

May's instability in large economies

José Moran¹ and Jean-Philippe Bouchaud²

¹*Centre d'Analyse et de Mathématique Sociales, EHESS, 54 Boulevard Raspail, 75006 Paris, France*

²*Capital Fund Management, 23 Rue de l'Université, 75007 Paris, France*



(Received 1 March 2019; revised manuscript received 7 July 2019; published 18 September 2019)

Will a large economy be stable? Building on Robert May's original argument for large ecosystems, we conjecture that evolutionary and behavioural forces conspire to drive the economy towards marginal stability. We study networks of firms in which inputs for production are not easily substitutable, as in several real-world supply chains. Relying on results from random matrix theory, we argue that such networks generically become dysfunctional when their size increases, when the heterogeneity between firms becomes too strong, or when substitutability of their production inputs is reduced. At marginal stability and for large heterogeneities, we find that the distribution of firm sizes develops a power-law tail, as observed empirically. Crises can be triggered by small idiosyncratic shocks, which lead to “avalanches” of defaults characterized by a power-law distribution of total output losses. This scenario would naturally explain the well-known “small shocks, large business cycles” puzzle, as anticipated long ago by Bak, Chen, Scheinkman, and Woodford.

DOI: [10.1103/PhysRevE.100.032307](https://doi.org/10.1103/PhysRevE.100.032307)

I. INTRODUCTION

Why is the output of large economies so volatile? Why do small idiosyncratic fluctuations lead to large business cycles? These questions have been at the forefront of economic research for decades [1–4].

Naively, one would expect that the fluctuations of an economy made of N independent sectors should decay rather quickly, as $N^{-1/2}$ [4,5] because of the central limit theorem. In order to explain why fluctuations survive at the aggregate level, three families of explanations have been proposed in the literature. The first one is that aggregate fluctuations are driven by global shocks, that affect all firms and sectors simultaneously. However, it is often not clear what these shocks might be¹ and, when identified, they appear too small to be responsible for the observed volatility of the aggregate industrial production. Bernanke *et al.* [3] have called this the *small shocks, large cycles puzzle*. One interesting possibility is that these shocks are self-fulfilling prophecies [6], perhaps due to collective opinion shifts or trust collapse; see, e.g., [7–10] for various strands of literature on the subject.

Another resolution has been proposed by Gabaix [11] and, in a slightly different context, by Wyart and Bouchaud [12]; see also [13]. The argument is that the fat-tailed distribution of sizes of independent firms and sectors slows down the regression of fluctuations from the standard $N^{-1/2}$ behavior to $N^{-\alpha}$, with $\alpha \leq 1/2$ related to the tail exponent of the distribution. Although some empirical support for this scenario has been put forth [11,14], other works suggest that *network*

effects are in fact dominant [15–17], as idiosyncratic shocks can cascade along the input-output network and eventually become macroscopic.² One particular stigma of these network effects is the strong covariation of fluctuations across different sectors [22]—but see also [23].

While the cascade story is enticing, the baseline Cobb-Douglas network model proposed by Acemoglu *et al.* [1,15] is, in our view, not convincing. Indeed, the only way to escape the $N^{-1/2}$ decay of fluctuations within this framework is to assume that the supply network is very unbalanced, i.e., that a few sectors are crucial suppliers to the whole economy [24]. This somehow throws out the baby with the bathwater, as it reintroduces the idea of aggregate shocks in disguise. A possible way out was proposed in Ref. [25]: by introducing myopia and frictions in the Cobb-Douglas network model, it was found that the general equilibrium solution of Acemoglu *et al.* [15] is only stable in a certain region of parameters, outside of which large fluctuations emerge *endogenously*, i.e., without any microscopic shocks. These fluctuations arise from a breakdown of coordination between the different sectors and illustrate how some mechanisms present in the real world may cause the economy to be intrinsically unstable.

This idea (that the economy may in itself be unstable and turbulent) was in fact already mentioned in 1948 by Hawkins [26] (see also [27]) and picked up again by Bak, Chen, Scheinkman, and Woodford [28,29] in the context of “self-organized critical” (SOC) states in complex systems [30]. In such a state, small microscopic perturbations give rise to macroscopic fluctuations—like avalanches in sand piles.

¹As Cochrane quipped [2], “What shocks are responsible for economic fluctuations? Despite at least two hundred years in which economists have observed fluctuations in economic activity, we still are not sure.”

²Similar network effects have been argued to be at the origin of systemwide breakdowns of the banking sector [18–21].

Similar ideas have emerged in the context of theoretical ecology. In his seminal paper, *Will a Large Complex System be Stable?*, May [31] argued that a large number of very different species can lead an ecosystem to instabilities and mass extinctions. May's paradigm has recently been made much more explicit in the context of a generalized Lotka-Volterra model in Refs. [32–34], where it is shown that the system indeed spontaneously evolves towards a marginally stable state that is anomalously sensitive to small perturbations. Unfortunately, this stream of ideas has not gained much traction in the economics literature, perhaps for lack of a convincing modeling framework. (See however [18–21] for financial network models with explicit references to ecosystems, and [35] for a very recent contribution.)

The aim of the present work is to present an economically motivated model where marginal stability appears naturally, and leads to an amplification of small, idiosyncratic shocks along the input-output network. Interestingly, our model is closely related to the ecological models alluded to above, and builds upon the classic—but still extremely active—field of random matrix theory, that describes the statistical properties of the eigenvalues and eigenvectors of certain families of random matrices, here related to the input-output matrix. In particular, the *feasibility* of an equilibrium, defined as the existence of an economically sound set of prices and production outputs, depends strongly on the eigenvalues of such a matrix. We define here the stability of the economy as the resilience of such an equilibrium to idiosyncratic shocks. We find that for a fixed number of firms N , increased interlinkages, profit maximization, and/or reduced substitutability drive the system at the edge of instability. Similarly, increasing the number of firms at fixed productivity also leads to a critical state. At criticality, small idiosyncratic shocks can lead to bankruptcy avalanches, which, depending on the topology of the network and the heterogeneity of firms' productivity, can be either small and localized or systemwide, with all possible gradations.

In the present paper, we only describe the equilibrium (or absence thereof) aspects of our model, leaving the analysis of its truly dynamical features—crucial when crises occur—for a forthcoming publication [36].

II. MODEL

We consider N firms with a given input-output network determined by the technology available to firms. The “technology network” is a directed graph where nodes represent firms and where a directed edge $j \rightarrow i$ exists if i needs goods produced by j for its own production. Note that this framework allows for self-loops, with an edge $i \rightarrow i$ existing if a firm produces one of its own inputs. The node $i = 0$ conventionally represents households, and supplies firms with labor while consuming a part of their output. The link $j \rightarrow i$ carries a “stoichiometric weight” J_{ij} , measuring the number of j goods needed to make an unit of i 's production. The set of suppliers of i is thus given by $\{j/J_{ij} \neq 0\}$, while the set of clients is $\{j/J_{ji} \neq 0\}$. The production of i , π_i , is given by a so-called CES (constant elasticity of substitution) function,

which reads [37]³

$$\pi_i = z_i \left(\sum_j a_{ij} \left(\frac{J_{ij}}{Q_{ij}} \right)^{1/q} \right)^{-q}, \quad \text{with} \quad \sum_j a_{ij} = 1, \quad (1)$$

where z_i is the firm's productivity, Q_{ij} is the number of goods firm i buys from firm j , and $a_{ij} \geq 0$ are weight parameters. The parameter q measures the global substitutability of the different inputs. When $q \rightarrow 0$, no substitutes are available and Eq. (1) reduces to the classical Leontief production function:

$$\pi_i = z_i \min_j \left(\frac{Q_{ij}}{J_{ij}} \right). \quad (2)$$

The Cobb-Douglas function $\pi_i = z_i \prod_j (Q_{ij}/J_{ij})^{a_{ij}}$ corresponds to $q \rightarrow \infty$ and is often used to describe the average aggregate production of economic sectors [1], or of the economy as a whole. In a Cobb-Douglas economy, the loss of a fraction f of good j can always be compensated for by an increase of any other good k by a factor $1/f^{a_{ij}/a_{ik}}$. In the Leontief case, on the other hand, the loss of a fraction f of good j cannot be compensated and translates to an immediate loss of the same fraction f of total production π_i . It models a situation where redundancy is costly. Firms therefore choose their suppliers with parsimony and cannot “rewire” (i.e., find alternative suppliers) on short time scales in the real economy. For example, in the aftermath of the 2011 tsunami and Fukushima Daiichi nuclear power plant disaster, the shortage of a few, seemingly unimportant components had a severe impact on the car industry [38,39], very dependent on products manufactured in Japan. The incident highlighted how competition led firms to have a very tight supply-chain strategy, as in the words of an observer: “In the race to provide better quality at lower prices, manufacturers picked very narrow, optimized supply chains” which caused them to be very dependent on the “one supplier that had the best product at the lowest price” [40]. At the time of the disruption, firms had to swiftly rethink their supply-chain strategy and large-scale rewirings of the production network took place. The influence of possible rewirings is beyond the scope of this paper, but for preliminary work in that direction see [41].⁴ In the following, we will for simplicity focus on the extreme case of a Leontief production function $q \rightarrow 0$, but will show that our results hold true in a range $q \in [0, q_c]$, where the critical value q_c depends on the network and on the productivities.

Calling p_i the price of the goods produced by firm i , its profit \mathcal{P}_i reads

$$\mathcal{P}_i = p_i \pi_i - \sum_{j \neq 0} Q_{ij} p_j - Q_{i0} p_0, \quad (3)$$

where p_0 is the labor wage. Optimizing the profit with respect to all inputs Q_{ij} (including labor Q_{i0}) leads, for $q \rightarrow 0$, to the

³For a more in-depth exploration of these production functions, see Appendix A.

⁴One could consider a case where firms have several possible suppliers j within the same sector J and write $\pi_i = z_i \min_j \max_{j \in J} (\frac{Q_{ij}}{J_{ij}})$. This extension will be studied at a later stage of the project.

condition

$$\forall(i, j), \exists \gamma_i \geq 0 \quad \text{such that} \quad Q_{ij} = \gamma_i J_{ij} \quad (4)$$

which can also be interpreted as saying that given an output level $\gamma_i := \pi_i/z_i$, the optimal choice for inputs Q_{ij} is to pick them proportionally to their stoichiometric weight, as buying more would result in waste. In this case, profit can be written as

$$\mathcal{P}_i = \gamma_i \left(z_i p_i - \sum_{j \neq 0} J_{ij} p_j - J_{i0} p_0 \right). \quad (5)$$

We now assume that households' optimal consumption of good i , given a certain utility function and a vector of prices, are given by $C_i > 0$.⁵

As standard in the literature, we now assume the following:

(i) Market clearing, i.e., every good that is produced is either consumed by households or bought by other firms for their own production. Hence

$$\pi_i = \sum_{j \neq 0} Q_{ji} + C_i \longrightarrow z_i \gamma_i - \sum_{j \neq 0} J_{ji} \gamma_j = C_i \quad (> 0). \quad (7)$$

(ii) Competitive equilibrium, i.e., competition drives profits to zero. Hence

$$\mathcal{P}_i = 0 \longrightarrow z_i p_i - \sum_{j \neq 0} J_{ij} p_j = V_i \quad (> 0), \quad (8)$$

where we have defined $V_i = J_{i0} p_0$ and imposed that $\gamma_i \neq 0, \forall i$ [otherwise Eq. (7) cannot be satisfied]. One could also model firms attempting to impose markups to reach a positive profit equal to a fraction φ_i of its sales $z_i \gamma_i p_i$. This simply amounts to shifting z_i to $z_i(1 - \varphi_i)$ in Eq. (8).

Now, in order for the equilibrium to make sense, the solutions to Eqs. (7) and (8) must be such that $\gamma_i > 0$ and $p_i > 0, \forall i$; i.e., that equilibrium prices and quantities must be strictly positive. As first noted by Hawkins and Simon [27], this is not automatic and requires the matrix \mathbf{M} , defined by $(\mathbf{M})_{ij} = z_i \delta_{ij} - J_{ij}$, to be a so-called “ M matrix,”⁶ i.e., such that *all its eigenvalues have non-negative real parts* [42]. Therefore some conditions on productivities and linkages must be fulfilled for the economy to work.

This condition is the equivalent, in an ecological context, of May's stability criterion that allows the equilibrium population of all species to be strictly positive [32,33]. Rather interestingly, Eqs. (7) and (8) are identical, *mutatis mutandis*, to the equation determining the equilibrium size of species in a generalized Lotka-Volterra model [34].

⁵For example, for a utility function $\mathcal{U} = \sum_i \theta_i \ln(C_i)$ and a certain budget B , the optimal consumption C_i is

$$C_i = \frac{B}{\sum_j \theta_j} \frac{\theta_i}{p_i} := \frac{\Gamma_i}{p_i}, \quad (6)$$

but any other type of utility function would work in our model, as long as consumption levels were strictly positive.

⁶Note that if \mathbf{M} is an M matrix, \mathbf{M}' is also an M matrix. An interesting property of an M matrix is that all the elements of its inverse are non-negative.

In the case of a more general CES production function with $q \geq 0$, the competitive equilibrium equation reads

$$(z_i p_i)^\zeta - \sum_{j \neq 0} a_{ij}^{q\zeta} (J_{ij} p_j)^\zeta = V_i \quad (> 0), \quad \zeta := \frac{1}{1+q}, \quad (9)$$

which boils down to Eq. (8) when $q = 0$ [for a detailed proof, see Eq. (B5) in Appendix C]. Interestingly, setting $\hat{p}_i = p_i^\zeta$, one finds again that the condition for an admissible equilibrium is that the matrix $(\hat{\mathbf{M}})_{ij} = z_i^\zeta \delta_{ij} - a_{ij}^{q\zeta} J_{ij}^\zeta$ is an M matrix.

Note that since $\sum_{j \neq 0} a_{ij} < 1$, the Perron-Frobenius theorem ensures that Cobb-Douglas networked economies (such as those considered in Acemoglu *et al.* [15], and corresponding to $q \rightarrow \infty$), *always* have an admissible equilibrium, for any network and any productivities. Therefore, the type of shock propagation that takes place in our model has no counterpart in a Cobb-Douglas economy.

III. STABILITY CONDITIONS FOR MODEL NETWORKS

Here and below we will for simplicity focus on the Leontief case, commenting on the more general case $q > 0$ in the conclusion. In order to gain some intuition on the stability conditions, let us first consider a random directed network, where each supply link J_{ij} is equal to J with probability r and 0 with probability $1 - r$, and where all N firms have the same productivity z . The spectrum of \mathbf{M} in this case is well known when $N \gg 1$ and $r \sim \mathcal{O}(1)$. It consists of an isolated eigenvalue $\lambda_{\min} = z - JrN$ and a “sea” of complex eigenvalues uniformly distributed in a disk of radius $J\sqrt{r(1-r)N}$ centered at z (see, e.g., [43,44]). The stability condition therefore reads $z > JrN$, i.e., productivity must be large enough for the economy to function. The most unstable eigenvector, corresponding to eigenvalue λ_{\min} , is the uniform vector $(1/\sqrt{N}, \dots, 1/\sqrt{N})$. As will be clear below, this corresponds to a case where crises are system wide. The same qualitative result holds when productivities are weakly heterogeneous, i.e., $z_i = z(1 + \epsilon_i)$ with $\epsilon_i \ll 1$ [albeit λ_{\min} is slightly shifted downwards by an amount $\mathcal{O}(z^2 \epsilon^2 / JN)$].

More interesting—but more complex!—is the case where the average number of suppliers $c = rN$ is of order unity [i.e., when $r = \mathcal{O}(N^{-1})$]. In the *random regular network* (RRN) where each firm has exactly c suppliers (and c customers) chosen randomly among the $N - 1$ other firms, one knows that the spectrum of \mathbf{M} again consists of an isolated eigenvalue $\lambda_{\min} = z - Jc$ and a “sea” of complex eigenvalues distributed in a disk of radius $J\sqrt{c}$ centered at z .⁷ When heterogeneity is introduced, either topological (i.e., letting the number of suppliers or customers vary) or because the couplings J and the productivities z fluctuate, there are no exact results available, in particular in the case where \mathbf{M} is not a symmetric matrix; see [45] for a very recent survey.

When \mathbf{M} is symmetric, exact results are still scarce but a huge amount of work has been done in the physics, mathematics, and computer science literature to characterize

⁷In this case, however, the density of complex eigenvalues is not uniform but is given by $\rho(\lambda) \propto (c^2 - |\lambda|^2)^{-2}$ for $|\lambda| < \sqrt{c}$ [45].

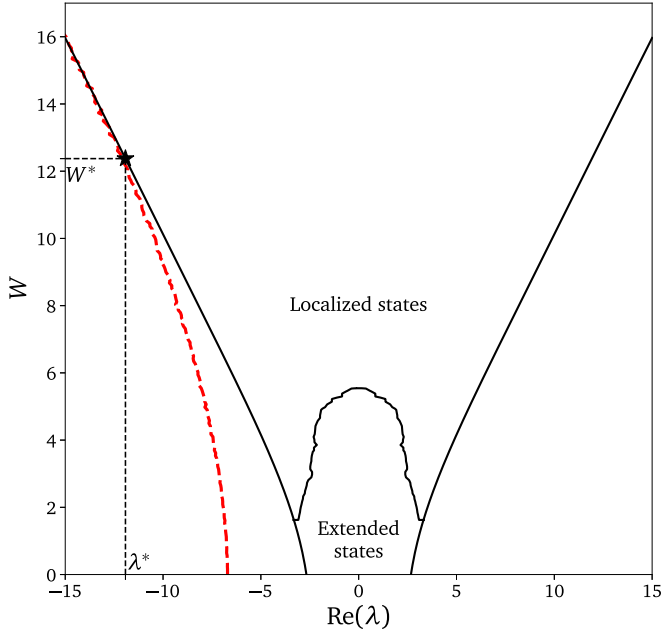


FIG. 1. Numerical results for the structure of the eigenstates of a directed random regular network (RRN) with $N = 2000$ firms with $c = 7$ suppliers and clients each and $J = 1$. Productivities z_i are uniformly distributed in an interval $[z - W/2, z + W/2]$, but the x axis is centered around z . Notice that the eigenstates contained in the bulk get localized as W increases. The leftmost dashed red curve corresponds to the isolated eigenvalue that gets absorbed at the point marked by a black star in the graph, corresponding to values $W^* \approx 12.4$ and $\lambda^* \approx -12$. The boundary between extended and localized states is defined here by $H = 5/N$. Compare to Fig. 1 in [55], for the case of undirected RRNs.

the eigenvalues and eigenvectors of such random matrices [46–55]. The reason is that such symmetric random matrices appear in many physical situations, such as the vibration spectrum of amorphous solids or the energy spectrum of quantum systems with impurities. Such random matrices also appear in graph theory and computer science. The problem of estimating the extremal eigenvalue is of special importance, as it appears in many different problems (such as epidemic or rumor spreading [56]—or crisis propagation as in the present work); the associated eigenvector is related to the concept of node centrality in network theory [57]; see also [58] and references therein.

In the sequel, we will call Δ^2 the variance of fluctuations (of connectivity, productivity, etc.). From the host of results accumulated in the last decades, the following general scenario is expected (see [55] and Fig. 1 for the case of random regular graphs):

- (i) For $\Delta = 0$, all eigenvalues except one have their real part confined in a certain interval \mathcal{I} ($= [z - \sqrt{c}, z + \sqrt{c}]$ in the RRN example), while the isolated eigenvalue is located to the left of this interval, at a nonzero distance g from its edge.
- (ii) As Δ increases, the interval \mathcal{I} broadens and its edges become somewhat blurred, while the isolated eigenvalue gets closer and closer to the lower edge of \mathcal{I} (see Fig. 1).
- (iii) Beyond a certain critical value Δ_c , the isolated eigenvalue is “eaten up” by \mathcal{I} and disappears [this is called, in a

different context, the Baik-Ben Arous-Péché (BBP) transition [59]].

Furthermore, as soon as Δ is nonzero, the interval \mathcal{I} is further subdivided into three intervals $\mathcal{I}_-, \mathcal{I}_0, \mathcal{I}_+$ (with \mathcal{I}_0 possibly empty; see Fig. 1), where the structure of the corresponding eigenvectors is markedly different. In the central part \mathcal{I}_0 , eigenvectors are *extended*, or *delocalized*, whereas in the extreme parts $\mathcal{I}_-, \mathcal{I}_+$, eigenvectors are *localized*. In a hand-waving manner, “localized” means that most of the norm of the vector is concentrated on a few nodes (firms), whereas “delocalized” means that the norm is well spread out over all nodes. More precisely, calling v_1, v_2, \dots, v_N the component of a normalized vector $|V\rangle$, the localized or delocalized nature of $|V\rangle$ is captured by its Herfindahl index H [called the inverse participation ratio (IPR) in the physics literature]:

$$H(|V\rangle) = \sum_i^N |v_i|^4. \quad (10)$$

A localized eigenvector is such that $H(|V\rangle) = \mathcal{O}(1)$ in the limit $N \rightarrow \infty$ whereas a delocalized eigenvector has $H(|V\rangle) = \mathcal{O}(N^{-1})$. The importance of this distinction for crisis propagation in the context of our model will become clear below.

Owing to the structure of \mathbf{M} , the Perron-Frobenius theorem ensures that its leftmost eigenvalue λ_{\min} is real with a real positive eigenvector $u_i > 0$. As stated above, λ_{\min} must be positive for \mathbf{M} to be an M matrix, i.e., for all prices and all quantities to be positive. As $\lambda_{\min} \rightarrow 0$, the economy becomes more and more fragile to external shocks. Let us for example consider the case where the productivity z_i of some firms decrease by $-\varepsilon \Delta_i < 0$, and/or that some of the stoichiometric weights J_{ij} increase by some amount $\varepsilon \Delta_{ij} > 0$. Using standard perturbation theory to first order in ε ,⁸ one finds that the leftmost eigenvalue is shifted as

$$\lambda_{\min} \longrightarrow \lambda_{\min} - \varepsilon \left[\sum_i \Delta_i u_i^2 + \sum_{i \neq j} \Delta_{ij} u_i u_j \right]. \quad (11)$$

Since both correction terms are negative, this formula shows that as the system becomes marginally stable, any local decrease of productivity or increase of required inputs tips the system towards the unstable region. A certain number of prices or quantities then become negative. Intuitively, the physiognomy of these “crises” will depend on the localized or delocalized nature of the eigenvector $|U\rangle$ corresponding to λ_{\min} , as we now discuss.

The next order correction to Eq. (11) is of order ε^2/g , where g is the gap between λ_{\min} and the next eigenvalue of \mathbf{M} , therefore first order perturbation theory is only valid provided $\varepsilon \ll g$. Now, two cases must be distinguished, depending on the strength Δ of the heterogeneities:

- (i) When $\Delta < \Delta_c$, the leftmost eigenvalue is isolated, in which case $g > 0$ even when $N \rightarrow \infty$. The first order result Eq. (11) is then valid when ε is small enough. Furthermore,

⁸Indeed, given an eigenvector \vec{u} corresponding to an eigenvalue λ of a matrix \mathbf{B} subject to a perturbation $\mathbf{B} \rightarrow \mathbf{B} + \varepsilon \mathbf{P}$, the first order correction to λ in epsilon reads $\lambda \rightarrow \lambda + \varepsilon \vec{u}^\dagger \mathbf{P} \vec{u}$.

the associated eigenvector $|U\rangle$ is delocalized. From Eq. (11), one deduces that a localized shock—say on firm ℓ alone—decreases λ_{\min} by $\sim -\delta z_\ell/N$. The system is unstable when $\delta z_\ell > N\lambda_{\min}$, but for this condition to be compatible with $\delta z_\ell \ll g$, one must also require $N\lambda_{\min} \ll g$. When destabilized, the shock propagates over the whole system, because of the delocalized nature of $|U\rangle$. In the case of a small *global* productivity shock $\delta z_i = \delta z, \forall i$, the destabilization occurs as soon as $\delta z > \lambda_{\min}$.

(ii) When $\Delta > \Delta_c$, the leftmost eigenvalue is at the edge of the interval \mathcal{I}_- , such that the gap $g(N)$ generically goes to zero as $N \rightarrow \infty$. Furthermore, the associated eigenvector $|U\rangle$ is now localized, usually centered around particularly low productivity or high connectivity firms (called the Lifschitz regions in the physics literature [55,60,61]). In this case, however, first order perturbation theory breaks down as soon as $\varepsilon \sim g(N)$, so one must have recourse to numerical simulations to characterize the associated crisis patterns—see the next section and Appendix D for a comparison with empirical data.

Finally, the following remarks should be useful to get an intuition about crisis propagation in our model. Note that one can write prices and outputs using the inverse matrix \mathbf{M}^{-1} and its transpose. Hence, the price response to some generic perturbations δy (for example to productivity, household consumption, etc.) can be expressed using the eigenvalues and eigenvectors as

$$\delta p_i = \sum_{\alpha} \ell_i^{\alpha} \frac{1}{\lambda_{\alpha}} \langle r^{\alpha} | \delta y \rangle, \quad (12)$$

where $\ell^{\alpha}, r^{\alpha}$ are, respectively, the left and right eigenvectors of \mathbf{M} associated to eigenvalue λ_{α} . Similarly, for production

$$\delta y_i = \sum_{\alpha} r_i^{\alpha} \frac{1}{\lambda_{\alpha}} \langle \ell^{\alpha} | \delta y \rangle. \quad (13)$$

In the limit where λ_{\min} touches zero with a finite gap g , one can approximate these responses as

$$\delta p_i \approx \frac{\ell_i^{\min} \langle r^{\min} | \delta y \rangle}{\lambda_{\min}}; \quad \delta y_i \approx \frac{r_i^{\min} \langle \ell^{\min} | \delta y \rangle}{\lambda_{\min}}. \quad (14)$$

Hence, the amplitude of the response of prices depends on the overlap $\langle r^{\min} | \delta y \rangle$ and is localized on the left eigenvector ℓ^{\min} , and vice versa for production. This will be illustrated using real data in Appendix D.

In order to understand intuitively the divergence of the response to perturbations, consider the simple case where $\forall i, z_i = z$. One can expand \mathbf{M}^{-1} in the stable region as

$$\mathbf{M}^{-1} = \frac{1}{z} \sum_{k=0}^{\infty} \left(\frac{\mathbf{J}}{z} \right)^k \quad (15)$$

with $(\mathbf{J})_{ij} = J_{ij}$, since the stability condition implies that the spectral radius of \mathbf{J} is smaller than z . Now, the term $(\mathbf{J}^k)_{ij}$ consists of the sum of all paths of length k linking firm j to firm i . Marginal stability corresponds to this sum becoming divergent, with paths of all lengths contributing to $(\mathbf{M}^{-1})_{ij}$. This interpretation also holds in the case of heterogeneous

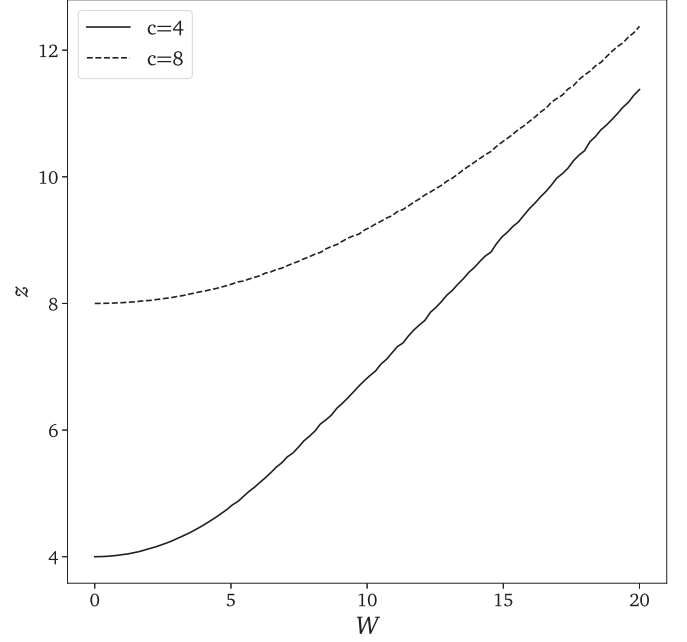


FIG. 2. Plot of the average productivity $z = f_c(W)$ needed to stabilize the economy in the case of a directed RRN with productivities uniformly distributed in $[z - W/2, z + W/2]$, for connectivities $c = 4$ and $c = 8$. Notice that we find $f_c(0) = c$ and a linear behavior of $f_c(W)$ as $W \rightarrow \infty$, as expected. Economies with $N = 2048$ firms were simulated. Error bars are too small to be visible.

z_i 's.⁹ The instability is therefore related to a situation where shocks can propagate over paths of arbitrary length in the input-output network. This is closely related to second order phase transitions in physics, where correlations extend over macroscopic distances and the response to small perturbations diverges; see, e.g., [62].

IV. NUMERICAL RESULTS: BROAD DISTRIBUTION OF FIRM SIZES AND CRISES

We will consider the simplest model of a random regular network, where each firm has exactly c suppliers and c customers, each chosen randomly among the $N - 1$ other firms (other types of networks will be discussed in [36]). Each firm has a random productivity uniformly chosen in the interval $[z - W/2, z + W/2]$, such that z is the average productivity and $\Delta = W/2\sqrt{3}$. The stoichiometric coefficients J_{ij} are taken to be all equal to J . Without loss of generality, J can be set to unity. In addition, we set $V_i = 1$ for simplicity and take the households' consumption $C_i = 1/p_i$ as obtained from a logarithmic utility function with identical preference for all products.

When $W = 0$, the spectrum of \mathbf{M} can be computed, as discussed in the previous section, with an isolated leftmost eigenvalue given by $\lambda_{\min} = z - Jc$. As W increases, the spectrum evolves as shown in Fig. 1. In the case of $c = 7$, the isolated eigenvalue disappears when $W = W^* \approx 15.5$, and the

⁹Indeed one can always write an M matrix as $\mathbf{M} = z_{\max} \mathbf{1} - \mathbf{B}$ where \mathbf{B} is non-negative, and expand the series as $\mathbf{M}^{-1} = \frac{1}{z_{\max}} \sum_{k=0}^{\infty} \left(\frac{\mathbf{B}}{z_{\max}} \right)^k$.

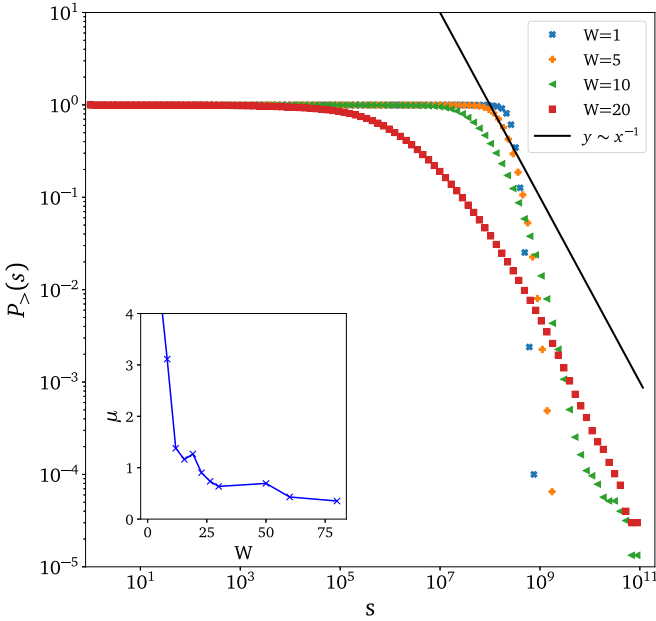


FIG. 3. Log-log cumulative distribution of firm sizes $P_{>}(S)$ defined by sales $S_i = z_i \gamma_i p_i$, along with a curve corresponding to power law $S^{-\mu}$ with exponent $\mu = 1$ (Zipf) for comparison. For small values of W the distribution of firm sizes is sharply peaked at a value of order $1/\varepsilon = 10^8$. Increasing W causes the distribution to get fatter tails with an apparent power-law exponent μ that decreases with W . Here, $c = 4$, $z = f_c(W)$, $N = 1500$. Inset: power-law exponent μ as a function of W , along the critical line of the model.

edge of the spectrum corresponds to a localized eigenvector. In the following, we fix the average productivity to $z = f_c(W)$ such that [the function $f_c(W)$ is shown in Fig. 2]. The model then only depends on two parameters: the connectivity c and the productivity heterogeneity W . We will study along this critical line different characteristics of the corresponding economy.

Two quantities are of particular interest for this paper (a more throughout account of the results will be reported in [36]). One is the distribution of firm size, defined as the total sales $S_i = z_i \gamma_i p_i$. Quite interestingly, while this distribution has thin tails when W is small, it becomes fat-tailed (Zipf-like) as W increases, as observed empirically [63], but with an exponent that appears to vary with W and c (see Fig. 3, inset). The emergence of a power-tailed firm size distribution is a consequence of the criticality of the model, but requires the extreme eigenvectors to be localized and heterogeneous, as it is the case for W sufficiently large.

The second quantity is the distribution of crises amplitudes \mathcal{A} , defined as the total size of the firms that are such that their equilibrium price becomes negative after an idiosyncratic shock of amplitude $-\delta z_\ell$ hitting a certain firm ℓ . (Shocks on the coefficients J_{ij} lead to qualitatively similar results.) While in a fraction of cases nothing much happens, an avalanche can develop where a number of firms “go under,” in the sense that their equilibrium price becomes negative. Conditioned to such events, the distribution of \mathcal{A} is found to be of three types (see Fig. 4):

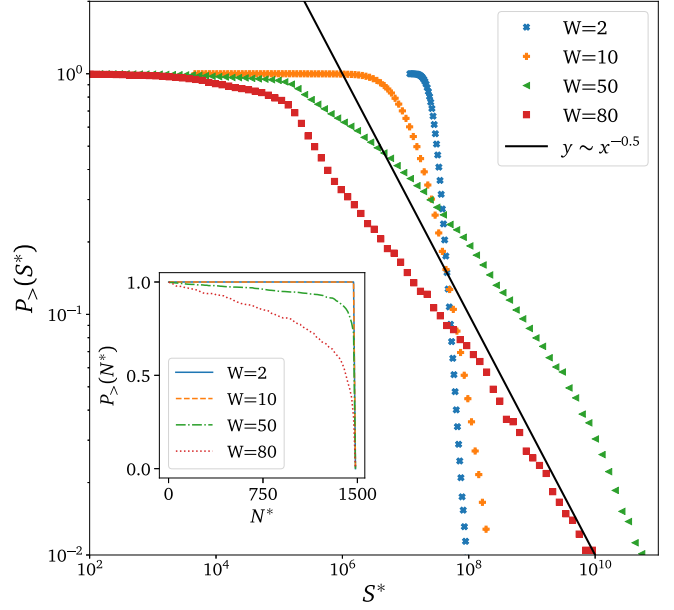


FIG. 4. Log-log cumulative distribution of avalanche sizes as defined by the total sales of firms gone under after a shock, for $c = 4$, $z = f_c(W)$, $N = 1500$, and $\delta z_\ell = 0.05$. One again observes a broad, power-law tailed distribution of casualties for large enough W 's. Inset: cumulative distribution of the number of firms N^* that have gone under after a shock. Notice that $W = 2$ and $W = 10$ correspond to mostly systemwide avalanches (i.e., $N^* \approx 1500$), while larger values of W correspond to avalanches of all sizes.

(1) Mostly “systemwide,” where a substantial fraction of the output is wiped out. This occurs when $W < W^*$ and $\delta z_\ell > N \lambda_{\min}$, as expected from our general discussion;

(2) Thin-tailed, where avalanches are restricted to particularly fragile firms connected to ℓ . This corresponds to $W > W^*$, and weak perturbations $\delta z_\ell < g$, in which case only one or a few localized eigenvectors close to the edge propagate the crisis.

(3) Fat-tailed, where small crises coexist with large crises (a feature of self-organized criticality, as recalled in the Introduction). This happens when $\delta z_\ell \gg g$, i.e., when a large collection of eigenstates are mobilized in the crisis propagation.

The generic existence of three crisis scenarios is, we believe, quite interesting. In particular, the possibility that a small, idiosyncratic shock can lead to systemwide trouble, or else to avalanches of all sizes, has potentially deep consequences on our understanding of the business cycle and on crisis prevention policies. Of course, the above analysis postulates that the economy is close to criticality, i.e., that $\lambda_{\min} \rightarrow 0$. Why this should be the case is obviously the crux of the matter, and will be discussed in the next section.

It is also interesting to plot the total consumption of households as a function of λ_{\min} in the two cases above: delocalized ($W < W^*$) vs localized ($W > W^*$) crises. In the first case, the whole economy grinds to a halt as $\lambda_{\min} \rightarrow 0$, as expected. In the second case, only a fraction of the total consumption (mostly coming from firms represented in the corresponding localized eigenvector) is affected. See Fig. 5 for an illustration of this point.

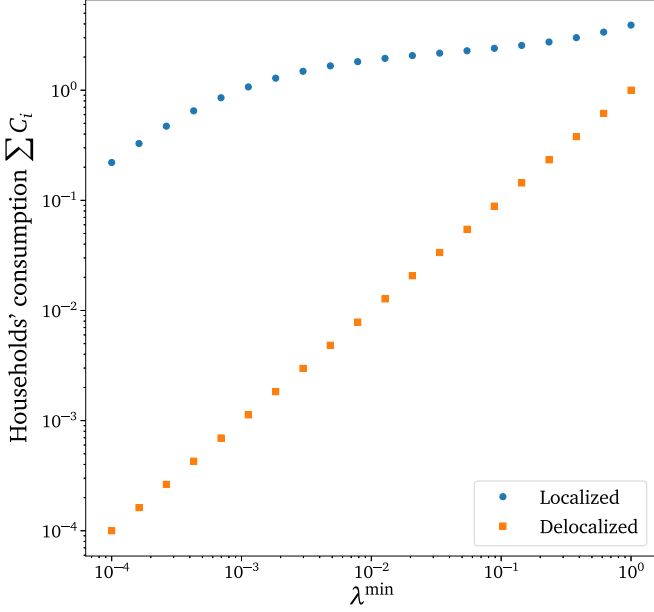


FIG. 5. Total consumption of households vs λ_{\min} in the localized and delocalized cases. Intuitively, one expects the prices of goods represented in the eigenvector corresponding to λ_{\min} to behave as λ_{\min}^{-1} , while a logarithmic utility function implies that consumption scales as the inverse of the price, leading naturally to the consumption of those goods to be proportional to λ_{\min} . In the delocalized case, all goods are concerned and thus global consumption plummets, as seen in the graph (orange squares). On the other hand, only the handful of goods associated to λ_{\min} see their consumption decline in the delocalized case, corresponding to the blue dots. An economy with 1000 firms and a connectivity $c = 4$ was used for this plot, with $W = 0$ for the localized case and $W = 12$ for the delocalized case.

Our general scenario is strongly reminiscent of similar ideas in an ecological context, where the disappearance of a single species can lead to mass extinctions mediated by network effects [31,34]. A major difference, however, is that the economic network is not static and can in principle adapt to new conditions on relatively short time scales (at least compared to evolutionary time scales). With our framework, we expect that if a supplier undergoes some difficulty (i.e., its equilibrium production is found to be negative), its customers may choose to rewire and look for alternatives. Furthermore, one expects that the market clearing and zero profit conditions will be temporarily violated. An extension of the present model that takes such dynamical effects into account would certainly be extremely interesting (see [25,41] for preliminary work in that direction). But what is clear is that if rewiring takes time and/or is costly, the “paper crises” found above could indeed materialize as actual defaults, or at least acute difficulties. Since economic frictions are substantial and rewiring cannot be instantaneous, we expect that the present scenario could be relevant to understanding real world crises [38–40].

Let us finally come back to the case of partial substitutability, i.e., when the parameter q appearing in Eq. (1) is strictly larger than zero. Since the matrix $(\hat{\mathbf{M}})_{ij} = z_i^\zeta \delta_{ij} - a_{ij}^{q\zeta} J_{ij}^\zeta$ is a continuous function of q , it is clear that if the smallest eigenvalue λ_{\min} is negative for $q = 0$, it will remain so for a certain

range of q . We have checked numerically on some examples that this is indeed the case; the economy is only stabilized when q exceeds a (problem dependent) value $q_c > 0$. Not surprisingly, allowing for more substitutability can stabilize an otherwise unfeasible economy. An in-depth study of this new threshold will be presented in our forthcoming work [36], but we expect all the properties reported in this section to hold true for all $q > 0$ when the system is close to criticality.

V. MARGINAL STABILITY: DISCUSSION AND CONCLUSION

In this section, we will motivate our claim that generic economies—like many other complex systems, see, e.g., [30,34,64–66]—might “self-organize” to sit, at least temporarily, close to the boundary of the stable region, i.e., satisfy the marginal stability criterion $\lambda_{\min} \rightarrow 0$. Several types of evolutionary forces act to that effect. One is simply the creation of new firms, that lead to an *effective* reduction of productivity and increase of connectivity. To see this, consider that the economy consists of N firms in equilibrium and add an additional firm indexed by \star , with productivity z_\star , labor requirements $J_{\star 0} = V_\star/p_0$, and links $J_{\star i}, J_{j\star}$ to the N pre-existing firms. The equilibrium condition for price p_\star is

$$p_\star = \frac{V_\star}{z_\star} + \sum_{j=1}^N \frac{J_{\star j}}{z_\star} p_j. \quad (16)$$

Plugging this result in the new equilibrium conditions for the N original firms yields

$$\left(z_i - \frac{J_{i\star} J_{\star i}}{z_\star}\right) p_i - \sum_{j=1}^N \left(J_{ij} + \frac{J_{i\star} J_{\star j}}{z_\star}\right) p_j = V_i + \frac{J_{i\star} V_\star}{z_\star}, \quad (17)$$

which means that the addition of a firm amounts in effect to *decreasing* all original productivities: $z_i \rightarrow z_i - \frac{J_{i\star} J_{\star i}}{z_\star}$ and *increasing* all stoichiometric coefficients: $J_{ij} \rightarrow J_{ij} + \frac{J_{i\star} J_{\star j}}{z_\star}$. As clear from Eq. (11), this can only decrease the smallest eigenvalue of the matrix \mathbf{M}_N^\star that describes the pre-existing firms with the new firm added. One concludes that a growing economy can only become more unstable with time. This argument is actually closely related to May’s original argument about the stability of large ecologies [31].

In fact, one can show that as the number of links to the most connected node of the network increases, the smallest eigenvalue of \mathbf{M} decreases [58], until the instability threshold is reached. In this case, the fragility of the network comes from the most central hubs, a scenario akin to, but different from, the one of Acemoglu *et al.* [15]. This effect might be amplified if firms systematically favor links toward hubs (as suggested in [40]), leading to a “scale-free” input-output network [67]. Interestingly, a stability-constrained growth mechanism for networks, whereby a node is freely added to the network if it does not destabilize the system but induces rewirings in the network until stability is found again if it does, has been found to generate such scale-free networks [68].

The second evolutionary effect is, even for a fixed size N , the complexification of the production process, i.e., technology progress means that a wider array of products are needed as inputs. If the average productivity z remains the

same while the average connectivity c increases, the system eventually reaches the instability point (which in the simplest case reads $z = Jc$). Hence productivity must increase at some minimum rate for the economy to remain stable. But since increasing productivity is costly, one can postulate that the average productivity z will tend to hover around the minimal viable threshold, and sometimes lagging behind, leading to occasional endogenous crises. Similarly, as mentioned after Eq. (2), firms tend to optimize their portfolios of suppliers, thereby reducing their redundancy but, by the same token, reducing the effective substitution effects captured by the CES parameter q . As $q \rightarrow q_c^+$, the economy will again become unstable.

Finally, we have assumed that firms are perfectly competitive and that equilibrium corresponds to zero profit. Now, in more realistic situations, firms attempt to realize positive profits and distribute dividends. As already noted, if the profit target of firm i is a certain fraction φ_i of its total sales $S_i = z_i \gamma_i p_i$, Eq. (8) remains identical but with a decreased effective productivity $z_i \rightarrow z_i(1 - \varphi_i)$. As firms attempt to maximize their profits, the average effective productivity goes down, until the marginal stability point is reached and a crisis ensues. After the crisis, economic actors revert to more reasonable levels of markups (i.e., reduce φ_i), which makes the economy viable again—until the next crisis.

One could probably come up with other mechanisms that push the economy towards instability, see for example [69,70]. Our conjecture is that evolutionary and behavioral forces repeatedly drive the economy close to marginal stability. As anticipated by Bak, Chen, Scheinkman, and Woodford [28–30] and confirmed in this paper, this scenario would be a natural explanation of the broad (Zipf-like) distribution of firm sizes, and of the “small shocks large business cycle” puzzle, that both suggest some kind of criticality. Crises should then be understood as intrinsically nonlinear events, where feedback loops of arbitrary size contribute to propagating and amplifying idiosyncratic shocks.

There are many directions to explore further. The most important one is, in our opinion, to endow the model with some realistic dynamics, partly along the lines of [25], that would include frictions, myopia, imperfect market clearing, rewiring, etc. This would make the model more realistic, and is a prerequisite to calibration on empirical data, since within the present static setting crises are signaled by the appearance of negative prices, beyond which the model ceases to make sense.

ACKNOWLEDGMENTS

We are indebted to F. Benaych-Georges, G. Biroli, J. Bonart, G. Bunin, C. Colon, R. Farmer, A. Kirman, A. Landier, A. Mandel, M. Marsili, J. P. Nadal, F. Roy, A. de Sanctis, A. Secchi, D. Sharma, M. Tarzia, D. Thesmar, and F. Zamponi for countless illuminating discussions on the topics of this paper. We thank in particular X. Gabaix for many detailed comments, and for insisting that we should investigate the model beyond its Leontief limit.

APPENDIX A: CES PRODUCTION FUNCTIONS

Of standard usage in economics, the constant elasticity of substitution (CES) functions are a family of production

functions giving the total production of a firm i given inputs Q_{ij} from its suppliers.¹⁰ In the most general setting, the CES production function is defined as

$$\pi_i = z_i \left[\sum_j a_{ij} \left(\frac{Q_{ij}}{J_{ij}} \right)^{-1/q} \right]^{bq}, \quad (\text{A1})$$

where z_i is the productivity level of firm i , the a_{ij} s are weight coefficients satisfying $\sum_j a_{ij} = 1$, and the J_{ij} terms are stoichiometric coefficients defining the number of inputs from j required to make a unit of i 's output. In the remaining terms, b is called the returns to scale: multiplying all inputs Q_{ij} by some coefficient K will make the whole output level to be multiplied by K^b . In the main body and all that follows we have chosen $b = 1$, corresponding to the so-called constant returns to scale case, but our analysis can be extended to other values of b . The effect of q , the degree of substitutability, deserves a more in-depth discussion through the study of the limits $q \rightarrow \infty$ and $q \rightarrow 0$, corresponding to the so-called Cobb-Douglas and Leontief production functions.

1. Perfect substitutability: Cobb-Douglas case

Take indeed the limit $q \rightarrow \infty$ as

$$\begin{aligned} \pi_i/z_i &= \exp \left[-q \ln \left(\sum_j a_{ij} \exp \left[-\frac{1}{q} \ln \left(\frac{Q_{ij}}{J_{ij}} \right) \right] \right) \right] \\ &\simeq \exp \left[-q \ln \left(1 - \frac{1}{q} \sum_j a_{ij} \ln \left(\frac{Q_{ij}}{J_{ij}} \right) \right) \right] \\ &\simeq \exp \left(\sum_j a_{ij} \ln \left(\frac{Q_{ij}}{J_{ij}} \right) \right) = \prod_j \left(\frac{Q_{ij}}{J_{ij}} \right)^{a_{ij}} \end{aligned} \quad (\text{A2})$$

and one retrieves the ubiquitous Cobb-Douglas production function. In this setting, one can easily check that if a given input Q_{ij} from a firm j is multiplied by an amount $f < 1$ then the output need not drop if any other input Q_{ik} is multiplied by $f^{-a_{ik}/a_{ij}}$.

2. Unsubstitutable inputs: Leontief case

Take instead the limit $q \rightarrow 0$, and consider $j^* = \arg \min_j (Q_{ij}/J_{ij})$ to get

$$\begin{aligned} \pi_i/z_i &= \left(a_{ij^*} \left(\frac{Q_{ij^*}}{J_{ij^*}} \right)^{-1/q} + \sum_{j \neq j^*} a_{ij} \left(\frac{Q_{ij}}{J_{ij}} \right)^{-1/q} \right)^{-q} \\ &= \frac{Q_{ij^*}^{qb}}{J_{ij^*}^{qb}} \left(a_{ij^*} + \sum_{j \neq j^*} a_{ij} \left(\frac{J_{ij^*} Q_{ij}}{Q_{ij^*} J_{ij}} \right)^{-1/q} \right)^{-q} \\ &\xrightarrow{q \rightarrow 0} \frac{Q_{ij^*}}{J_{ij^*}} = \min_j \left(\frac{Q_{ij}}{J_{ij}} \right), \end{aligned} \quad (\text{A3})$$

¹⁰This may also include labor inputs, which we conventionally choose to correspond to the index $j = 0$.

where the total output of firm i is determined by its scarcest input.

The CES production function therefore bridges these two limiting cases of which only the Cobb-Douglas case has been studied in the network literature. We will now study the competitive equilibrium equations for values of $q \in [0; \infty)$.

APPENDIX B: COMPETITIVE EQUILIBRIUM EQUATIONS AND HAWKINS-SIMON CONDITION

Our problem is first to maximize the profits,

$$\mathcal{P}_i = \pi_i p_i - \sum_j Q_{ij} p_j, \quad (\text{B1})$$

for each firm subject to the constraint given by Eq. (A1). Computing first the derivative of π_i with respect to $Q_{il=j}$ and substituting using Eq. (A1) yields

$$\frac{\partial \pi_i}{\partial Q_{ij}} = z_i^{-1/q} a_{ij} J_{ij}^{1/q} Q_{ij}^{-(1+q)/q} \pi_i^{(1+q)/q}, \quad (\text{B2})$$

which can now be used to set $\frac{\partial \mathcal{P}_i}{\partial Q_{ij}} = 0$, i.e.,

$$p_i z_i^{-1/q} a_{ij} J_{ij}^{1/q} Q_{ij}^{-(1+q)/q} \pi_i^{(1+q)/q} = p_j$$

$$Q_{ij} = \left(\frac{p_i}{p_j} a_{ij} \right)^{q/(1+q)} J_{ij}^{1/(q+1)} z_i^{-1/(1+q)} \pi_i. \quad (\text{B3})$$

One can now check this solution in the Leontief and Cobb-Douglas limiting cases,

$$Q_{ij} \begin{cases} \xrightarrow{q \rightarrow \infty} \frac{p_i}{p_j} a_{ij} \pi_i \\ \xrightarrow{q \rightarrow 0} J_{ij} \frac{\pi_i}{z_i} \end{cases}, \quad (\text{B4})$$

retrieving the condition $Q_{ij} = J_{ij} \gamma_i$ defined in Eq. (4) of the main body, where $\gamma_i := \pi_i / z_i$ is the firm's output level.

In the Leontief case the optimal input is necessarily determined by the firm's desired output level, while it is determined by the input's price in the Cobb-Douglas case.

One needs now to impose a competitive equilibrium by setting all optimized profits to 0 as

$$\pi_i p_i = \sum_j Q_{ij} p_j$$

$$\pi_i (z_i^{1/(1+q)} p_i^{1/(1+q)}) = \pi_i \left(\sum_j a_{ij}^{q/(1+q)} J_{ij}^{1/(1+q)} p_j^{1/(1+q)} \right) \quad (\text{B5})$$

corresponding to Eq. (9). The existence of a solution $p_i > 0$ for this equation is equivalent to saying that the matrix $(\tilde{\mathbf{M}})_{ij} = z_i^\zeta \delta_{ij} - a_{ij}^{q\zeta} J_{ij}^\zeta$, with $\zeta = 1/(1+q)$, is a so-called M matrix. Once prices are determined, imposing market clearing $\pi_i = \sum_j Q_{ji}$ also leads to a similar equation,

$$z_i \gamma_i - \sum_j \left(\frac{p_j}{p_i} a_{ji} \right)^{q\zeta} J_{ji}^{\zeta} z_j^{q\zeta} \gamma_j = Q_{j0}, \quad (\text{B6})$$

which would require the matrix $(\tilde{\mathbf{M}})_{ij} = z_i \delta_{ij} - (p_i/p_j)^{q\zeta} J_{ji}^{\zeta} z_j^{q\zeta}$ to be an M matrix. In the Leontief $q \rightarrow 0$ case

we have the relation $\tilde{\mathbf{M}} = \hat{\mathbf{M}}^\top$ and the sufficient condition to have both a competitive zero profit equilibrium and market clearing is that $\hat{\mathbf{M}}$ be an M matrix.

On the other hand, in the Cobb-Douglas case, it is easy to see that $(\hat{\mathbf{M}})_{ij} = \delta_{ij} - a_{ij}$ is always an M matrix, and thus the prices are chosen so that $(\tilde{\mathbf{M}})_{ij} = z_i \delta_{ij} - \frac{p_j}{p_i} a_{ji} z_j$ is an M matrix to have a market-clearing equilibrium.

APPENDIX C: FUNCTIONAL ECONOMIES

In [26,27] the authors proved that a necessary and sufficient condition for an equation such as Eq. (8) to have solutions is that all of the principal minors of the matrix \mathbf{M} be positive, while the authors in [42] proved that this condition is equivalent to all of the eigenvalues of \mathbf{M} having a positive real part.

Regarding the positivity of all principal minors of \mathbf{M} , Hawkins and Simon gave an economic interpretation of this condition by claiming that it is equivalent to saying that “the group of industries corresponding to each minor must be capable of supplying more than its own needs for the group of products produced by this group of industries” [27]. In this section we will provide a rewording of this in terms of an effective medium equation and Schur complements.

1. Effective medium for one firm

Consider first a firm i satisfying Eq. (7) written as

$$\mathbf{M} |P\rangle = |V\rangle \quad (\text{C1})$$

and consider the matrix $\mathbf{M}^{(i)}$ with row and column i removed, as well as $|P^{(i)}\rangle$ and $|V^{(i)}\rangle$ the vectors with i removed, $|J_{i\leftarrow}\rangle = (J_{i0}, \dots, J_{iN})$ and $|J_{i\rightarrow}\rangle = (J_{0i}, \dots, J_{Ni})$. The previous equation can now be written as

$$z_i p_i - \langle J_{i\leftarrow} | P^{(i)} \rangle = V_i,$$

$$\mathbf{M}^{(i)} |P^{(i)}\rangle - p_i |J_{i\rightarrow}\rangle = |V^{(i)}\rangle; \quad (\text{C2})$$

multiplying now the second line by $(\mathbf{M}^{(i)})^{-1}$, taking the product with $\langle J_{i\leftarrow}|$, and subtracting from the first line leads to

$$[z_i - \langle J_{i\leftarrow} | (\mathbf{M}^{(i)})^{-1} | J_{i\rightarrow} \rangle] p_i = V_i + \langle J_{i\leftarrow} | (\mathbf{M}^{(i)})^{-1} | V^{(i)} \rangle, \quad (\text{C3})$$

which can be interpreted as i being a unique isolated firm, albeit with an effective “renormalized” productivity $\tilde{z}_i = z_i - \langle J_{i\leftarrow} | (\mathbf{M}^{(i)})^{-1} | J_{i\rightarrow} \rangle < z_i$. The Hawkins-Simons condition implies that $\mathbf{M}^{(i)}$ must also fulfill the same conditions as \mathbf{M} , in particular $(\mathbf{M}^{(i)})^{-1}$ has positive components. For the whole economy to be functional, the effective productivity of each firm must be positive.

2. Sectoral interpretation

One can also extend this analysis to different economical sectors. Consider for simplicity two sectors, so that \mathbf{M} has the following block structure:

$$\mathbf{M} := \begin{pmatrix} \mathbf{M}_1 & -\mathbf{J}_{12} \\ -\mathbf{J}_{21} & \mathbf{M}_2 \end{pmatrix}, \quad (\text{C4})$$

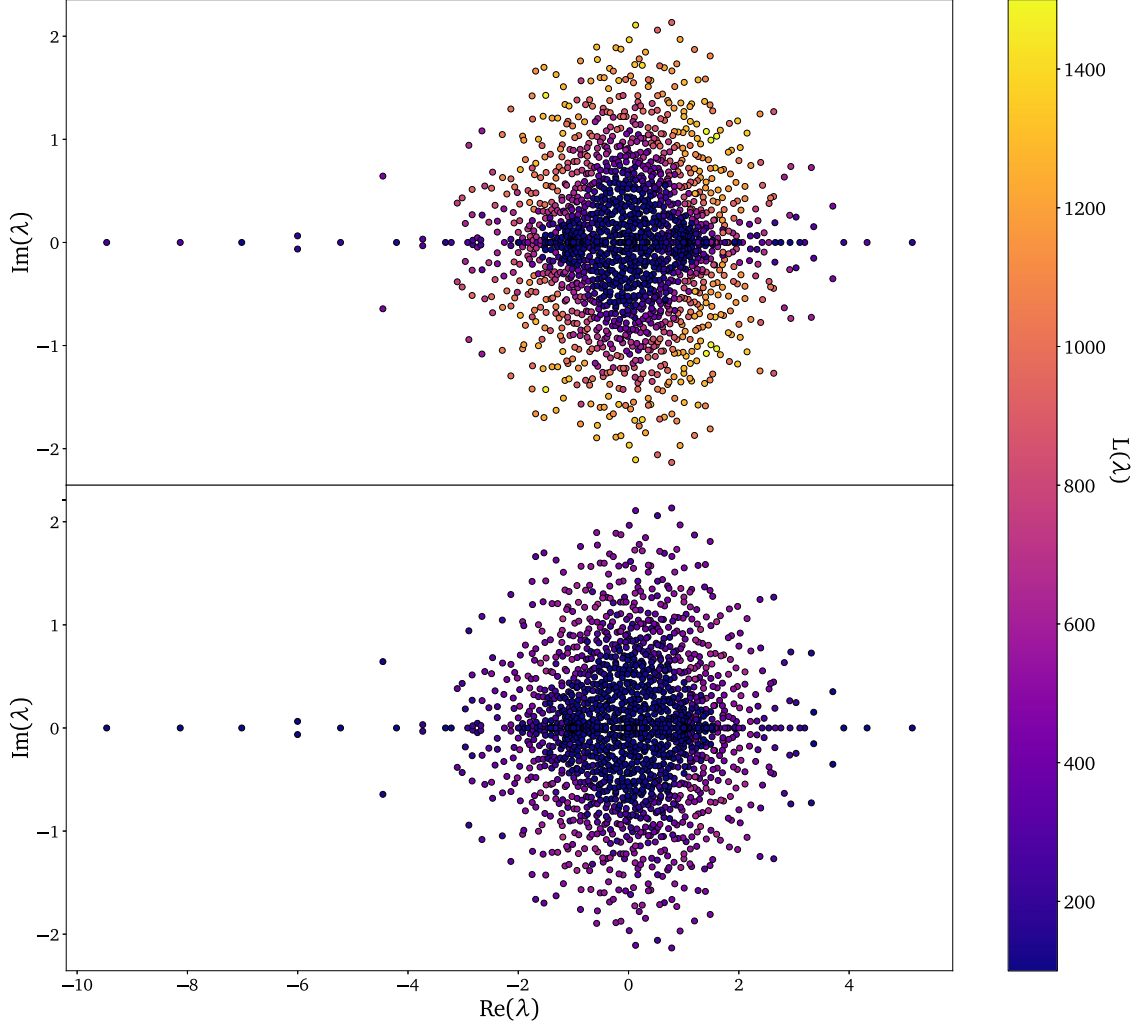


FIG. 6. Spectrum of $-A$ in the complex plane, with the color of each eigenvalue λ given by $L(\lambda)$, defined as the inverse of the inverse participation ratio H of the corresponding eigenvector, e.g., the number of firms over which the eigenvector is effectively spread. The color on the top plot corresponds to the right eigenvectors, while the one on the bottom corresponds to the left eigenvectors. Notice the different localization profiles on left and right eigenvectors.

where the matrix \mathbf{M}_1 is made of firms from sector 1 and inter-linkages between them, while the \mathbf{J} matrices link sector 1 with sector 2. Calling $|P\rangle = (|P_1\rangle, |P_2\rangle)$ and $|V\rangle = (|V_1\rangle, |V_2\rangle)$, we can write as in Eq. (C2) that

$$(\mathbf{M}_1 - \mathbf{J}_{12}\mathbf{M}_2^{-1}\mathbf{J}_{21})|P_1\rangle = |V_1\rangle + \mathbf{J}_{21}\mathbf{M}_2^{-1}|V_2\rangle \quad (\text{C5})$$

and thus the Hawkins-Simon condition is fulfilled also if both \mathbf{M}_2 and $\mathbf{M}_1 - \mathbf{J}_{12}\mathbf{M}_2^{-1}\mathbf{J}_{21}$ are M matrices for any choice of sector partitioning in the economy. In other words, it is not just that \mathbf{M}_1 must be an M matrix by producing enough goods for the consumption of firms in sector 1, but it must also produce enough goods for all of the other sectors.

APPENDIX D: REAL INPUT-OUTPUT NETWORKS

In this section we will attempt to confront our model with available, but partial, data. At this stage, this is more of an exercise that gives color to the general framework presented in the body of the paper.

Since our stability criterion depends on spectral properties of a matrix describing the entirety of the economy, we need in principle highly detailed network data to present a full analysis. Most available data, however, consist of input-output tables describing sale and purchase relationships between entities, be they firms or larger entities such as sectors or even countries, or simple relational data describing who is in a client-supplier relationship with whom, with the latter having a significantly larger coverage. While it allows us to some degree to infer the importance of certain firms in the network, it does not correspond directly to the J_{ij} coefficients that appear in our formalism.

The total productivity factors z_i , which measure the efficiency with which a firm turns inputs into outputs, are also hard to deduce from actual production data. In particular, detailed output data are seldom available, and so productivity measures must be constructed from revenue data only, allowing for potential errors between the actual and the inferred productivity levels. For a detailed discussion, see [71].

TABLE I. Firms with the 20 most important contributions to the eigenvector $|r^{\lambda_{\min}}\rangle$, their respective contributions to the Herfindahl, and their reported sales for the year 2015 in thousands of 2015 USD.

Company name	Sales (kUSD)	Herfindahl contribution $ r_i^{\lambda_{\min}} ^4$
HP Inc.	1.40×10^6	3.62×10^{-3}
Samsung Electronics Co. Ltd.	1.37×10^6	3.21×10^{-3}
Airbus SE	6.69×10^5	1.10×10^{-3}
Boeing Co., The	1.12×10^6	9.37×10^{-4}
General Electric Co.	1.85×10^6	9.28×10^{-4}
Intel Corp.	4.99×10^5	8.85×10^{-4}
Microsoft Corp.	5.61×10^5	8.74×10^{-4}
Apple Inc.	2.20×10^6	5.75×10^{-4}
International Business Machines Corp.	7.93×10^5	4.53×10^{-4}
QUALCOMM Inc.	1.57×10^5	4.23×10^{-4}
Google Inc.	7.51×10^5	2.18×10^{-4}
Sony Corp.	2.73×10^5	2.17×10^{-4}
United Technologies Corp.	7.68×10^5	1.92×10^{-4}
Verizon Communications Inc.	7.73×10^5	1.67×10^{-4}
Siemens AG	6.43×10^5	1.49×10^{-4}
Honeywell International Inc.	3.58×10^5	1.11×10^{-4}
Safran SA	1.02×10^5	1.07×10^{-4}
Taiwan Semiconductor Manufacturing Co. Ltd.	2.25×10^5	1.04×10^{-4}
ARM Holdings PLC	7.63×10^3	5.83×10^{-5}
LG Electronics Inc.	3.45×10^5	5.80×10^{-5}

Detailed data are becoming increasingly available, and we intend to show that spectral analysis of production networks is in principle possible.

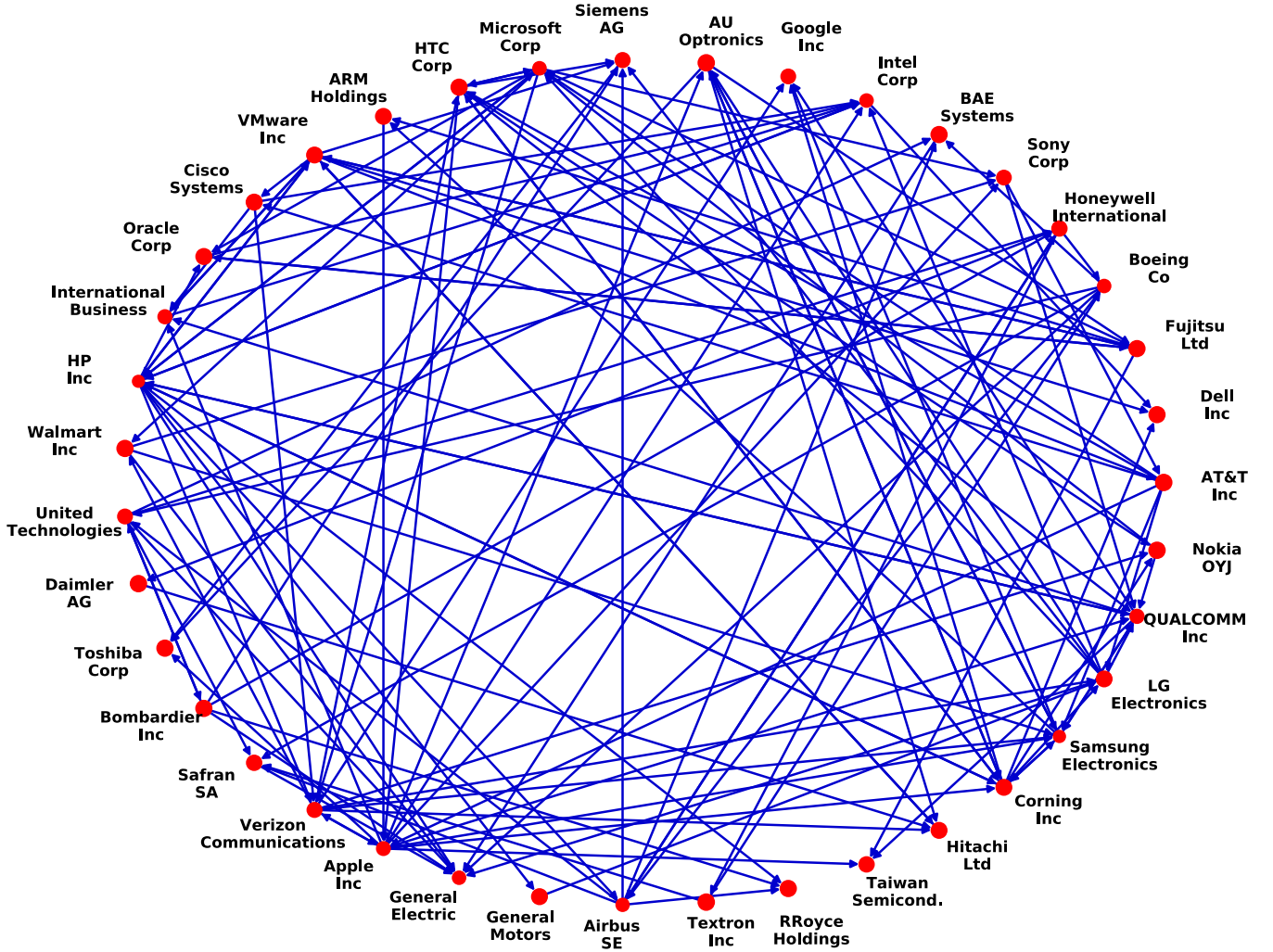
1. Dataset and definitions

We use the FactSet Supply Chain Relationships database to build a supply chain network. The FactSet dataset contains a list of relational data between firms, stating that firms A and B

have a client-supplier relation, if they are in competition or if they have a joint venture. It is built by collecting information from primary public sources such as SEC 10-K annual filings, investor presentations, and press releases, and covers about 23 000 publicly traded companies with over 325 000 relationships. Since the relationships are inferred from data released to the public, we cannot be sure that it is an exhaustive database of all the relationships between firms, but the subset of relationships is deemed important by the firm themselves.

TABLE II. Firms with the 20 most important contributions to the eigenvector $|l^{\lambda_{\min}}\rangle$, their respective contributions to the Herfindahl, and their reported sales for the year 2015 in thousands of 2015 USD.

Company name	Sales (kUSD)	Herfindahl contribution $ l_i^{\lambda_{\min}} ^4$
Avnet Inc.	2.52×10^5	8.94×10^{-4}
Apple Inc.	2.20×10^6	6.61×10^{-4}
Esprinet SpA	1.59×10^4	2.72×10^{-4}
Arrow Electronics Inc.	1.59×10^5	2.39×10^{-4}
3M Co.	2.50×10^5	1.11×10^{-4}
Computacenter PLC	2.74×10^4	8.06×10^{-5}
Data#3 Ltd.	1.25×10^3	7.86×10^{-5}
CDW Corp., DE	1.11×10^5	6.33×10^{-5}
Emulex Corp.	1.26×10^3	6.31×10^{-5}
Stratasys Ltd.	?	6.30×10^{-5}
Xilinx Inc.	3.04×10^4	5.45×10^{-5}
JB Hi-Fi Ltd.	?	5.19×10^{-5}
Entropic Communications Inc.	5.68×10^2	3.64×10^{-5}
Proximus SADP	6.49×10^4	3.46×10^{-5}
Nuance Communications Inc.	2.89×10^4	2.97×10^{-5}
Premier Farnell Ltd.	5.70×10^3	2.92×10^{-5}
Ingram Micro Inc.	3.24×10^5	2.86×10^{-5}
Systex Corp.	6.62×10^2	2.83×10^{-5}
Northrop Grumman Corp.	2.16×10^5	2.73×10^{-5}
CANCOM SE	8.35×10^3	2.67×10^{-5}

FIG. 7. Subgraph of the 38 firms with the largest contributions to $|r^{\min})$.

Such links between firms have a finite duration in time and have thus a beginning and an end date. For our study, we have chosen the set of client-supplier relationships during the whole year of 2015. This allows us to build a graph \mathcal{G} where a link $i \rightarrow j$ exists whenever i is reported to be a supplier of j or when j is reported to be a client of i . This graph \mathcal{G} consists of 237 weakly linked subgraphs,¹¹ of which we select the largest subgraph \mathcal{G}_0 . The graph \mathcal{G}_0 consists of 10 447 firms with 40 300 relationships between them. The remaining subgraphs are very small and consist of at most a few tens of firms.

From \mathcal{G}_0 we construct an adjacency matrix \mathbf{A} defined by

$$(\mathbf{A})_{ij} = \begin{cases} 1 & \text{if } j \rightarrow i \in \mathcal{G}_0, \\ 0 & \text{otherwise} \end{cases}, \quad (\text{D1})$$

which carries the topology we expect from the matrix $(\mathbf{M})_{ij} = z_i \delta_{ij} - J_{ij}$ defined in the main text. Indeed if for simplicity one supposes that all firms have the same productivity z and that

$J_{ij} = 1$ for all links in the graph, then one has $\mathbf{M} = z\mathbf{1} - \mathbf{A}$. A full study of the properties of \mathbf{M} would therefore require the supplementary knowledge of the values taken by the J_{ij} and z_i terms, i.e., have access to the dollar amount of products exchanged between firms and the total production of each firm.

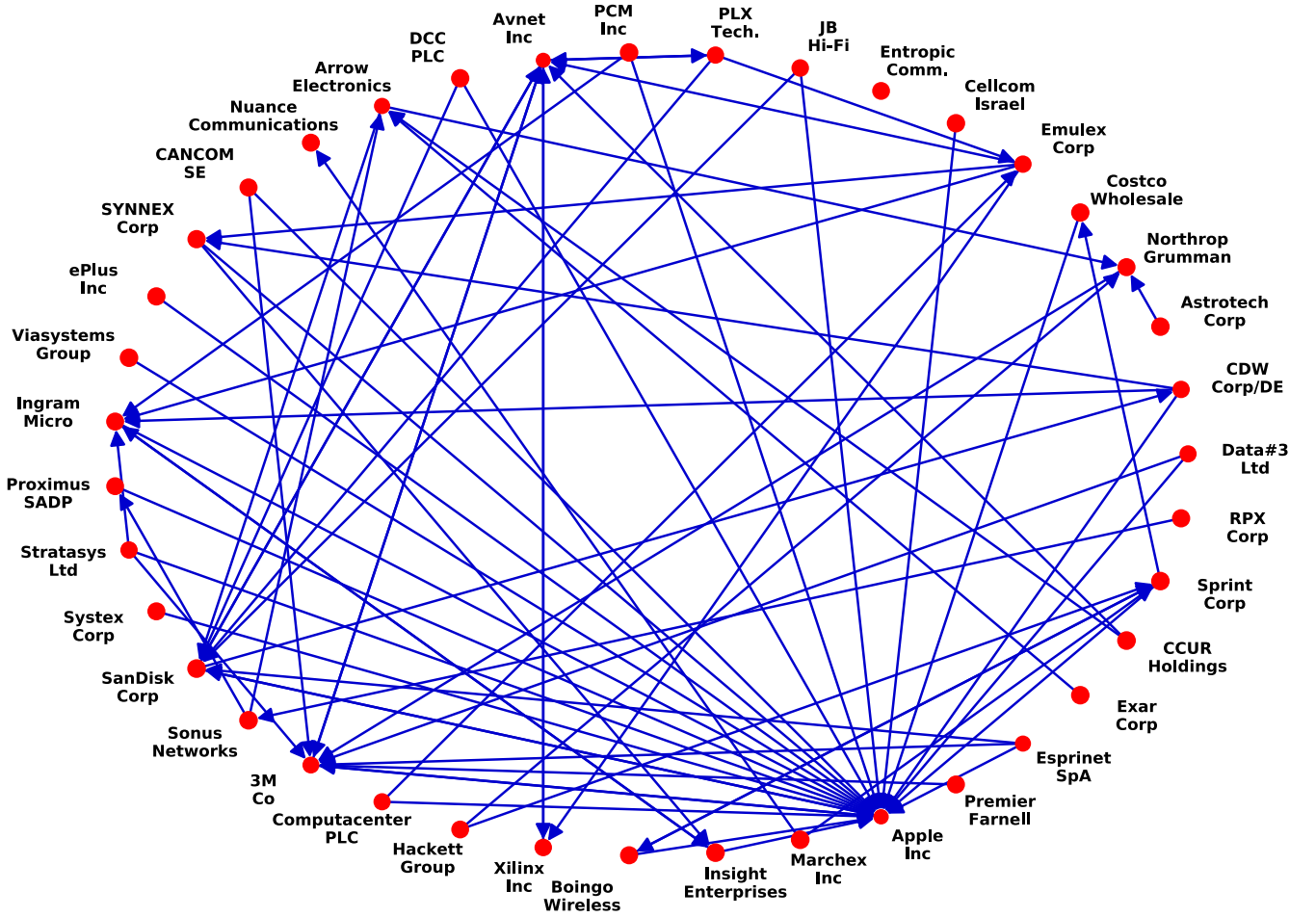
2. Spectral study of the adjacency matrix

As the main body of our paper suggests, it is interesting to look into the properties of the eigenvalues and eigenvectors of the matrix \mathbf{M} , which under the hypothesis presented above are the same as those of the matrix $-\mathbf{A}$, up to a shift by z on the real axis for the eigenvalues.

A notion introduced in our paper is that of localized and delocalized eigenvectors. To quantify the localization properties we introduce the inverse participation ratio, or IPR,¹² H of a normalized complex-valued eigenvector $|v^\lambda\rangle = (v_1^\lambda, \dots, v_N^\lambda)$

¹¹A weakly linked subgraph is such that any two nodes in it can be linked by a directed path.

¹²Akin to the Herfindahl index used in the economics literature.

FIG. 8. Subgraph of the 38 firms with the largest contributions to $|l^{\min}\rangle$.

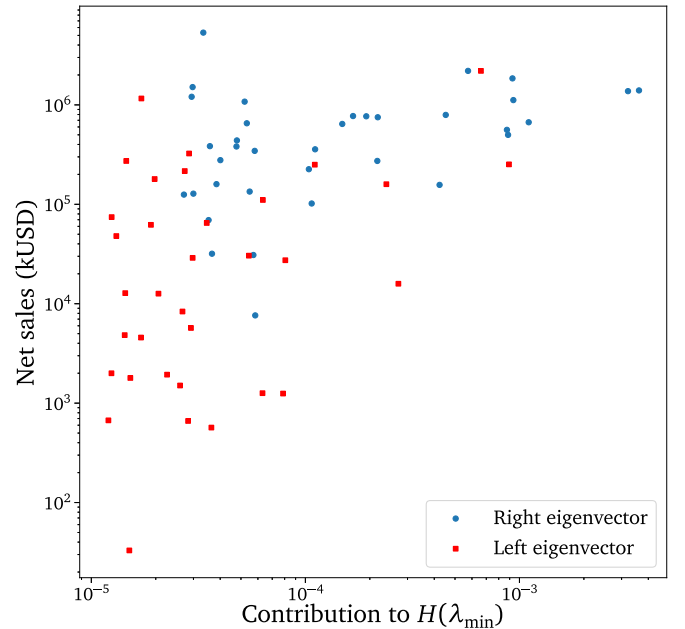
associated to an eigenvalue λ as

$$H(\lambda) := \sum_i ||v_i^\lambda||^4. \quad (\text{D2})$$

Indeed, for a perfectly delocalized eigenvector we should have $\forall i, v_i^\lambda = 1/\sqrt{N}$ and therefore an $H(\lambda) = 1/N$, while an eigenvector localized on a single site $|v^\lambda\rangle = (1, 0, \dots, 0)$ has trivially a value $H(\lambda) = 1$. It follows then that $L(\lambda) = 1/H(\lambda)$ is a measure of the number of sites over which a given eigenvector associated to the eigenvalue λ is spread out.

We have computed the eigenvalues and eigenvectors of $-\mathbf{A}$, as visible in Fig. 6, where one can see qualitatively that the spectral properties of the real adjacency matrix are not far from the simplistic random regular graph case shown in Fig. 1: we see indeed a bulk to the right hand side with isolated eigenvectors to the left of it, and with varying localization properties of the eigenvectors across the spectrum.

Of capital interest to our study are the leftmost eigenvectors $|l^{\min}\rangle$ and $|r^{\min}\rangle$, which owing to the Perron-Frobenius theorem have real positive components and are associated to a real eigenvalue λ_{\min} , as can also be seen in Fig. 6. In the graph we are studying, we have found that the eigenvectors are spread out over $L_r(\lambda_{\min}) \simeq 40$ firms for the right eigenvector and $L_l(\lambda_{\min}) \simeq 185$ for the left eigenvector. For both eigenvectors, we have listed the 20 firms with the most important

FIG. 9. Log-log plot of the sales of the 38 main firms contributing to $|v^{\lambda_{\min}}\rangle$ vs their contribution to $H(\lambda_{\min})$, suggesting that firms with a higher overlap with $|v^{\lambda_{\min}}\rangle$ have higher sales.

contributions on Tables I and II, as well as their yearly reported sales for the year 2015.¹³ We have also represented the subgraph of the 38 firms with the largest contributions to these eigenvectors and their interlinkages in Figs. 7 and 8. This allows for a better understanding of Eq. (14): if a shock hits firms represented in Fig. 7 this will be reflected in the prices of goods produced by firms represented in Fig. 8. This corresponds fairly well to basic intuition, as firms in Fig. 7 correspond roughly to firms producing goods (such as electronics and software) that we expect to be purchased, after possible transformations along the supply chain, by retail and communication firms in Fig. 8. The same holds by the inverting the roles of the firms, as idiosyncratic shocks to firms contributing to the left eigenvector will have an effect on the output of firms represented in the right eigenvector.

¹³In a few cases these data were not available.

Although the data available to us are not as complete as we could wish for a full study of the matrix \mathbf{M} , we think nevertheless that it supports some of our qualitative conclusions. For example, in the main text we argue that the power-law firm size distribution could be a feature explained by the proximity of an instability, with the largest firms being those that have a high overlap with the leftmost eigenvectors of \mathbf{M} . To test this, we plot the sales of the 36 firms with the largest contribution to $|r^{\min}\rangle$ and 34 firms with the largest contribution to $|l^{\min}\rangle$ against their contributions of these firms to the IPR or Herfindahl of these vectors¹⁴ in Figure 9, suggesting an increasing relation between the overlap of a firm with $|v^{\lambda-\min}\rangle$ and its sales.

¹⁴This is none other than the quartic root of the scalar product between a firm's vector and $|l, r^{\lambda-\min}\rangle$.

- [1] J. B. Long and C. I. Plosser, Real business cycles, *J. Political Econ.* **91**, 39 (1983).
- [2] J. H. Cochrane, Shocks, *Carnegie-Rochester Conference Series on Public Policy* **41**, 295 (1994).
- [3] B. Bernanke, M. Gertler, and S. Gilchrist, The financial accelerator and the flight to quality, *Rev. Econ. Stat.* **78**, 1 (1996).
- [4] R. E. Lucas, Understanding business cycles, in *Essential Readings in Economics*, edited by S. Estrin and A. Marin (Macmillan Education UK, London, 1995), pp. 306–327.
- [5] B. Dupor, Aggregation and irrelevance in multi-sector models, *J. Monetary Econ.* **43**, 391 (1999).
- [6] R. E. A. Farmer, *The Macroeconomics of Self-fulfilling Prophecies*, Macroeconomics of Self-fulfilling Prophecies (MIT Press, Cambridge, MA, 1999).
- [7] W. A. Brock and S. N. Durlauf, Discrete choice with social interactions, *Rev. Econ. Stud.* **68**, 235 (2001).
- [8] J.-P. Bouchaud, Crises and collective socio-economic phenomena: Simple models and challenges, *J. Stat. Phys.* **151**, 567 (2013).
- [9] K. Anand, A. Kirman, and M. Marsili, Epidemics of rules, rational negligence and market crashes, *Eur. J. Finance* **19**, 438 (2013).
- [10] J. da Gama Batista, J.-P. Bouchaud, and D. Challet, Sudden trust collapse in networked societies, *Eur. Phys. J. B* **88**, 55 (2015).
- [11] X. Gabaix, The granular origins of aggregate fluctuations, *Econometrica* **79**, 733 (2011).
- [12] M. Wyart and J.-P. Bouchaud, Statistical models for company growth, *Phys. A (Amsterdam, Neth.)* **326**, 241 (2003).
- [13] J. Sutton, The variance of firm growth rates: the ‘scaling’ puzzle, *Phys. A (Amsterdam, Neth.)* **312**, 577 (2002).
- [14] J. di Giovanni, A. A. Levchenko, and I. Mejean, Large firms and international business cycle comovement, *Am. Econ. Rev.* **107**, 598 (2017).
- [15] D. Acemoglu, V. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi, The network origins of aggregate fluctuations, *Econometrica* **80**, 1977 (2012).
- [16] D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi, Networks, shocks, and systemic risk, in *The Oxford Handbook of the Economics of Networks*, edited by Y. Bramoullé, A. Galeotti, and B. Rogers (Oxford University Press, Oxford, 2016).
- [17] D. Acemoglu, U. Akcigit, and W. Kerr, Networks and the macroeconomy: An empirical exploration, *NBER Macroecon. Annu.* **30**, 273 (2015).
- [18] A. G. Haldane and R. M. May, Systemic risk in banking ecosystems, *Nature (London)* **469**, 351 (2011).
- [19] P. Gai and S. Kapadia, Contagion in financial networks, *Proc. R. Soc. A* **466**, 2401 (2010).
- [20] P. Tasca and S. Battiston, Diversification and financial stability, CCSS Working Paper No. 11-001, 2011.
- [21] F. Caccioli, P. Barucca, and T. Kobayashi, Network models of financial systemic risk: A review, *J. Comput. Social Sci.* **1**, 81 (2018).
- [22] A. T. Foerster, P.-D. G. Sarte, and M. W. Watson, Sectoral versus aggregate shocks: A structural factor analysis of industrial production, *J. Political Econ.* **119**, 1 (2011).
- [23] V. Carvalho and X. Gabaix, The great diversification and its undoing, *Am. Econ. Rev.* **103**, 1697 (2013).
- [24] D. Acemoglu, A. Ozdaglar, and A. Tahbaz-Salehi, The network origins of large economic downturns, NBER Working Papers 19230, National Bureau of Economic Research, Inc., 2013.
- [25] J. Bonart, J.-P. Bouchaud, A. Landier, and D. Thesmar, Instabilities in large economies: aggregate volatility without idiosyncratic shocks, *J. Stat. Mech.: Theory Exp.* (2014) P10040.
- [26] D. Hawkins, Some conditions of macroeconomic stability, *Econometrica* **16**, 309 (1948).
- [27] D. Hawkins and H. A. Simon, Note: Some conditions of macroeconomic stability, *Econometrica* **17**, 245 (1949).
- [28] P. Bak, K. Chen, J. Scheinkman, and M. Woodford, Aggregate fluctuations from independent sectoral shocks: Self-organized criticality in a model of production and inventory dynamics, NBER Working Paper No. 4241, 1992.

- [29] J. Scheinkman and M. Woodford, Self-organized criticality and economic fluctuations, *Am. Econ. Rev.* **84**, 417 (1994).
- [30] P. Bak, *How Nature Works: The Science of Self-Organized Criticality* (Springer Science+Business Media, New York, 1996).
- [31] R. M. May, Will a large complex system be stable? *Nature (London)* **238**, 413 (1972).
- [32] G. Bunin, Ecological communities with Lotka-Volterra dynamics, *Phys. Rev. E* **95**, 042414 (2017).
- [33] G. Bunin, Interaction patterns and diversity in assembled ecological communities, [arXiv:1607.04734](https://arxiv.org/abs/1607.04734).
- [34] G. Biroli, G. Bunin, and C. Cammarota, Marginally stable equilibria in critical ecosystems, *New J. Phys.* **20**, 083051 (2018).
- [35] M. Nirei and J. Scheinkman, Self-organization of inflation volatility, Social Science Research Network, 2019.
- [36] J. Moran and J.-P. Bouchaud (unpublished).
- [37] D. R. Raval, The micro elasticity of substitution and non-neutral technology, *RAND J. Econ.* **50**, 147 (2019).
- [38] J. Reed and B. Simon, Car components hit by Japan aftershock, *Financial Times* (2011), <https://www.ft.com/content/9ac4d7e2-595f-11e0-bc39-00144feab49a>.
- [39] D. Fisher, Japan disaster shakes up supply-chain strategies, Working Knowledge: Harvard Business School (2011), <https://hbswk.hbs.edu/item/japan-disaster-shakes-up-supply-chain-strategies>.
- [40] N. Tajitsu, Five years after Japan quake, rewiring of auto supply chain hits limits, *Reuters* (2016), <https://www.reuters.com/article/us-japan-quake-supplychain-idUSKCN0WW09N>.
- [41] C. Colon and J.-P. Bouchaud, Transition from plasticity to instability in the structure of economic networks (unpublished).
- [42] M. Fiedler and V. Ptak, On matrices with non-positive off-diagonal elements and positive principal minors, *Czech. Math. J.* **12**, 382 (1962).
- [43] V. L. Girko, Circular law, *Theory Probab. Appl.* **29**, 694 (1985).
- [44] T. Tao and V. Vu, Random matrices: the circular law, *Commun. Contemp. Math.* **10**, 261 (2008).
- [45] F. L. Metz, I. Neri, and T. Rogers, Spectral theory of sparse non-Hermitian random matrices, *J. Phys. A: Math. Theor.* (2019), doi:[10.1088/1751-8121/ab1ce0](https://doi.org/10.1088/1751-8121/ab1ce0).
- [46] E. P. Wigner, Random matrices in physics, *SIAM Rev.* **9**, 1 (1967).
- [47] G. Biroli and R. Monasson, A single defect approximation for localized states on random lattices, *J. Phys. A: Math. Gen.* **32**, L255 (1999).
- [48] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Spectra of “real-world” graphs: Beyond the semicircle law, *Phys. Rev. E* **64**, 026704 (2001).
- [49] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74**, 47 (2002).
- [50] T. Rogers and I. P. Castillo, Cavity approach to the spectral density of non-Hermitian sparse matrices, *Phys. Rev. E* **79**, 012101 (2009).
- [51] Y. Fyodorov and A. Mirlin, Localization in Ensemble of Sparse Random Matrices, *Phys. Rev. Lett.* **67**, 2049 (1991).
- [52] R. Kühn, Spectra of sparse random matrices, *J. Phys. A: Math. Theor.* **41**, 295002 (2008).
- [53] F. Benaych-Georges and S. Péché, Largest eigenvalues and eigenvectors of band or sparse random matrices, *Electron. Commun. Probab.* **19**, 9 (2014).
- [54] I. Neri and F. L. Metz, Eigenvalue Outliers of Non-Hermitian Random Matrices with a Local Tree Structure, *Phys. Rev. Lett.* **117**, 224101 (2016).
- [55] G. Biroli, G. Semerjian, and M. Tarzia, Anderson model on Bethe lattices: Density of states, localization properties and isolated eigenvalue, *Prog. Theor. Phys. Suppl.* **184**, 187 (2010).
- [56] D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos, Epidemic thresholds in real networks, *ACM Trans. Inf. Syst. Security* **10**, 1 (2008).
- [57] M. Newman, *Networks: An Introduction* (Oxford University Press, New York, 2010).
- [58] C. Castellano and R. Pastor-Satorras, Relating Topological Determinants of Complex Networks to Their Spectral Properties: Structural and Dynamical Effects, *Phys. Rev. X* **7**, 041024 (2017).
- [59] J. Baik, G. B. Arous, and S. Péché, Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices, *Ann. Probab.* **33**, 1643 (2005).
- [60] I. M. Lifshitz, The energy spectrum of disordered systems, *Adv. Phys.* **13**, 483 (1964).
- [61] D. J. Thouless, Electrons in disordered systems and the theory of localization, *Phys. Rep.* **13**, 93 (1974).
- [62] J. Sethna, *Statistical Mechanics: Entropy, Order Parameters and Complexity*, Oxford Master Series in Physics (Oxford University Press, Oxford, 2006).
- [63] R. L. Axtell, Zipf distribution of U.S. firm sizes, *Science* **293**, 1818 (2001).
- [64] P. Le Doussal, M. Müller, and K. J. Wiese, Avalanches in mean-field models and the Barkhausen noise in spin-glasses, *Europhys. Lett.* **91**, 57004 (2010).
- [65] P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, and F. Zamponi, Fractal free energy landscapes in structural glasses, *Nat. Commun.* **5**, 3725 (2014).
- [66] M. Müller and M. Wyart, Marginal stability in structural, spin, and electron glasses, *Annu. Rev. Condens. Matter Phys.* **6**, 177 (2015).
- [67] E. Atalay, A. Hortaçsu, J. Roberts, and C. Syverson, Network structure of production, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5199 (2011).
- [68] J. I. Perotti, O. V. Billoni, F. A. Tamarit, D. R. Chialvo, and S. A. Cannas, Emergent Self-Organized Complex Network Topology Out of Stability Constraints, *Phys. Rev. Lett.* **103**, 108701 (2009).
- [69] A. De Martino, M. Marsili, and I. Pérez Castillo, Typical properties of large random economies with linear activities, *Macroecon. Dyn.* **11**, 34 (2007).
- [70] M. Bardoscia, G. Livan, and M. Marsili, Statistical mechanics of complex economies, *J. Stat. Mech.: Theory Exp.* (2017) 043401.
- [71] C. Syverson, What determines productivity? National Bureau of Economic Research Working Paper No. 15712, 2010.