique fixed point. □

Remark 1 $d_N(\cdot,\cdot)$ is not a metric on $\Re^\infty$ (it is a pseudo-metric). However, the range of $T_N(\cdot)$ is restricted to $\mathscr{M}(\Re^N)$. Since $d_N(\cdot,\cdot)$ defines a metric on $\Re^N$, the proof goes through.

Define $T:\mathscr{M}(\Re)\to\mathscr{M}(\Re)$ as

$$T(\mu)=\mathscr{L}\left(A_1+\sum_{i=1}^{P}\frac{(X_i-V_i)^+\wedge B_i}{1+P_i}\right), \tag{8}$$

where P_i and P are independent $\mathrm{Poi}(\alpha)$ random variables. We now state this section's main result:

Theorem 1 The limiting map $T(\cdot)$, defined by (8) is a contraction in the Wasserstein 1 metric, and hence has a unique fixed point. Let $T_N(\cdot)$ be defined by (4). Then, if λ^* is the unique fixed point of (8), and λ_N^* is the unique fixed point of $T_N(\cdot)$, $\lambda_N^*\xrightarrow{\mathscr{D}}\bigotimes_{i=1}^\infty\{\lambda^*\}_i$, as well as in W_p, for any $p\leq 2$, where $\bigotimes$ denotes the product measure and $\{\lambda^*\}_i\stackrel{\mathscr{D}}{=}\lambda^*$.

Proof Sketch: The proof consists broadly of the following steps:

1) Uniqueness of limiting fixed point: To show that λ^* is the unique fixed point of $T(\cdot)$ defined by (8) we show that it is a contraction. Fix measures λ_1, $\lambda_2\in\mathscr{M}(\Re)$. Observe that as consequence of (1), for any arbitrary coupling Π of λ_1,λ_2, $W_1(\lambda_1,\lambda_2)\leq E_{(X,Y)\sim\Pi}|X-Y|$. Let $W_{i,1}$ and $W_{i,2}$ be drawn i.i.d. from λ_1 and λ_2, respectively. Then, proceeding in a manner similar to the proof of Lemma 1,

$$\begin{aligned}W_1(T(\lambda_1),T(\lambda_2)) &\leq \mathbf{E}\left|\sum_{i=1}^{P}\frac{(A_i-V_i+W_{i,1})^+\wedge B_i-(A_i-V_i+W_{i,2})^+\wedge B_i)}{1+P_i}\right| \leq \mathbf{E}\left(\sum_{i=1}^{P}\frac{|W_{i,1}-W_{i,2}|}{1+P_i}\right)\\ &\leq \mathbf{E}\left(\frac{P}{1+P'}\right)W_1(\lambda_1,\lambda_2) \text{ where } P \text{ and } P' \text{ are independent } \mathrm{Poi}(\alpha).\end{aligned} \tag{9}$$

One can verify that (9) is $(1-\mathrm{e}^{-\alpha})W_1(\lambda_1,\lambda_2)$. Hence, $T(\cdot)$ is a contraction, and has a unique fixed point.

2) Constructing bank ancestries: By relative compactness, it is sufficient to argue the convergence of the projection $\pi_{1,2}^{-1}\lambda_N^*$ to $(\lambda^*\bigotimes\lambda^*)$. In equilibrium,

$$X_1^N=A_1+\sum_{j=1}^{N}\frac{(X_j^N-V_j)^+\wedge B_j}{1+D_j^N}\mathbf{I}(j\to 1) \quad \text{and} \quad X_2^N=A_2+\sum_{j=1}^{N}\frac{(X_j^N-V_j)^+\wedge B_j}{1+D_j^N}\mathbf{I}(j\to 2). \tag{10}$$

where the D_j^N are i.i.d. $\mathrm{Bin}(N-2,p_N)$ random variables. Since the indicators $\mathbf{I}(j\to 1)$ and $\mathbf{I}(j\to 2)$ are independent of each other, in equilibrium, the number of debtors of banks 1 and 2, henceforth referred to as parents of bank 1 and 2, are independent $\mathrm{Bin}(N-1,p_N)$ random variables, call them C_1^N and C_2^N respectively. Now, for any $i\in[n]$, call the set of all banks which reach i via at most k edges the ancestors of i of order k, denoted by $\mathscr{A}_i^N(k)$, and set $\mathscr{A}_i(0)=\{i\}$. Denote the joint ancestry of banks 1 and 2 up to k steps by $\mathscr{A}_{1,2}^N(k)$. Define $B_{1,2}^N=\inf\{k:\mathscr{A}_1^N(k)\cap\mathscr{A}_2^N(k)\neq\emptyset\}$. This denotes the smallest level at which banks 1 and 2 have a common ancestor. Then, if $k<B_{1,2}^N$, and if $\mathscr{A}_1(k)$ and $\mathscr{A}_2(k)$ are individually tree

like, $\mathscr{A}_{1,2}^N(k)$ consists of two disjoint trees. Suppose the latter holds. Now, consider two independent copies of λ^*:

$$Y_1 = A_1 + \sum_{i=1}^{P_1} \frac{(Y_{i,1}-V_{i,1})^+ \wedge B_{i,1}}{1+P_{i,1}} \quad \text{and} \quad Y_2 = A_2 + \sum_{i=1}^{P_2} \frac{(Y_{i,2}-V_{i,2})^+ \wedge B_{i,2}}{1+P_{i,2}}. \tag{11}$$

Recall that here, $Y_{i,1}$ and $Y_{i,2}$ are sampled from λ^*, i.i.d, the $(A_k, B_{i,k}, V_{i,k})$ are independent and $(P_{i,1}, P_{i,2}, P_1, P_2)$ are independent $\mathrm{Poi}(\alpha)$ random variables. In (11), for $k \in \{1,2\}$, replace the $P_{i,k}$ and P_k by $\mathrm{Bin}(N-2, p_N)$ and $\mathrm{Bin}(N-1, p_N)$, random variables, call them $D_{i,k}^N$ and C_k^N respectively, all else remaining the same:

$$Y_1^N = A_1 + \sum_{i=1}^{C_1^N} \frac{(Y_{i,1}^N-V_{i,1})^+ \wedge B_{i,1}}{1+D_{i,1}^N} \quad \text{and} \quad Y_2^N = A_2 + \sum_{i=1}^{C_2^N} \frac{(Y_{i,2}^N-V_{i,2})^+ \wedge B_{i,2}}{1+D_{i,2}^N}. \tag{12}$$

One can show that the distributional fixed point of (12) is unique, that is there is a unique distribution μ_N^* on the line which satisfies (12). We expand the $Y_{i,1}^N$ and $Y_{i,2}^N$ in (12) as

$$Y_{i,1}^N = A_{i,1} + \sum_{j=1}^{C_{i,1}^N} \frac{(Y_{i,j,1}^N-V_{i,j,1})^+ \wedge B_{i,j,1}}{1+D_{i,j,1}^N} \quad \text{and} \quad Y_{i,2}^N = A_{i,2} + \sum_{j=1}^{C_{i,2}^N} \frac{(Y_{i,j,2}^N-V_{i,j,2})^+ \wedge B_{i,j,2}}{1+D_{i,j,2}^N},$$

where all the random variables defined above are independent of each other. Continuing such an expansion, it can be seen that the ancestries of Y_1^N and Y_2^N form disjoint trees with roots at Y_1^N and Y_2^N, denoted by $\mathscr{T}_1^N$ and $\mathscr{T}_2^N$. Observe that up to depth $k < B_{1,2}^N$, the joint tree $(\mathscr{T}_1^N, \mathscr{T}_2^N)$ is similar in structure to $\mathscr{A}_{1,2}^N(k)$. The main idea behind the proof is to leverage this to construct a coupling between (X_1^N, X_2^N) and (Y_1^N, Y_2^N) and show that the distance between them vanishes. Since (Y_1^N, Y_2^N) converges to (Y_1, Y_2), defined in (11), (X_1^N, X_2^N) converges to (Y_1, Y_2) as $N \uparrow \infty$. *For simplicity of explanation, let the external liability $V_i = 0$, $p_N = \frac{\alpha}{N}$, and let the inter-bank liability B_i be bounded by* 1. A further outline is as follows:

3) Constructing the coupling: Set $r = 0$ and compute the fixed point of (4). Let the parents of (X_1^N, X_2^N) in $\mathscr{A}_{1,2}^N$ be enumerated as $((X_{1,1}^N, \ldots, X_{C_1^N,1}^N), (X_{1,2}^N, \ldots, X_{C_2^N,2}^N))$. Suppose $\mathscr{A}_1^N(1) \cap \mathscr{A}_2^N(1) = \emptyset$, that is, there is no common parent. Then, use the same random variable (C_1^N, C_2^N) (recall that C_1^N and C_2^N are independent $\mathrm{Bin}(N-1, p_N)$) to select number of parents of banks 1 and 2 in $(\mathscr{T}_1^N, \mathscr{T}_2^N)$. Enumerate these as $((Y_{1,1}^N, \ldots, Y_{C_1^N,1}^N), (Y_{2,1}^N, \ldots, Y_{C_2^N,2}^N))$. Since $\mathscr{A}_1^N(1) \cap \mathscr{A}_2^N(1) = \emptyset$, for $k \in \{1,2\}$, $(A_{i,k}, B_{i,k}, D_{i,k}^N)_{i=1}^{C_k^N}$ are independent over i and k. Use these as the external wealth, liabilities and out degree respectively of $(Y_{1,k}^N, \ldots, Y_{C_k^N,k}^N)$ for $k \in \{1,2\}$. Set $r \to (r+1)$ and repeat the above procedure for higher level ancestors, until $\mathscr{A}_1^N(r+1) \cap \mathscr{A}_2^N(r+1) \neq \emptyset$, that is until $r = B_{1,2}^N - 1$, and select all the subsequent random variables, (A_i, B_i, D_i^N) in $\mathscr{T}_1^N$ and $\mathscr{T}_2^N$ independent of $\mathscr{A}_{1,2}^N$ and of each other. This produces the correct marginals on (X_1^N, X_2^N) and (Y_1^N, Y_2^N) and hence is a coupling, call it π_N.

4) Bounding the distance: Call the distribution of (Y_1^N, Y_2^N) as Q_N and that of (X_1^N, X_2^N) as R_N. Suppose $B_{1,2}^N > 3$. For a coupling π, letting $\mathbf{E}_\pi(\cdot)$ denote the expectation when $((X_1^N, X_2^N), (Y_1^N, Y_2^N))$ are jointly drawn according to π,

$$W_1(Q_N, R_N) \leq \mathbf{E}_\pi(|X_1^N - Y_1^N|) + \mathbf{E}_\pi(|X_2^N - Y_2^N|), \text{ where the Wasserstein distance is taken under the } L_1 \text{ metric.} \tag{13}$$

Fix the coupling π_N. Since we have used the same random variables $(A_{i,1}, B_{i,1}, D_{i,1}^N)_{i=1}^{C_1^N}$ for the C_1^N parents of vertex 1 in both $\mathscr{T}_1^N$ and $\mathscr{A}_{1,2}^N$, the first term in (13) is bounded above by $\mathbf{E}_{\pi_N}\left(\sum_{j=1}^{C_1^N} \frac{|X_{j,1}^N - Y_{j,1}^N|}{1+D_{j,1}^N}\right)$, where $X_{j,1}^N$ and $Y_{j,1}^N$ are the parents of 1 in $\mathscr{A}_{1,2}^N$ and $\mathscr{T}_1^N$ respectively. For $l \in \{1, \ldots, C_1^N\}$, denote its $C_{l,1}^N$

parents in $\mathscr{T}_1^N$ and $\mathscr{A}_{1,2}^N$ by $(X_{l_1,1}^N,\ldots,X_{l_{C_{l,1}}^N,1}^N)$ and $(Y_{l_1,1}^N,\ldots,Y_{l_{C_{l,1}}^N,1}^N)$. Since $\mathscr{A}_1^N(2)\cap\mathscr{A}_2^N(2)=\emptyset$, set identical $(A_{l_r,1},B_{l_r,1},D_{l_r,1}^N)_{k=1}^{C_{l,1}^N}$, for its parents in both $\mathscr{T}_1^N$ and $\mathscr{A}_{1,2}^N$, and bound $|X_{l,1}^N-Y_{l,1}^N|$ by $\sum_{r=1}^{C_{l,1}^N}\frac{|X_{l_r,1}^N-Y_{l_r,1}^N|}{1+D_{l_r,1}^N}$, where $X_{l_r,1}^N$ and $Y_{l_r,1}^N$ are the parents of $X_{l,1}^N$ and $Y_{l,1}^N$ respectively.

Recall that for any bank k, $\mathscr{C}_k^N$ denotes its parents. Now, repeating the above step $B_{1,2}^N-3$ more times, one can bound the first term in (13) by

$$\mathbf{E}\left(\sum_{r_1\in\mathscr{C}_1^N}\sum_{r_2\in\mathscr{C}_{r_1}^N}\cdots\sum_{r_{B_{1,2}^N-1}\in\mathscr{C}_{r_{B_{1,2}^N-2}}^N}|X_{r_{B_{1,2}^N-1},1}^N\wedge B_{r_{B_{1,2}^N-1},1}-Y_{r_{B_{1,2}^N-1},1}^N\wedge B_{r_{B_{1,2}^N-1},1}|\cdot\prod_{l=1}^{B_{1,2}^N-1}\frac{1}{(1+D_{r_l}^N)}\right),$$

Since B_i is bounded by 1, a further upper bound is

$$\mathbf{E}\left(\sum_{r_1\in\mathscr{C}_1^N}\sum_{r_2\in\mathscr{C}_{r_1}^N}\cdots\sum_{r_{B_{1,2}^N-1}\in\mathscr{C}_{r_{B_{1,2}^N-2}}^N}\prod_{l=1}^{B_{1,2}^N-1}\frac{1}{(1+D_{r_l}^N)}\right).\tag{14}$$

Now, since $\mathscr{A}_1^N(l)$ and $\mathscr{A}_2^N(l)$ are trees, and $l<B_{1,2}^N$, the vertices r_l in the product in (14) are all distinct, and hence the $D_{r_l}^N$ are i.i.d, $\text{Bin}(N-2,p_N)$. Then, the expectation of the product $\prod_{l=1}^{B_{1,2}^N-1}\frac{1}{(1+D_{r_1}^N)}$ equals $\left(\mathbf{E}\left(\frac{1}{1+D_1^N}\right)\right)^{B_{1,2}^N-1}$. This can be bounded, for a sufficiently large N by $2\left(\frac{1-\mathrm{e}^{-\alpha}}{\alpha}\right)^{B_{1,2}^N-1}$. Next, observe that the total number of terms in the sum in (14) is the size of $\mathscr{A}_1(B_{1,2}^N-1)$. Now, (14) may be bounded for a large enough N by $2\left(\frac{1-\mathrm{e}^{-\alpha}}{\alpha}\right)^{B_{1,2}^N-1}\cdot\mathbf{E}|\mathscr{A}_1(B_{1,2}^N-1)|$, which equals $2C_\alpha\left(1-\mathrm{e}^{-\alpha}\right)^{B_{1,2}^N-1}$, for some constant C_α. It can be seen that $B_{1,2}^N<\frac{\log N}{3\log\alpha}$ has a probability of $O(N^{-\frac{1}{3}})$. Similarly for $k<\frac{\log N}{3\log\alpha}$, the probability that either $\mathscr{A}_1^N(k)$ or $\mathscr{A}_2^N(k)$ is not tree-like is $O(N^{-\frac{1}{3}})$. Since B_i are bounded by 1, $W_1(Q_N,R_N)=O(N^{-\frac{1}{3}})$. □

Remark 2 Generalisations: The Erods-Renyi assumption on $\mathscr{G}_N$ is simplest non-trivial case of random banking network structure. Our results are also applicable when the out degrees are regularly varying and even when in-degrees are themselves fat-tailed and dependent on the out-degree. However, the key structural insights which we provide are seen most easily in the Erdos-Renyi case. One may also do away with the assumption that (A_i,B_i,V_i) are conditionally i.i.d. over i, and allow some heterogeneity, e.g. based on the geographical location of the banks. Similarly, one can introduce heterogeneity in inter-bank borrowing: e.g, if bank i has an out degree D, one could assign i.i.d. weights $(e_1,\ldots,e_D)$, and then, the amount the bank borrows from its jth creditor can be set to equal $\frac{e_j}{\sum_{r=1}^D e_r}B_i$, where the random variables e_j are statistically fitted from the data. This captures the heterogeneity in the inter-bank borrowing, and facilitates the use of our approximation even when granularities of the inter-banking structure are observed.

4 SIMULATION OF A DISTRIBUTIONAL FIXED POINT

Theorem 1 shows that a large banking network with uncertainty can be approximated in distribution by the cartesian product of the fixed point of a one dimensional operator. In order to study the properties of the finite network, it is essential to be able to draw samples from the limiting fixed point. The measure we want to sample from is the fixed point of (8). The map $T(\cdot)$ defined in (8) depends on the distribution of μ, as well as other random variables (A_i,B_i,V_i). Since it is difficult to obtain a closed form expression we resort to a Monte-Carlo algorithm which gets provably close. Before stating the algorithm, we first outline our general set-up. Let $T:\mathscr{M}(\Re)\to\mathscr{M}(\Re)$ be a contractive operator, with Lipschitz constant $\beta<1$. For $\gamma>0$, define $\mathrm{K}_M^\gamma=\{\lambda\in\mathscr{M}(\Re):\mathbf{E}_{X\sim\lambda}(|X|^\gamma)\le M\}$.

Assumption 2 There exists an $\gamma > 2$, such that $T(\mathscr{M}(\Re)) \subset \mathrm{K}_M^{\gamma}$.

Is is easy to see that Assumption 2 is satisfied by the limiting operator $T(\cdot)$ defined in (8). This may be relaxed by assuming that for any distribution with bounded moment of order γ, $T(\cdot)$ outputs a distribution with bounded moment of order γ. We now outline the motivation behind the Monte-Carlo algorithm. First, observe that by the contraction property, for any distribution λ_0, the sequence of iterates $\{T^k(\lambda_0)\}_{k=1}^{\infty}$ converges to the fixed point λ^* of $T(\cdot)$. Our strategy is to replace the iterates $T^k(\cdot)$ by their empirical samples of size n. Then, provided a sufficiently large n is selected, the resulting empirical sample distribution is close to $T^k(\cdot)$ with high probability. Since $T^k(\lambda_0)$ itself converges to the limiting fixed point at a geometric rate, the empirical $\hat{T}_n^k(\lambda_0)$ will be close to the true fixed point with high probability. To the best of our knowledge, such probabilistic guarantees for distributional fixed points are new (however, see Blanchet and Sigman (2011) for some related work).

Algorithm 1: Simulation of Distributional fixed point

1. Start with an arbitrary distribution $\lambda_0 = \hat{\lambda}_0$, a fixed n (to be specified later), and set $i = 0$.
2. For $j = 1,\ldots,n$, generate i.i.d. samples $\{T(\hat{\lambda}_i^n)\}_{j=1}^n$. Call the resulting empirical distribution $\hat{\lambda}_{i+1}^n$ and set $i \leftarrow i+1$.
3. Repeat 2. for $i = 1,2,\ldots,r$, (r to be specified later).
4. Output the final empirical distribution $\hat{\lambda}_r^n$.

The following theorem gives a probabilistic bound on the performance of Algorithm 1.

Theorem 2 For every $\varepsilon > 0$ and $\delta > 0$, such that for all $n \geq C' \left(\left(\delta \varepsilon^{\gamma-\kappa} \log^{-1} \varepsilon\right)^{-\frac{1}{(\gamma-\kappa)-1}} \right)$ and $r \geq \frac{\log \frac{\varepsilon}{2E(B_1)}}{\log \beta}$, we have $\mathbf{P}(W_1(\hat{\lambda}_r^n, \lambda^*) \geq \varepsilon) \leq \delta$, for all sufficiently small $\kappa > 0$. Here, C' is a computable system dependent constant.

The following result, which gives a concentration bound for the error between the true and empirical distribution is required for the proof:

Theorem 3 [Fournier and Guillin (2015), Theorem 2] Suppose $\exists \gamma > 2$ such that $\mathbf{E}_{X\sim\lambda}(|X|^{\gamma}) < \infty$. Let $\hat{\lambda}_n$ be the empirical distribution generated by n i.i.d. samples from λ. Then, $\mathbf{P}(W_1(\hat{\lambda}_n, \lambda) \geq x) \leq a(n,x)\mathbf{I}(x \leq 1) + b(n,x)$, where $a(n,x) = K_1 \mathrm{e}^{-K_2 n x^2}$ and $b(n,x) = K_1 n (nx)^{-(\gamma-\kappa)}$, for all $\kappa \in (0,1)$ for some constants K_1 and K_2, which depend on γ, $\mathbf{E}_{X\sim\lambda}(|X|^{\gamma})$ and κ.

Proof of Theorem 2 For any distribution μ on the line, let $\hat{T}_n(\mu)$ denote the empirical distribution of size n drawn from $T(\mu)$. Then, observe that for all i, $\hat{\lambda}_{i+1}^n$, as defined in Algorithm 1 is $\hat{T}_n(\hat{\lambda}_i^n)$. Now, consider the following sequence of bounds where λ^* is the fixed point of $T(\cdot)$ which we wish to sample from, and where we define recursively $\lambda_k = T(\lambda_{k-1})$:

$$
\begin{aligned}
W_1(\lambda^*, \hat{\lambda}_r^n) &\leq W_1(\lambda^*, \lambda_r) + W_1(\lambda_r, \hat{\lambda}_r^n) \text{ by the triangle inequality} \\
&\leq \beta^r W_1(\lambda^*, \lambda_0) + W_1(\lambda_r, \hat{\lambda}_r^n) \text{ since } \lambda^* \text{ is the fixed point, by contraction} \\
&\leq \beta^r W_1(\lambda^*, \lambda_0) + W_1(T(\lambda_{r-1}), T(\hat{\lambda}_{r-1}^n)) + W_1(T(\hat{\lambda}_{r-1}^n), \hat{T}_n(\hat{\lambda}_{r-1}^n)) \\
&\leq \beta^r W_1(\lambda^*, \lambda_0) + \beta W_1(\lambda_{r-1}, \hat{\lambda}_{r-1}^n) + W_1(T(\hat{\lambda}_{r-1}^n), \hat{T}_n(\hat{\lambda}_{r-1}^n)) \qquad (15) \\
&\leq \beta^r W_1(\lambda^*, \lambda_0) + \sum_{i=1}^{r} \beta^{i-1} W_1(T(\hat{\lambda}_{r-i}^n), \hat{T}_n(\hat{\lambda}_{r-i}^n)) \text{ , by (15) and contraction.} \qquad (16)
\end{aligned}
$$

The first term of (16) is deterministic for a given λ_0. Select r so that $\beta^r W_1(\lambda^*, \lambda_0) < \frac{\varepsilon}{2}$, e.g., select $r = r_{\varepsilon} = \left\lceil \frac{1}{\log \beta} \log \frac{\varepsilon}{2E(B_1)} \right\rceil$. Then, the first part of the theorem follows. To complete the proof, we shall bound the probability that $\sum_{i=1}^{r} \beta^{i-1} W_1(T(\hat{\lambda}_{r-i}^n), \hat{T}_n(\hat{\lambda}_{r-i}^n))$ exceeds $\frac{\varepsilon}{2}$ by δ. Define $A_{\varepsilon,\delta} = \bigcup_{i=1}^{r_{\varepsilon}} \left\{ W_1(T(\hat{\lambda}_i^n), \hat{T}_n(\hat{\lambda}_i^n)) \geq \right.$

$\frac{(1-\beta)\varepsilon}{2}\}$ and observe that by the union bound,

$$\mathbf{P}(A_{\varepsilon,\delta}) \leq \sum_{i=1}^{r_\varepsilon} \mathbf{P}\left(W_1(T(\hat{\lambda}_i^n), \hat{T}_n(\hat{\lambda}_i^n)) \geq \frac{(1-\beta)\varepsilon}{2}\right). \tag{17}$$

Further, observe that restricted to $A^c_{\varepsilon,\delta}$, $\sum_{i=1}^{r_\varepsilon} \beta^{i-1} W_1(T(\hat{\lambda}_i^n), \hat{T}_n(\hat{\lambda}_i^n)) \leq \frac{\varepsilon}{2}$. We now develop uniform bounds for the probabilities in (17). By Assumption 2, $\sup_{\lambda \in \mathscr{M}(\Re)} \mathbf{E}_{T(\lambda)}(|X|^\gamma) \leq M$ for some $\gamma > 2$. Then from Lemma 13 of Fournier and Guillin (2015), the constant K_2 in Theorem 3 above works uniformly for all i. Hence, fix $i = 1$. We bound $\mathbf{P}\left(W_1(T(\hat{\lambda}_1^n), \hat{T}_n(\hat{\lambda}_1^n)) \geq \frac{(1-\beta)\varepsilon}{2}\right)$ by $\frac{\delta}{r_\varepsilon}$ which ensures $\mathbf{P}(A_{\varepsilon,\delta}) \leq \delta$. Set $n \geq C' \left(\delta \varepsilon^{(\gamma-\kappa)} \log^{-1} \varepsilon\right)^{-\frac{1}{(\gamma-\kappa)-1}}$. The rest of the proof now follows from an application of Theorem 3, replacing ε by $\frac{(1-\beta)\varepsilon}{2}$ and δ by $\frac{\delta}{r_\varepsilon}$. The constant C' can be calculated from the system parameters. □

5 COMPUTATIONAL GUARANTEES

In analysis of computational efficiency of a numerical algorithm it is often beneficial to study its behaviour constrained to a large but fixed budget, Γ, which asymptotically increases to infinity. We demonstrate how to distribute this budget well, so as to get a good upper bound on average error in the algorithm, and optimise this upper bound. Notice that the average computation required in Algorithm 1 is $n \times r$ (since step 2 is repeated r times, and we assume for simplicity that each sampling on average takes unit time). Taking expectations of (16)

$$\mathbf{E}(W_1(\lambda^*, \hat{\lambda}_r^n)) \leq \beta^r W_1(\lambda^*, \lambda_0) + \sum_{i=1}^{r} \beta^{i-1} \mathbf{E}(W_1(T(\hat{\lambda}_{r-i}^n), \hat{T}_n(\hat{\lambda}_{r-i}^n))). \tag{18}$$

Now, by Assumption 2 and Fournier and Guillin (2015), Theorem 1, the RHS of (18) is bounded above by

$$\beta^r W_1(\lambda^*, \lambda_0) + \frac{C}{1-\beta} n^{-\frac{1}{2}} + \text{ smaller terms}, \tag{19}$$

for some constant C, depending on the system parameters. Then, to derive a good upper bound on the computation required, one minimises (19) subject to $r \cdot n \leq \Gamma$. It can be seen that up to first order terms, the optimum is achieved by $r^* = C_1 \log \Gamma$ and $n^* = C_2 \frac{\Gamma}{\log \Gamma}$, for some constants C_1 and C_2. This guarantees an error of $\mathrm{K}\left(\sqrt{\frac{\log \Gamma}{\Gamma}}\right) +$ smaller terms, for a constant K which can be computed form C_1 and C_2.

Remark 3 Iteration dependent sampling To further improve performance, one may choose n dependent on the current iteration i. For instance, a carefully chosen increasing sequence $\{n_1, \ldots, n_r\}$ may be selected in Step 2 of the algorithm, instead of a uniform n. The intuition here is that since for small r, $T^r(\lambda_0)$ is far from the true fixed point, one can tolerate more noise in approximating it. The following reduces the error from order $\sqrt{\frac{\log \Gamma}{\Gamma}}$ to order $\Gamma^{-\frac{1}{2}}$. Let $r = \frac{\log \Gamma}{2 \log \beta}$, and set $n_k = \frac{\beta^{r-k}}{\sum_i \beta^i} \Gamma$ for $k \leq r$. From (18) and (19), the error in estimation now becomes $\frac{C_\beta}{\sqrt{\Gamma}}$, where C_β is a β dependent constant.

Approximating a large banking system We next derive the computation required to accurately approximate a banking system of size N. As a measure of performance, we use the average expected surplus of the network, defined as the net profit made by all the banks in the economy, $S^N \triangleq \frac{1}{N} \sum_{i=1}^N (X_i^N - V_i - B_i)^+$. A limiting approximation to this is $\mathbf{E}_{X \sim \lambda^*}(X - B_1 - V_1)^+$, the surplus computed by the limiting system. However, we can draw samples only from the approximation to the limiting distribution λ^*. Let $\hat{\lambda}^*_\Gamma$ be the random distribution output by Algorithm 1. In order to capture the two-sided of error in estimating the surplus, we consider the square error

$$\left(\mathbf{E}_{\{X_i\} \sim \lambda_N^*}\left(\frac{1}{N} \sum_{i=1}^N (X_i^N - B_i - V_i)^+\right) - \mathbf{E}_{X \sim \hat{\lambda}^*_\Gamma}(X - B - V)^+\right)^2. \tag{20}$$

The following guarantee is obtained:

Lemma 2 Consider the problem of approximating a banking network of size N using the limiting fixed point. Suppose that the total computing power of Γ. Then,

$$\left(\mathbf{E}_{\{X_i\}\sim\lambda_N^*}\left(\frac{1}{N}\sum_{i=1}^{N}(X_i^N - B_i - V_i)^+\right) - \mathbf{E}_{X\sim\hat{\lambda}_\Gamma^*}(X-B-V)^+\right)^2 = O_{\mathscr{P}}(N^{-1}) + O_{\mathscr{P}}\left(\frac{\log\Gamma}{\Gamma}\right). \tag{21}$$

From (21), it follows that for $\Gamma = O(N\log N)$, the two error terms are roughly of order $O_{\mathscr{P}}(N^{-1})$. Any further increase in order of Γ no longer reduces the order of the error. We can set $\Gamma = O(N)$ and achieve error $O_{\mathscr{P}}(N^{-1})$ by using the iteration dependent sampling outlined in Remark 3.

Proof Sketch: Bound (20) above by 2 times

$$\left(\mathbf{E}_{\{X_i\}\sim\lambda_N^*}\left(\frac{1}{N}\sum_{i=1}^{N}(X_i^N - B_i - V_i)^+\right) - \mathbf{E}_{X\sim\lambda^*}(X-B-V)^+\right)^2 + \left(\mathbf{E}_{X\sim\lambda^*}(X-B-V)^+ - \mathbf{E}_{X\sim\hat{\lambda}_\Gamma^*}(X-B-V)^+\right)^2. \tag{22}$$

Notice that for any (b,v), $f(x) = (x-b-v)^+ \in \mathrm{Lip}(1)$. By an application of the duality in (2), (22) is bounded above by $W_1^2(\hat{\lambda}_\Gamma^*, \lambda^*) + O(N^{-1})$. This together with the explanation following (19) completes the proof. □

Observe that the squared bias between the finite and the limiting system:

$$\left(\mathbf{E}_{\{X_i\}\sim\lambda_N^*}\left(\frac{1}{N}\sum_{i=1}^{N}(X_i^N - B_i - V_i)^+\right) - \mathbf{E}_{X\sim\lambda^*}(X-B_1-V_1)^+\right)^2 = O(N^{-1}).$$

Recall that the finite banking network has a complicated dependence among the wealth of the banks due to inter-connectivity in the network. This is a function of the number of counter-parties of a typical bank. Because the in and out degree distributions of various banks have bounded second moments, the dependence disappears rapidly. This results in a $O(N^{-1})$ rate of decay of the square of the bias error.

6 NUMERICAL SIMULATIONS

We now present a numerical study to validate the theory developed in the previous sections, and provide some structural insights. We assume that B_i have a fat tailed distribution with a complimentary CDF $\bar{F}_B(x) \sim x^{-3}$, and that external liabilities are zero. The external assets $A_i = \varepsilon_i \cdot E_i$, where E_i has a complimentary CDF $\bar{F}_E(x) \sim x^{-3}$ and ε_i is a binary random variable, taking value 1 with probability 0.9, and 0 otherwise, independent of E_i. Here, ε_i represents the idiosyncratic shock to the banks external assets, and we assume that occurrence of a shock wipes out external assets held by a bank. Lastly, in all experiments, A_i and B_i are taken to be independent of each other.

In the first set of experiments, we test the accuracy of the approximation produced by the algorithm. Throughout, we set the average connectivity $\alpha = 5$. Table 1 shows the behaviour of the network surplus and default probability as a function of its size. It can be seen that all else kept the same, the average surplus remains somewhat constant with network size, while the variance error decays inversely as system size. Table 2 shows the performance of Algorithm 1 as a function of computational power. Recall that the limiting surplus computed by Algorithm 1, $S(\hat{\lambda}^*) \triangleq \mathbf{E}_{X\sim\hat{\lambda}^*}(X-B_1)^+$ is a random variable, where $\hat{\lambda}^*$ is the output random measure. To find the limiting expected surplus, we run Algorithm 1 L times independently, and return the estimate $\hat{S} = \frac{1}{L}\sum_{k=1}^{L} S(\hat{\lambda}_k^*)$ and compute the limiting default probability in a similar manner.

Even when there are as few as 200 banks in the network, Tables 1 and 2 show that the estimated mean surplus and default probabilities in the finite system are well approximated by the limiting system. Table 2

Table 1: The expected surplus and default probability of a finite banking network of size N are estimated by simulating 2000 networks and evaluating the empirical average. Further, a confidence interval is evaluated around the estimated mean as $1.96 \times \sigma_N$, where σ_N is the estimated standard deviation.

N	Estimated mean surplus	Estimated default probability
100	1.339 ± 0.178	0.126 ± 0.063
200	1.341 ± 0.128	0.126 ± 0.043
500	1.341 ± 0.083	0.127 ± 0.027
1000	1.341 ± 0.058	0.127 ± 0.019

Table 2: Here, we calculate the expected surplus and average default probabilities computed by the random distribution output by Algorithm 1. The quantity (r,n) indicates that step 2 of the algorithm has been run r times, and at each step, n samples have been drawn. Here, $L = 200$ independent runs of Algorithm 1 are used to compute the limiting expected surplus and default probabilities. Confidence intervals around the mean are found similar to those in Table 1. It is seen that the width of the confidence interval shrinks as $n^{-\frac{1}{2}}$.

Computation (r,n)	Limiting expected surplus	Limiting default probability
$(3,500)$	1.335 ± 0.095	0.125 ± 0.031
$(4,1000)$	1.341 ± 0.074	0.125 ± 0.023
$(4,2000)$	1.341 ± 0.051	0.127 ± 0.016
$(5,3000)$	1.338 ± 0.038	0.127 ± 0.011

further suggests that only a small amount of computational effort is required to accurately approximate these.

In the second set of experiments, we study the effect of connectivity on system surplus and default probability. It is observed (see Table 3) that highly connected systems have lesser default probabilities as well as larger surpluses. The intuition behind this is as follows: higher connectivity in the network allows a diversification of risk throughout the network, and thus decreases the probability of knock-on effects triggering a cascade of defaults. Our results also show that the limiting approximation works well over a wide range of connectivities.

Table 3: Connectivity parameter α is variable. Here, $N = 500$, and 2000 Erdos Renyi networks are simulated, and the average surplus and default probabilities are estimated. Similarly, limiting average surplus and default probability are computed using Algorithm 1, and a confidence interval is constructed around the mean as before.

α	Limiting surplus	Estimated surplus	Limiting default probability	Estimated default Probability
2	1.177 ± 0.096	1.171 ± 0.091	0.199 ± 0.034	0.204 ± 0.035
5	1.335 ± 0.095	1.341 ± 0.083	0.125 ± 0.031	0.127 ± 0.027
7	1.345 ± 0.092	1.350 ± 0.083	0.117 ± 0.026	0.113 ± 0.027
10	1.347 ± 0.090	1.355 ± 0.082	0.106 ± 0.027	0.104 ± 0.025

Remark 4 In the case where $A_i = 0$ a.s. it is easy to see that δ_0 is the unique fixed point of (8). We find that the Wasserstein-1 distance between the simulated fixed point with $(r,n) = (4,1000)$ and 0 random variable is 0.022, thus validating the algorithm on an example where the true limiting fixed point is known.

7 Conclusions and Future Work

We analysed a large inter-connected network of banks in presence of incomplete information. We developed distributional limit theorems to approximate the behaviour of a large banking network, and a Monte-Carlo algorithm for simulating this limiting behaviour. We found that even when there are a few hundred banks in the network, the approximations worked well and that these distributional limits provided structural insights into the behaviour of the system. Typically, degree distribution in banking networks are seen to have fat tails. Further, there often exist banks which account for a finite fraction of the systems wealth. To this end in an extended version of this work, we will present a more realistic model where the in and out degrees of banks in the network are positively correlated, and there exist large banks interacting with every other bank in the network, and will derive limiting results.

REFERENCES

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

Allen, F., and D. Gale. 2000. "Financial contagion". *Journal of political economy* 108(1):1–33.

Amini, H., R. Cont, and A. Minca. 2016. "Resilience to contagion in financial networks". *Mathematical Finance* 26(2):329–365.

Anand, K., B. Craig, and G. Von Peter. 2015. "Filling in the blanks: Network structure and interbank contagion". *Quantitative Finance* 15(4):625–636.

Blanchet, J., and K. Sigman. 2011. "On exact sampling of stochastic perpetuities.". *Journal of Applied Probability* 48(A):165–182.

Cont, R., and A. Moussa. 2010. "Network structure and systemic risk in banking systems". *Edson Bastos e*.

Eisenberg, L., and T. Noe. 2001. "Systemic risk in financial systems". *Management Science* 47(3):236–249.

Fournier, N., and A. Guillin. (2015). "On the rate of convergence in Wasserstein distance of the empirical measure". *Probability Theory and Related Fields* 162 no 3-4:707–738.

Gai, P., and S. Kapadia. 2010. "Contagion in financial networks.". *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 466* 2120:2401–2423.

Glasserman, P., and H. P. Young. 2016. "Contagion in financial networks.". *Journal of Economic Literature* 54(3):779–831.

Haldane, A. G., and R. M. May. 2011. "Systemic risk in banking ecosystems". *Nature* 469(7330):351.

Kavitha, V., I. Saha, and S. Juneja. 2018. "Random Fixed Points, Limits and Systemic risk". In *IEEE Conference on Decision and Control (CDC)*, 5813–5819. IEEE.

Villani, C. 2003. *Topics in optimal transportation. No. 58*. American Mathematical Soc.

AUTHOR BIOGRAPHIES

ANAND DEO is a PhD candidate in the School of Technology and Computer Science at the Tata Institute of Fundamental Research, Mumbai, India. He received a Bachelor's degree in Electronics Engineering from Mumbai University. His research interests are broadly in applied probability, and specifically in finance and Operations Research. His email address is anand.deo@tifr.res.in.

SANDEEP JUNEJA is a Professor and Dean at the School of Technology and Computer Science in Tata Institute of Fundamental Research in Mumbai. He is currently on the editorial board of Stochastic Systems. Earlier he has been on editorial boards of Mathematics of Operations Research, Management Science and ACM TOMACS. His research interests lie in applied probability including in sequential decision making using bandit methods, financial mathematics, Monte Carlo methods, and game theoretic analysis of queues. His email address is juneja@tifr.res.in.