How Damage Diversification Can Reduce Systemic Risk

Rebekka Burkholz,* Antonios Garas,† and Frank Schweitzer‡

ETH Zurich, Chair of Systems Design
Weinbergstrasse 56/58, 8092 Zurich, Switzerland

We consider the problem of risk diversification in complex networks. Nodes represent e.g. financial actors, whereas weighted links represent e.g. financial obligations (credits/debts). Each node has a risk to fail because of losses resulting from defaulting neighbors, which may lead to large failure cascades. Classical risk diversification strategies usually neglect network effects and therefore suggest that risk can be reduced if possible losses (i.e., exposures) are split among many neighbors (exposure diversification, ED). But from a complex networks perspective diversification implies higher connectivity of the system as a whole which can also lead to increasing failure risk of a node. To cope with this, we propose a different strategy (damage diversification, DD), i.e. the diversification of losses that are imposed on neighboring nodes as opposed to losses incurred by the node itself. Here, we quantify the potential of DD to reduce systemic risk in comparison to ED. For this, we develop a branching process approximation that we generalize to weighted networks with (almost) arbitrary degree and weight distributions. This allows us to identify systemically relevant nodes in a network even if their directed weights differ strongly. On the macro level, we provide an analytical expression for the average cascade size, to quantify systemic risk. Furthermore, on the meso level we calculate failure probabilities of nodes conditional on their system relevance.

I. INTRODUCTION

In the course of globalization and technical advancement, systems become more interconnected and system components more dependent on the functioning of others [13, 26]. In particular for socio-economic networks [25] and financial networks [11, 14] we observe an increase in coupling strength and complexity at the same time. Examples are global supply chains, but also technical systems, as, e.g., power-grids in the USA and Europe [6].

Increased dependence, under normal conditions, has advantages for the efficient operation of systems. But it also makes systems more vulnerable if some of their components break down. Precisely, the failure of a few components can get amplified such that it leads to system wide failure cascades. Classical risk management theories suggest that higher risk diversification reduces the failure risk of the component and consequently the one of the system as a whole. But these conclusions are based on the assumption that a component only depends on other components that are independent among each other.

This assumption does not hold in most real world systems. Therefore, we model such dependences by a network where system components are represented as nodes and direct interactions as links between nodes. Even if the neighbors of a node are not linked directly, they can be coupled indirectly through the network. This effect increases if nodes have a high degree, i.e. are well diversified in a financial context.

In order to study simple risk diversification strategies in the context of systemic risk a model introduced by Watts [28] has been adopted to financial networks of interbank lending [4, 10, 23]. We call this the *exposure diversification* approach (ED), and in this paper we will contrast it with an approach where the *impact* of a failing node is diversified, instead. In a banking network this would correspond to a policy where each financial institution is only allowed to get into a limited amount of debt. This *damage diversification* approach (DD) has the potential to reduce systemic risk significantly, since it counterbalances the failure amplification caused by hubs. Such hubs, because of their large number of neighbors, considerably affect the network in case of failure. Consequently, policy discussions center around the question of how to prevent the failure of hubs, e.g. by increasing their robustness (i.e. capital buffers in finance) [3, 23]. Our findings from the DD approach suggest to complement such regulatory efforts by the mitigation of the *impact* of the failures of well connected nodes.

In this paper we present simulations as well as analytic derivations of network ensemble averages in the limit of infinite network size, where two quantities on the system level are given: a) the degree distribution, which defines the number of direct neighbors of a node and thus limits their respective diversification strategies, and b) the distribution of robustness which is later defined by the failure threshold. The analytic method (also known as heterogeneous mean field or branching process approximation) has been derived for the ED approach [12]. But, it does not capture processes where the impact of a failing neighbor depends on its specific properties (e.g., its degree or robustness) as it is required for the treatment of the DD variant. Therefore, we extend the branching process approximation to the latter case and generalize it for the application to weighted random network models,

---

* rburkholz@ethz.ch
† agaras@ethz.ch
‡ fschweitzer@ethz.ch

whose weight statistics could be deduced from data, taking also into account different properties of neighboring nodes. This way, we generalize the analytic treatment to match application scenarios, by better understanding the role of different risk diversification strategies. Our analysis provides an interesting mesoscopic perspective with the comparison of failure probabilities of nodes with different diversification strategies, which should accompany the study of macroscopic measures, such as the average cascade, since it is crucial for the identification of system-relevant nodes.

II. MODELING EXPOSURE VERSUS DAMAGE DIVERSIFICATION

In a weighted network with N nodes a link with (non-zero) weight w_{ji} between two nodes j and i represents an exposure of i to its network neighbor j. Each node j can fail either initially or later in (discrete) time t because of a propagating cascading process. Its (binary) state then switches from $s_i(t) = 0$ (ok) to $s_i(t) = 1$ (failed), without the possibility to recover.

If the node j fails, its neighbor i faces the loss w_{ji}. The total amount of i's losses sums up to its fragility

$$\phi_i(t+1) = \sum_j w_{ji} s_j(t). \tag{1}$$

If ϕ_i exceeds the threshold θ_i (i.e., $\phi_i \geq \theta_i$), then node i fails as well. Hence, θ_i expresses the *robustness* of node i. In this way a cascade of failing nodes develops over time, which can even span the whole network.

We measure the cascade size by the final fraction of failed nodes

$$\rho_N = \lim_{t \to \infty} \frac{1}{N} \sum_{i=1}^{N} s_i(t), \tag{2}$$

when no further failures are triggered. Any cascade stops after at most N time steps, since at least one node needs to fail at a time to keep the process ongoing.

The cascade dynamics are fully deterministic for given thresholds and weights on a fixed network. Still, many systems do not remain constant over time, and may also fluctuate by their exposure to large cascades. Consequently, it is reasonable to quantify the risk of large cascades with respect to macroscopic distributions that allow for microscopic variations of the weighted network and the thresholds. The average cascade size with respect to these distributions then defines our measure of systemic risk.

In such a setting, we study the influence of two diversification variants. The difference between the ED and the DD approach is in defining the weights w_{ij}. Precisely,

$$w_{ij}^{(\text{ED})} = \frac{1}{k_j} \; ; \quad w_{ij}^{(\text{DD})} = \frac{1}{k_i} \tag{3}$$

for *exposure diversifications* (ED) and for *damage diversifications* (DD) respectively. Here, k_i denotes the degree of node i, i.e., the number of its neighbors. We note that in general $w_{ij}^{(\text{ED})} \neq w_{ji}^{(\text{ED})}$ and $w_{ij}^{(\text{DD})} \neq w_{ji}^{(\text{DD})}$. Still, we call the network undirected. This differs from the approach for directed networks [1, 2, 10] where the neighbors whose failures impact a node are distinct from the ones who face a loss in case of the node's failure. In this case one node is exposed to the other, but not vice versa. In our case, however, once there is a link between two nodes, each can impact the other, but the amount of the loss can be different.

In the ED case every neighbor is treated identical. I.e. a failure of any neighboring node j exposes a node i to the same risk of a loss $1/k_i$. The higher the degree k_i of node i, the better it diversifies its total exposure of $\sum_{j \in \text{nb}(i)} 1/k_i = 1$, where $\text{nb}(i)$ denotes the neighbors of node i. I.e., in the ED case, single failures of neighbors j become less harmful to node i if it has a larger number of neighbors. On the other hand, the failure of a hub impacts many other nodes and is thus problematic from a system perspective.

In the DD case the impact of a hub is effectively reduced. I.e., the failure of a hub j has total impact $\sum_{i \in \text{nb}(j)} 1/k_j = 1$, which reduces the impact on a neighboring node i to $1/k_j$ instead of 1 in the ED case. Hence, better diversified nodes damage each of their neighbors effectively less in case of a failure.

We are interested in how the heterogeneity of such diversification strategies and the heterogeneity of thresholds impacts systemic risk for large systems. In both model variants the diversification strategies are determined by the degrees of the nodes.

Consequently, we study the fraction of failed nodes as an average over a whole class of networks characterized by a fixed degree distribution $p(k)$ and a fixed threshold distribution $F_\Theta(\theta)$ in the limit of infinitely large networks ($N \to \infty$). The network generation method with fixed $p(k)$ is known as configuration model [18, 20], where all possible network realizations (without multiple edges and self-loops, but with a prescribed degree sequence) are equally likely. The thresholds are then assigned to nodes independently of each other, and independently of their degree according to $F_\Theta(\theta)$, although the independence of the degree is not a necessary assumption.

For the ED approach, the average fraction of failed nodes at the end of a cascade can be calculated on random networks with given degree distribution $p(k)$ and threshold distribution $F_\Theta(\theta)$ [12]. To obtain the results, a branching process approximation was used, also known as heterogeneous mean field approximation (HMF) or as local tree approximation (LTA) [7]. This approximation was studied in many subsequent works. It was generalized for directed and undirected weighted networks [1, 15], it was shown to be accurate even for clustered networks with small mean inter-vertex distance [17], and the influence of degree-degree correlations has been investigated [7, 22]. According to a general

framework introduced by Lorenz et al. [16], the ED and DD approach belong to the constant load class, where the ED is called the inward variant, while the DD is identified as outward variant. Still, the risk reduction potential of the latter has not been understood so far, since a system's exposure to systemic risk has been only explored on fully-connected [16] or regular [27] networks, where both model variants coincide.

In order to study the DD approach on more general networks, we generalized and extended the existing approximations, which for the case of ED were proven to be exact [1, 15]. Now, we can treat more general processes where the directed weights in an undirected network can depend on properties of both nodes, the failing as well as the loss facing one. Here, in contrast to [1], nodes can depend on each other, and in contrast to [15] they do so in an non-symmetric way.

We show in Section IV that our approach leads to very good agreements with simulations on finite Erdös-Rényi networks and can also be applied to a more realistic setting where, e.g., the nodes' degrees follow a scale free distribution.

III. ANALYTIC FRAMEWORK

A. Local tree approximation

In the configuration model a node i is characterized only by its degree k_i so that its failure probability $\mathbb{P}(F|k_i)$ depends solely on this information. Hence the (average) final fraction of failed nodes is of the form

$$\rho^* = \sum_k p(k)\mathbb{P}(F|k). \qquad (4)$$

The quantity ρ^* allows for two different interpretations. On the system level, ρ^* measures the final fraction of failed nodes, and $\mathbb{P}(F|k)$ the fraction of failed nodes with degree k. On the node level, ρ^* can be seen as probability for a node to be failed, but if the node's degree k is known, then its actual failure probability is given by $\mathbb{P}(F|k)$. In the following, we proceed with successively decomposing $\mathbb{P}(F|k)$ into sums over products between conditional probabilities that assume more information about the network neighborhood, and probabilities that the neighborhood is in the assumed state.

1. The conditional failure probability

The computation of $\mathbb{P}(F|k)$ relies on the assumption of an infinite network size ($N \to \infty$), since the clustering coefficient for the configuration model vanishes in the limit, if the second moment of the degree distribution is finite [19]. Consequently, the topology simplifies to locally tree-like networks [7], where neighbors of a node are not connected among each other, as illustrated in
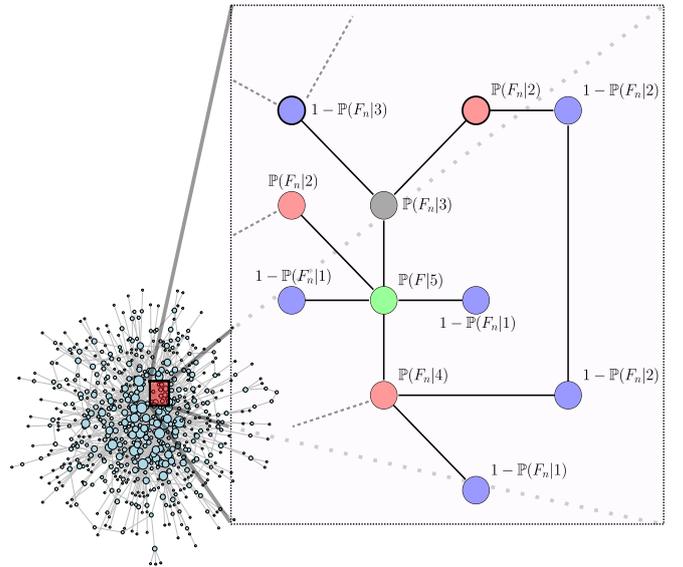

FIG. 1. Illustration of the local tree approximation. The green node is the focal node. Its failure probability $\mathbb{P}(F|k = 5)$ can be computed according to Equation (5), and depends on the state of its neighbors: Here, the two red ones and the gray one have failed, while the two blue ones are still functional. The neighbors' conditional failure probabilities $\mathbb{P}(F_n|k)$ rely on the failure probabilities of their own neighbors without regarding the green focal point.

Fig. 1. The node under consideration with degree k is colored in green and is also called the focal node. Its failure probability $\mathbb{P}(F|k)$ decomposes into a sum over two factors, due to the combination of a locally tree-like network structure with the assumption that the local neighborhood defines the state of a node already:

$$\mathbb{P}(F|k) = \sum_{n=0}^{k} \mathbb{P}(F|k,n)b(n,k,\pi). \qquad (5)$$

The factor $b(n,k,\pi)$ describes the general state of the neighborhood, namely the probability that among the k neighbors of a node exactly n have failed. The factor $\mathbb{P}(F|k,n)$ gives the probability that a node with degree k fails after n of its neighbors have failed. Therefore, it takes into account the ability of a node to withstand shocks (i.e., failing neighbors).

If some of the neighbors of a node would be connected, which violates the local tree-like assumption, we would need to consider all possible (temporal) orders of their failures. Instead, the configuration model allows for assigning every neighbor the same failure probability π, since no degree-degree correlations are present there, and each neighbor's failure is independent of the failure of the others because of the locally tree-like network structure. Consequently, the number of failed neighbors of a node is binomially distributed so that n neighbors can fail with

probability

$$b(n, k, \pi) = \binom{k}{n} \pi^n (1 - \pi)^{k-n}.$$

However, the probability $\mathbb{P}(F|k, n)$ that such an event causes the failure of the considered node with degree k may depend on specific properties of the neighbors, like e.g., their degrees k_i and their failure probabilities $\mathbb{P}(F_n|k_i)$. By allowing this dependence, we introduce a generalization of the existing heterogeneous mean field approximation which enables the analytical treatment of processes where failing nodes have different influences on their neighbors according to their degree.

Also the conditional failure probability $\mathbb{P}(F|k, n)$ can be decomposed into the sum over two factors

$$\mathbb{P}(F|k, n) = \sum_{\mathbf{k_n} \in I^n} p(\mathbf{k_n}|k, n) \mathbb{P}(F|k, \mathbf{k_n}), \qquad (6)$$

where the sum runs over all possible configurations of neighbors' degrees denoted by I. $\mathbf{k_n}$ is an abbreviation for a vector $(k_1, \cdots, k_n)$ of failed neighbors' degrees and takes values in I. Generally, $I = \mathbb{N}$ or $I = [c] := \{1, \cdots, c\}$ in the presence of a finite cutoff c. Such a cutoff is inevitable in numerical computations or in the observation of (finite) real world systems, while it guarantees the finiteness of the second moment of $p(k)$.

The probability $p(\mathbf{k_n}|k, n)$ captures the precise state of the neighborhood given that exactly n neighbors have failed. More precisely, it is the probability that the n failed neighbors of a node with degree k have degrees $\mathbf{k_n}$. This factor may depend on the failing probabilities of neighbors $\mathbb{P}(F_n|k_1), \cdots, \mathbb{P}(F_n|k_n)$, while conditioned on these failures, we denote $\mathbb{P}(F|k, \mathbf{k_n})$ the probability that a node with degree k fails. The latter is determined by the specific cascading model, and will be discussed later for our two diversification models.

2. The state of the neighborhood

In order to compute the failure probability

$$\mathbb{P}(F|k) = \sum_{n=0}^{k} b(n, k, \pi) \sum_{\mathbf{k_n} \in I^n} p(\mathbf{k_n}|k, n) \mathbb{P}(F|k, \mathbf{k_n}) \quad (7)$$

we need to derive the state of the neighborhood as described by the average failure probability of a neighbor π, and the failure probability $p(\mathbf{k_n}|k, n)$ that the n failed neighbors have degrees $\mathbf{k_n}$. Both depend on the degree distribution of a neighbor $p_n(k)$ as well as its failure probability $\mathbb{P}(F_n|k)$.

It is important to note that a neighbor with degree k (illustrated by the gray node in Fig. 1) does not fail with probability $\mathbb{P}(F|k)$, since one of its links leads to the focal node (illustrated by the green node in Fig. 1). Conditional on the event that the focal node has not failed yet,

only the remaining $k - 1$ neighbors of the gray neighbor (colored with bold fringe) could have caused the failure of the gray neighbor of the focal node. This model property is called the *Without Regarding Property* (WOR) by Hurd and Gleeson [15]. Therefore, a neighbor's failure probability $\mathbb{P}(F_n|k)$ is

$$\mathbb{P}(F_n|k) = \sum_{n=0}^{k-1} b(n, k-1, \pi) \sum_{\mathbf{k_n} \in I^n} p(\mathbf{k_n}|k, n) \mathbb{P}(F|k, \mathbf{k_n}).$$
$$(8)$$

It remains to calculate $p(\mathbf{k_n}|k, n)$ and the unconditional failure probability of a neighbor π. Both depend on the degree distribution $p_n(k)$ of a neighbor. Because of the local tree approximation, it is independent of the degree distribution of the other nodes in the network:

$$p_n(k) := \frac{kp(k)}{z}, \qquad (9)$$

where $z := \sum_k kp(k)$ denotes the normalizing average degree. $p_n(k)$ is proportional to the degree k in the configuration model, because each of a neighbor's k links could possibly connect the neighbor with the focal node (see, e.g., [19]).

We, therefore, obtain the unconditional failure probability π of a neighbor by

$$\pi = \sum_k p_n(k) \mathbb{P}(F_n|k) = \sum_k \frac{kp(k)}{z} \mathbb{P}(F_n|k). \qquad (10)$$

Similarly, the degree distribution of a neighbor conditional on its failure can be written as $\mathbb{P}(F_n|k)p(k)k/z\pi$ so that we can calculate the probability $p(\mathbf{k_n}|k, n)$ that the n failed neighbors have degrees $\mathbf{k_n}$ by

$$p(\mathbf{k_n}|k, n) = \prod_{j=1}^{n} \frac{p(k_j)k_j \mathbb{P}(F_n|k_j)}{z\pi}, \qquad (11)$$

since the neighbors are independent of each other, according to the locally tree-like network structure.

3. Fixed point iteration for the conditional failure probability

In short, the vector $\mathbb{P}(\mathbf{F_n}|\mathbf{k}) = (\mathbb{P}(F_n|k))_{k \in \{1, \cdots, c\}}$ turns out to be a fixed point of a vector valued function $\mathbf{L}(\mathbf{p})$ so that for the k−th component we have:

$$\mathbb{P}(F_n|k) = \sum_{n=0}^{k-1} b(n, k-1, \pi) \sum_{\mathbf{k_n} \in I^n} \mathbb{P}(F|k, \mathbf{k_n}) \cdot$$
$$\cdot \prod_{j=1}^{n} \frac{p(k_j)k_j \mathbb{P}(F_n|k_j)}{z\pi} = L_k(\mathbb{P}(\mathbf{F_n}|\mathbf{k})).$$

Such a fixed point exists according to the Knaster-Tarski Theorem, since $\mathbf{L}(\mathbf{p})$ is monotone with respect to a partial ordering and maps the complete lattice $[0, 1]^I$ onto itself. [21]

Thus, starting from an initial vector $\mathbb{P}(\mathbf{F_n}|\mathbf{k})^{(\mathbf{0})}$ which is defined by the considered cascading model, we can compute the fixed point iteratively by

$$\mathbb{P}(\mathbf{F_n}|\mathbf{k})^{(\mathbf{t+1})} = \mathbf{L}\left(\mathbb{P}(\mathbf{F_n}|\mathbf{k})^{(\mathbf{t})}\right), \tag{12}$$

with

$$\pi^{(t)} = \sum_k p_n(k)\mathbb{P}(F_n|k)^{(t)}. \tag{13}$$

Each iteration step (t) corresponds to one discrete time step of the cascading process so that

$$\rho^{(t)} = \sum_k p(k)\mathbb{P}(F|k)^{(t)} \tag{14}$$

can be interpreted as average fraction of failed nodes in the network at time t. Note that the relation between $\mathbb{P}(\mathbf{F}|\mathbf{k})^{(\mathbf{t})}$ and $\mathbb{P}(\mathbf{F_n}|\mathbf{k})^{(\mathbf{t})}$ is described by Equation (7) and Equation (11).

4. Simplification for homogeneous failure probability

In case the impact of a failing neighbor does not depend on its degree, the failure probability $\mathbb{P}(F|k,\mathbf{k_n}) = \mathbb{P}(F|k,n)$ does not depend on the degrees $\mathbf{k_n}$ of its n failed neighbors and Equation (12) can be simplified to

$$\begin{aligned}
\mathbb{P}(F_n|k) &= \sum_{n=0}^{k-1} b(n,k-1,\pi)\mathbb{P}(F|k,n) \\
&\quad \prod_{j=1}^{n}\frac{1}{\pi}\sum_{k_j\in I}\frac{p(k_j)k_j\mathbb{P}(F_n|k_j)}{z} \\
&= \sum_{n=0}^{k-1} b(n,k-1,\pi)\mathbb{P}(F|k,n),
\end{aligned} \tag{15}$$

using Equation (10). Inserting this into Equation (10) leads to the fixed point equation

$$\pi = \sum_k \frac{kp(k)}{z}\sum_{n=0}^{k-1} b(n,k-1,\pi)\mathbb{P}(F|k,n) \tag{16}$$

which in this case involves the scalar π instead of the vector $\mathbb{P}(\mathbf{F_n}|\mathbf{k})$ in Equation (12). With this information the final fraction of failed nodes can be computed as

$$\rho^* = \sum_k p(k)\sum_{n=0}^{k} b(n,k,\pi)\mathbb{P}(F|k,n), \tag{17}$$

as already known from the literature [7, 12]. Still, this simpler approach is not able to capture the cascade dynamics of the damage diversification model.

5. The ability of a node to withstand a shock

The only piece missing in our derivations is the model specific probability $\mathbb{P}(F|k,\mathbf{k_n})$ that a node with degree k fails exactly after n of its neighbors with degrees $\mathbf{k_n}$ have failed. This probability captures the failure dynamics, and is thus defined by a node's fragility $\phi(k,\mathbf{k_n})$ and its threshold $\Theta(k)$. The given information about the degrees $k, \mathbf{k_n}$ can in principle enter both variables, although the cumulative threshold distribution $F_{\Theta(k)}$ tends to depend solely on properties of the node itself, e.g. the degree k. Because a node fails, if its fragility exceeds its threshold, we have

$$\mathbb{P}(F|k,\mathbf{k_n}) = \mathbb{P}\left(\Theta(k)\le\phi(k,\mathbf{k_n})\right). \tag{18}$$

More generally, with respect to known weight distributions $p_{W(k_j,k)}$ for given degree k_j of a neighbor and the degree k of a node, Equation (18) reads:

$$\begin{aligned}
\mathbb{P}(F|k,\mathbf{k_n}) &= \mathbb{P}\left(\Theta(k)\le\sum_{j=1}^{n}W(k_j,k)\right) \\
&= \int F_{\Theta(k)}(w)\left(p_{W(k_1,k)}*\cdots*p_{W(k_n,k)}\right)(w)\,dw.
\end{aligned} \tag{19}$$

The last equation holds if the weight distributions $p_{W(k_j,k)}$ are independent, and the $*$ denotes a convolution.

In the simpler case of our two model variants, the weights $W(k_j,k)$ are completely deterministic. In accordance with the definition of the weights in Equation (3) we calculate for the ED case

$$\mathbb{P}(F|k,\mathbf{k_n})^{(\mathrm{ED})} = \mathbb{P}(F|k,n)^{(\mathrm{ED})} = F_{\Theta(k)}\left(\frac{n}{k}\right) \tag{20}$$

which is independent of the neighbors' degrees. Consequently, the calculation of the average final fraction of failed nodes can be simplified as outlined in Section III A 4.

For the DD case the fixed point iteration needs to take into account all degrees of the failed neighbors, since they define the loss $1/k_j$ the focal node faces. Thus, we have

$$\mathbb{P}(F|k,\mathbf{k_n})^{(\mathrm{DD})} = F_{\Theta(k)}\left(\sum_{j=1}^{n}\frac{1}{k_j}\right). \tag{21}$$

6. DD case: Correct Heterogeneous Mean Field Approximation (cHMF)

The probability of the failure of a node or neighbor with degree k and n failed neighbors $\mathbb{P}(F|k,n)^{(\mathrm{DD})}$ given by Equation (6) needs to be re-calculated for each fixed point iteration step. For the DD case, this involves the calculation of the convolution of the impact distribution p_{imp} of a failed neighbor, which depends on the iteratively
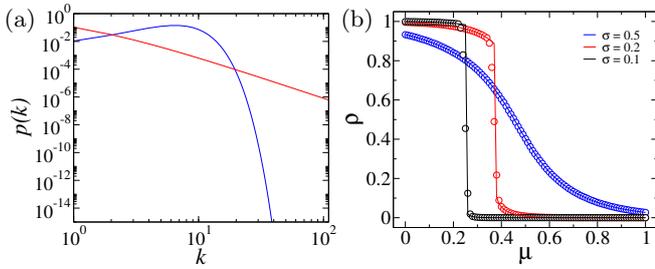
FIG. 2. (a) The studied degree distributions: Poisson distribution with parameter $\lambda = 8$ and cutoff degree $c = 50$ (blue), scale free distribution with exponent $\gamma = 3$ and maximal degree $c = 200$ (red) in log-log scale. (b) Comparison of the final fraction of failed nodes obtained by simulations (symbols) on networks with 1000 nodes as average over 2000 independent realizations with numerical results from the cHMF (lines) for the DD case on Poisson random graphs. The thresholds Θ are normally distributed with mean μ and standard deviation σ ($\Theta \sim \mathcal{N}(\mu, \sigma)$).

updated failure probability $\mathbb{P}(F_n|k)$. More precisely, one failed neighbor inflicts the loss $1/k$ with probability

$$p_{\text{imp}}\left(\frac{1}{k}\right) = \mathbb{P}(F_n|k)\frac{kp(k)}{z\pi} \qquad (22)$$

(that is conditioned on its failure) independently of the other failed neighbors. Thus, the fragility $\phi(k, n)$ of a node with degree k and n failed neighbors is distributed according to the n-th convolution of this impact distribution p_{imp}^{*n} and we have

$$\mathbb{P}(F|k, n) = \mathbb{P}\left(\Theta(k) \leq \phi(k, n)\right) = \sum_l p_{\text{imp}}^{*n}(l)F_\Theta(l),$$

where the inner sum runs over all possible values of the fragility.

Since the convolutions are computationally demanding (in terms of time and especially memory), we approximate p_{imp}^{*n} by first binning it to an equidistant grid and then using Fast Fourier Transformations (FFT) in order to take advantage of the fact that convolutions correspond to simple multiplications in Fourier space (see for instance [9, 24]). Although this is numerically accurate enough for the calculation of the final fraction of failed nodes, for the reporting of the vectors $\mathbb{P}(\mathbf{F}|\mathbf{k})$ and $\mathbb{P}(\mathbf{F_n}|\mathbf{k})$ we use a more precise direct convolution of the binned impact distributions p_{imp}, as is described in the Appendix. Fig. 2 shows that our numerical results coincide with simulations.

7. Neglecting the neighbors' degrees in the failure probability (simpHMF)

Considering the computational complexity of the cHMF approach we described above, it is worth asking whether we can approximate it with a simpler version as,
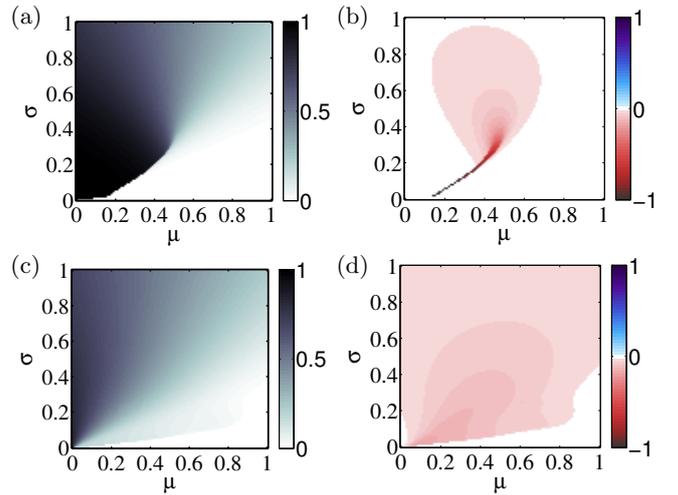

FIG. 3. (a) Fraction of failed nodes obtained by (simpHMF) for the DD case on Poisson random networks with $\lambda = 8$ and $c = 50$. (b) Difference between the correct version (cHMF) and a). (c) Fraction of failed nodes obtained by (simpHMF) for the DD case on scale free networks with $\gamma = 3$ and $c = 200$. (d) Difference between the corresponding correct version (cHMF) version and c). The thresholds Θ are normally distributed with mean μ and standard deviation σ ($\Theta \sim \mathcal{N}(\mu, \sigma)$).

e.g., the one described in Section III A 4, and still obtain reasonably good results.

This would require the probability $\mathbb{P}(F|k, \mathbf{k_n})$ to be independent of the neighbors' degree. Consequently, we would assume that every failed neighbor transfers a load $1/k$ with probability

$$p_{\text{simp}}\left(\frac{1}{k}\right) = \frac{kp(k)}{z}, \qquad (23)$$

although its degree does not need to coincide with k. Therefore, by computing

$$\mathbb{P}(F|k, \mathbf{k_n}) = \sum_l p_{\text{simp}}^{*n}(l)F_\Theta(l) = \mathbb{P}(F|k, n) \qquad (24)$$

we can calculate the failure probability initially without the need to update it in each fixed point iteration. Although this approach is more convenient, as shown in Fig. 3, it is inadequate for the damage diversification variant, especially in combination with skew degree distributions (as, e.g., in case of scale free networks). This is because if we follow this simplified calculation, we lose the risk reducing effect by hubs that are connected to more nodes, and we would draw opposite conclusions about systemic risk. So, as shown in Fig. 3, it is crucial to use the correct HMF to explain our simulation results.

IV.   NUMERICAL RESULTS

We calculate the final fraction of failed nodes for Poisson random graphs and scale free networks with degree distributions

$$p_P(k) := \frac{1}{S_P}\frac{\lambda^k}{k!}, \quad p_S(k) := \frac{1}{S_S}\frac{1}{k^\gamma} \qquad (25)$$

for $k \in \{1, \cdots, c\}$ with normalizing constants

$$S_P := \sum_{k=1}^{c}\frac{\lambda^k}{k!} \quad \text{and} \quad S_S := \sum_{k=1}^{c}\frac{1}{k^\gamma}, \qquad (26)$$

as shown in Fig. 2(a).

The Poisson random graphs are of interest as limit of the well studied Erdös-Rényi random graphs [8], and in comparison to the simulations serve as benchmark for our method (see Fig. 2(b)), while the class of scale free networks is more realistic with respect to real world networks [5].

Similar to [28] and [12] we study normally distributed thresholds $\Theta \sim \mathcal{N}(\mu, \sigma^2)$ with mean μ and standard deviation σ, but, we explore the role of the thresholds' heterogeneity as well as their mean size more extensively. Although our analytic framework also applies to more general cases, here we assume the thresholds to be independent from the degree k of a node.

A.   System Failures

Complementary to [28] we find that increasing the standard deviation can also reduce the cascade size - even though more nodes fail initially, similarly to what was reported for fully connected networks in [16], where the ED and DD cases coincide. The initial fraction of failed nodes is determined by the nodes with negative thresholds ($\Theta \leq 0$) and is thus given by $F_\Theta(0) = \Phi(\frac{-\mu}{\sigma})$, where Φ denotes the cumulative distribution function of the standard normal distribution, as illustrated in the center of Fig. 4.

More pronounced is the sudden regime shift from a region of negligible systemic risk to an almost complete break-down of the system that occurs for Poisson random graphs for moderate robustness ($\mu < 0.4$) and increasing threshold heterogeneity (measured by σ) as shown in the first row of Fig. 4. The same phenomenon can be found in fully-connected or regular random networks [16], but for a more realistic degree distribution, like the scale free distribution with exponent $\gamma = 3$, the regime shift is smoothed out as shown in the third row of Fig. 4.

The low average degree of scale free networks, in our case $z = 1.3643$, seems to make the systems less vulnerable. Also our computations for Poisson random graphs with the same average degree (and thus a parameter $\lambda = 0.6571$, i.e. below the percolation threshold) lead to similar results as for the case of scale free networks. But,

the system size N needs to be considered as well before concluding anything with respect to system safety. Our simulations on scale free networks consisting of $N = 1000$ nodes show similar sudden regime shifts as those for Erdös-Rényi random graphs. But, still, the degree distributions of the simulated networks also tend to have higher average degrees than the one used here.

We can verify the simulation results with the help of the cHMF by taking a representative degree distribution of simulated networks as input. But, the degree distribution of a simulated network would only converge to the theoretical one used in Equation (25) in the limit of infinite network size. In reality though, it might approximate it well for $N = 10^7$.

Still, we can conclude that in very large systems ($N > 10^7$) where few big hubs coexist with a majority of small degree nodes, the lower connectivity can significantly reduce systemic risk, especially in case of the DD.

As shown in Fig. 4 the DD variant exposes the system to a lower risk than the ED, with a minor exception at the phase transition line for the case of Poisson random graphs.

Still, for both our model variants, a higher risk diversification for every node does not lead to better outcomes in general. In fully connected networks - where every node is connected with everyone else and maximal risk diversification is realized - the system is exposed to a higher risk of large cascades in comparison with the studied scale free networks. The same holds for Poisson random graphs except for threshold parameters close to the phase transition, where the change is so abrupt that we cannot draw any conclusions.

In fact, well diversified - and thus also well connected - nodes have a higher failure risk in regions of high systemic risk in both model variants.

B.   Individual failure probabilities

In ED a high diversification is expected to decrease the individual failure risk, since the failure of a high number of neighbors is less probable than the failure of fewer neighbors. As shown in Fig. 5(a) and (b), this intuition applies only to regions of lower systemic risk, but tends to saturate for increasing degree k as does the fragility $\phi \simeq 1/k$. What is not considered in this reasoning is the impact that a failure of a well diversified node has on the overall system stability. Once there is the chance that a few hubs fail, they trigger larger failure cascades in which the failure probability of a neighbor π increases. Once π is big enough, hubs face an even higher failure probability $\mathbb{P}(F|k)$ than nodes with smaller degrees as shown in Fig. 5(a). But, also this effect saturates for increasing degrees.

Mathematically, we see that the shape of the individual failure probabilities $\mathbb{P}(F|k)$ for ED are defined by only two system sizes: π, which indicates the overall system
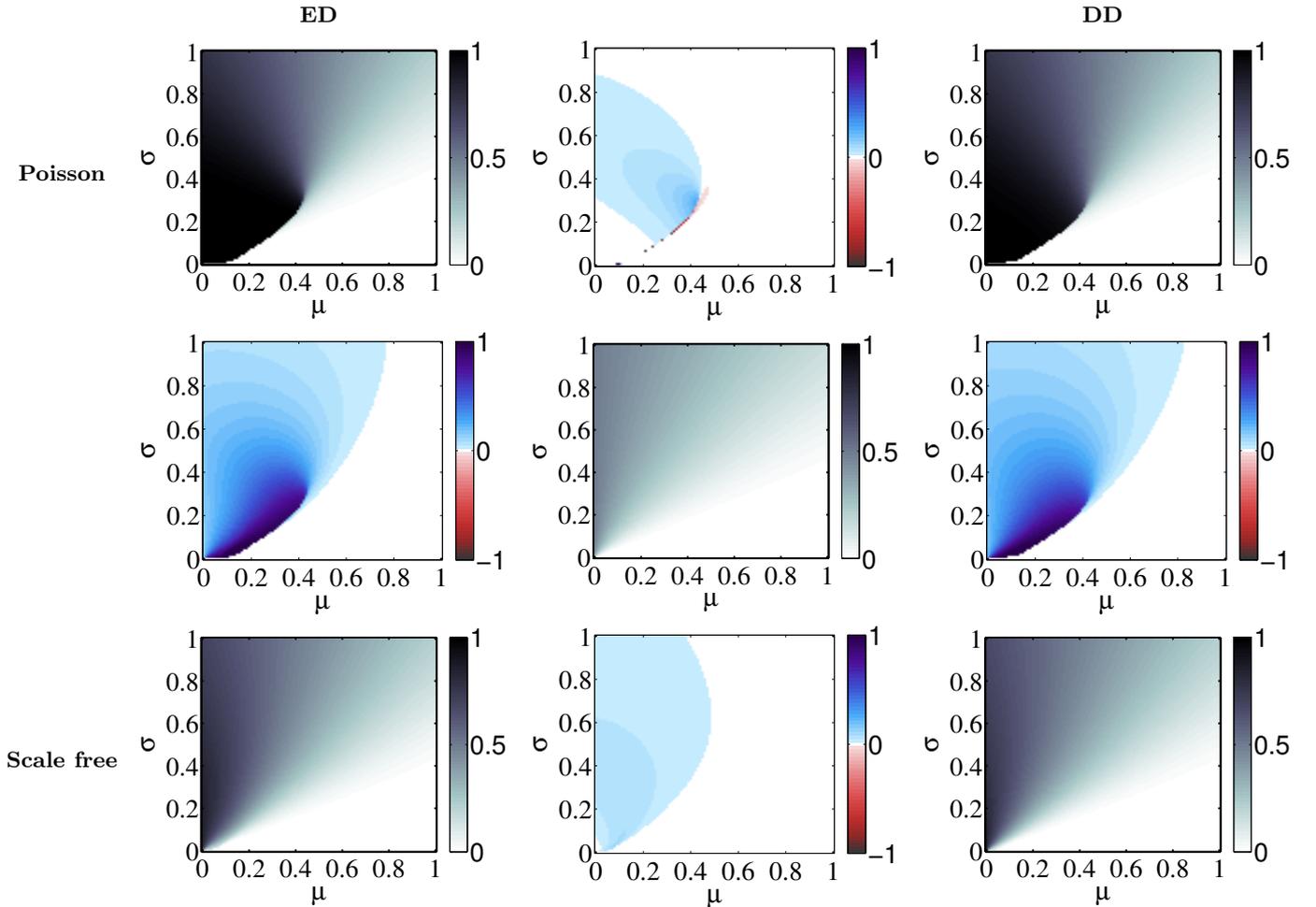
FIG. 4. Phase diagrams for the final fraction of failed nodes ρ for different degree distributions, diversification variants, and their differences. The thresholds Θ are normally distributed with mean μ and standard deviation σ ($\Theta \sim \mathcal{N}(\mu, \sigma)$). The darker the color the higher is the systemic risk. First row: Poisson distribution with parameter $\lambda = 8$ and cutoff degree $c = 50$ for the ED (left) and DD (right). The middle shows their difference $\rho^{(\text{ED})} - \rho^{(\text{DD})}$. Third row: Scale free distribution with exponent $\gamma = 3$ and maximal degree $c = 200$ for the ED (left) and DD (right). The middle shows their difference $\rho^{(\text{ED})} - \rho^{(\text{DD})}$. Second row: The difference between the diagrams with Poisson and Scale free degree distributions for the ED variant (left). Similarly for the DD variant (right). In the middle the initial fraction of failed nodes $\rho(0)$ is illustrated. $\rho_o := \rho(0)$ is constant along the lines $\sigma = \mu/\Phi^{-1}(\rho_0)$.

state, and the threshold distribution:

$$\mathbb{P}(F|k) = \sum_{n=0}^{k} b(n, k, \pi) F_\Theta \left(\frac{n}{k} \right).$$

$\mathbb{P}(F|k)$ does not depend directly on the degree distribution, and thus the diversification strategies of the other nodes. Still, if π exceeds a certain threshold (which is close to μ), nodes with higher degree have a higher failure probability.

In contrast, hubs in DD have always a higher failure probability. Each additional neighbor introduces the possibility of a loss, if it fails. Thus the shape of the condi-

tional failure probability

$$\mathbb{P}(F|k) = \sum_{n=0}^{k} b(n, k, \pi) \sum_{l} p_{\text{imp}}^{*n}(l) F_\Theta(l)$$

will always look similar as the ones presented in Fig. 5(c). Consequently, too many nodes with high degrees would increase the vulnerability of the system. Still, the presence of hubs also decreases the overall failure risk by decreasing some of the possible losses. These losses are distributed by $p_{\text{imp}}^{*n}(l)$ and can thus reduce all $\mathbb{P}(F|k)$.

Especially less diversified nodes (which have a chance to survive also large failure cascades) profit from the diversification of others.
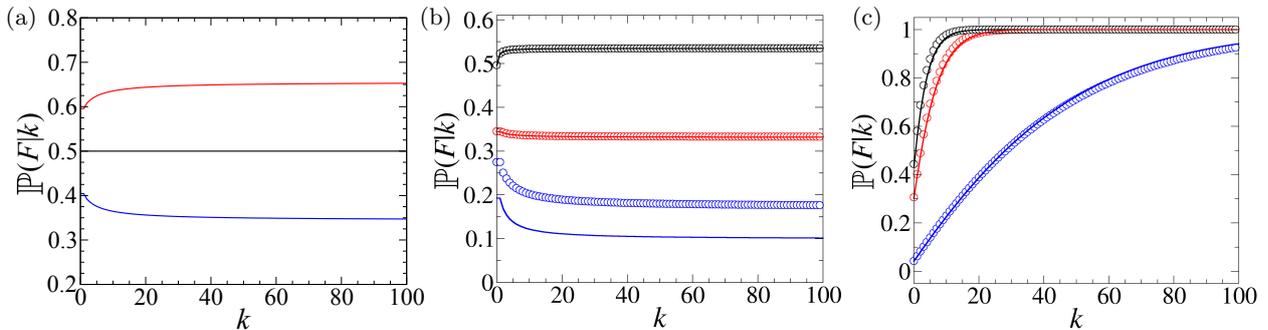
FIG. 5. (a) Conditional failure probability for ED and normally distributed thresholds with parameters $(\mu, \sigma) = (0.5, 0.25)$ with respect to a neighbor's failure probability of $\pi = 0.4$ (blue), $\pi = 0.5$ (black) and $\pi = 0.6$ (red). (b) Conditional failure probability for ED and normally distributed thresholds with parameters $(\mu, \sigma) = (0.3, 0.5)$ (black), $(\mu, \sigma) = (0.5, 0.6)$ (red), $(\mu, \sigma) = (0.5, 0.25)$ (blue) and π defined by cascade equilibrium. Lines indicate a scale free distribution with $\gamma = 3$ and $c = 200$ and thus $z = 1.3643$, while symbols belong to Poisson random graphs with $\lambda = 0.657$ and $c = 100$ (with the same z). (c) As in (b), but for DD.

V. DISCUSSION

Although risk diversification is generally considered to lower the risk of an individual node, on the system level it can even lead to the amplification of failures. With our work, we have deepened the understanding of cascading failure processes in several ways.

At first, we generalize the method to calculate the systemic risk measure, i.e. the average cascade size, to include directed and weighted interactions. This macro measure is complemented by a measure on the meso level, by calculating individual failure probabilities of nodes based on their degree (diversification).

This allows us to compare two different diversification mechanisms: ED (exposure diversification) and DD (damage diversification). As we demonstrate, nodes which diversify their exposures well (i.e. hubs), have a lower failure risk only as long as the system as a whole is relatively robust. But above a certain threshold for the failure probability of neighboring nodes, such hubs are at higher risk than other nodes because they are more exposed to cascading failures. This effect tends to saturate for large degrees.

In general, most regulatory efforts follow the *too big to fail* strategy, and focus on the prevention of the failures of systemic relevant nodes - the hubs. This is mainly achieved by an increase of the thresholds, i.e., capital buffers in a financial context, but, in reality it could be very costly.

With our study of another diversification strategy, the damage diversification, we suggest to accompany regulatory efforts by mitigating the failure of hubs. By limiting the loss that every node can impose on others, the damage potential of hubs and, thus, the overall systemic risk is significantly reduced. While this is systemically preferable, the DD strategy is a two-edged sword: Hubs face an increased failure risk, but many small degree nodes benefit from the diversification of their neighbors. This lowers the incentives for diversification as long as no other bene-

fit, e.g., higher gains in times of normal system operation, comes along with a high degree.

As we show, the systemic relevance of a node is not solely defined by its degree, or connectivity. The size of its impact in case of its failure and thus its ability to cause further failures is crucial. It is a strength of our approach that we can include this impact analytically and obtain a more refined and realistic identification of system relevant nodes.

If necessary, it is still straight forward to adopt our analytical expansions to the case of directed networks, where the failure of one node would impact another node, but not vice versa. Additionally, our approach can be transfered to degree-degree correlated networks. We would expect that a high degree assortativity in the DD could further reduce systemic risk, since many less diversified nodes could be saved by connections to hubs whose failures would impact their neighborhood only little, but this is outside the scope of the current paper, and will be addressed in a future work.

VI. APPENDIX

A. Approximation of the convolution

For the DD it is necessary but computational expensive to calculate the convolution $p_{\text{imp}}^{*^n}$, where

$$p_{\text{imp}}\left(\frac{1}{k}\right) = \mathbb{P}(F_n|k)\frac{kp(k)}{z\pi}$$

denotes the probability of a loss $\phi = 1/k$ caused by the failure of one neighbor with degree k. Given that n neighbors have failed, a node faces a loss ϕ_n, which is a random variable following the law p_{imp}^{*n}.

But the number, l, of ϕ_n's values with nonzero probability mass, which are of relevant size for good accuracy, as well as the number of accumulation points grows exponentially with the order n.

In the end we are interested in calculating the failure probability $\sum_l p_{\mathrm{imp}}^{*n}(l) F_\Theta(l)$ of a node so that it suffices to compute p_{imp}^{*n} accurately on an interval $[0, b] \subset \mathbb{R}$, where the threshold distribution F_Θ is effectively smaller than 1. Otherwise we consider the summand

$$\sum_{l>b} p_{\mathrm{imp}}^{*n}(l) F_\Theta(l) \simeq 1 - \sum_{l \leq b} p_{\mathrm{imp}}^{*n}(l).$$

We vary the parameters μ and σ of the threshold distribution between 0 and 1 so that we can savely set $b = 5$. Next, we partition $[0, b]$ into J small intervals $I_j =](j-1)h, jh]$ of length h, with $j = 1, \cdots, J$. For small enough h (here $h = 10^{-5}$) it is numerically precise enough to assume the approximation of p_{imp}^{*n} to be constant on J.

a. Convolution with the help of FFT In the standard (and faster) algorithm that we use, we simply bin p_{imp} to $[0, b]$ by

$$\hat{p}_{\mathrm{imp}}(jh) := \begin{cases} \sum_{k:(j-1)h < \frac{1}{k} \leq jh} p(k) & j = 1, \cdots, J, \\ 0 & \text{otherwise.} \end{cases}$$

Then, we apply the Fast Fourier Transformation (FFT) [9, 24], and take the n-th power of the distribution and transform it back in order to receive $\hat{p}_{\mathrm{imp}}^{*n}$.

This is numerically accurate enough for the calculation of the final fraction of failed nodes ρ^*, but often does not serve well, if we want to deduce the shape of $\mathbb{P}(F|k)$. For that purpose we have implemented a more precise alternative.

b. Alternative convolution algorithm Here we do not assume $\hat{p}_{\mathrm{imp}}^{*n}$ to be constant on an interval I_j, but uniformly distributed instead. For any (discrete) probability distribution p_X we define its approximation in $x \in [0, b]$ as

$$a\left(p_{\mathrm{imp}}(x)\right) := \sum_{j=1}^{J} \frac{x - (j-1)h}{h} \, \mathbb{1}_{\{(j-1)h < x \leq jh\}} \sum_{y:(j-1)h < y \leq jh} p_X(y).$$

Thus, we get for p_{imp}

$$a\left(p_{\mathrm{imp}}(x)\right) := \sum_{j=1}^{J} \frac{x - (j-1)h}{h} \, \mathbb{1}_{\{(j-1)h < x \leq jh\}} \sum_{k:(j-1)h < \frac{1}{k} \leq jh} \frac{\mathbb{P}(F_n|k)p(k)k}{z\pi}.$$

This is the initial distribution of an iteration in which we compute $\widehat{p_{\mathrm{imp}}^{*n}}$ in the n-th step by convoluting it first exactly with $\widehat{p_{\mathrm{imp}}^{*n-1}}$ of the previous step with the non-approximated p_{imp}. Afterwards we bin it to the intervals I_j again

$$\widehat{p_{\mathrm{imp}}^{*n}} := a\left(\widehat{p_{\mathrm{imp}}^{*n-1}} * p_{\mathrm{imp}}\right),$$

with

$$\widehat{p_{\mathrm{imp}}^{*1}} := a\left(p_{\mathrm{imp}}\right).$$

[1] Hamed Amini, Rama Cont, and Andreea Minca. Resilience to Contagion in Financial Networks. *arXiv:1112.5687*, 2011.

[2] Hamed Amini, Rama Cont, and Andreea Minca. Stress Testing the Resilience of Financial Networks. *Int. J. Theor. Appl. Finance*, 15:1250006, 2012.

[3] Nimalan Arinaminpathy, Sujit Kapadia, and Robert M May. Size and complexity in model financial systems. *Proc. Natl. Acad. Sci. U.S.A.*, 109:18338, 2012.

[4] Stefano Battiston, Domenico Delli Gatti, Mauro Gallegati, Bruce C. N. Greenwald, and Joseph E. Stiglitz. Credit Default Cascades: When Does Risk Diversification Increase Stability? *Journal of Financial Stability*, 8:138–149, 2012.

[5] M Boss, H Elsinger, M Summer, and S Thurner 4. Network topology of the interbank market. *Quant. Finance*, 4:677–684, 2004.

[6] Charles D Brummitt, Raissa M D'Souza, and E A Leicht. Suppressing cascades of load in interdependent networks. *Proc. Natl. Acad. Sci. U.S.A.*, 109:E680–E689, 2012.

[7] PS Dodds and JL Payne. Analysis of a threshold model of social contagion on degree-correlated networks. *Phys. Rev. E*, 79:066115, 2009.

[8] Paul Erdos and Alfred Renyi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[9] M. Frigo and S.G. Johnson. The Design and Implementation of FFTW3. *Proc. IEEE*, 93:216–231, 2005.

[10] P. Gai and S. Kapadia. Contagion in financial networks. *Proc. R. Soc. A*, 466:2401–2423, 2010.

[11] Antonios Garas, Panos Argyrakis, and Shlomo Havlin. The structural role of weak and strong links in a financial market network. *Eur. Phys. J. B*, 63:265–271, 2008.

[12] James P. Gleeson and Diarmuid Cahalane. Seed size strongly affects cascades on random networks. *Phys. Rev. E*, 75:1–4, 2007.

[13] Ian Goldin and Tiffany Vogel. Global Governance and Systemic Risk in the 21st Century: Lessons from the Financial Crisis. *Global Policy*, 1:4–15, 2010.

[14] Andrew G. Haldane and Robert M. May. Systemic risk in banking ecosystems. *Nature*, 469:351–355, 2011.

[15] TR Hurd and JP Gleeson. On Watts' cascade model with random link weights. *J. Complex Networks*, 1:1–24, 2013.

[16] Jan Lorenz, Stefano Battiston, and Frank Schweitzer. Systemic risk in a unifying framework for cascading processes on networks. *Eur. Phys. J. B*, 71:441–460, 2009.

[17] Sergey Melnik, Adam Hackett, Mason A. Porter, Peter J. Mucha, and James P. Gleeson. The unreasonable effectiveness of tree-based theory for networks with clustering. *Phys. Rev. E*, 83:036112, 2011.

[18] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, 6(1995):161–179, 1995.

[19] Mark E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010.

[20] Mark E. J. Newman, Steven H. Strogatz, and Duncan J Watts. Random graphs with arbitrary degree distributions and their applications. *Physcal Review E*, 64:26118, 2001.

[21] Definition of the partial ordering: Two vectors $\mathbf{x}$, $\mathbf{y} \in [0,1]^I$ are ordered as $\mathbf{x} \le \mathbf{y}$, if and only if $x_i \le y_i$ holds for all their components $i \in I$.

[22] Joshua Payne, Peter Dodds, and Margaret Eppstein. Information cascades on degree-correlated random networks. *Phys. Rev. E*, 80:026125, 2009.

[23] Tarik Roukny, Hugues Bersini, Hugues Pirotte, Guido Caldarelli, and Stefano Battiston. Default cascades in complex networks: topology and systemic risk. *Sci. Rep.*, 3:2759, 2013.

[24] Peter Ruckdeschel and Matthias Kohl. General Purpose Convolution Algorithm in S4-Classes by means of FFT. *J. Stat. Softw.*, 59:1–25, 2014.

[25] Frank Schweitzer, Giorgio Fagiolo, Didier Sornette, Fernando Vega-Redondo, Alessandro Vespignani, and Douglas R. White. Economic networks: the new challenges. *Science*, 325:422, 2009.

[26] Joseph E. Stiglitz. *Globalization and Its Discontents*. W.W. Norton & Company, 2002.

[27] Claudio J. Tessone, Antonios Garas, Beniamino Guerra, and Frank Schweitzer. How Big Is Too Big? Critical Shocks for Systemic Failure Cascades. *J. Stat. Phys.*, 151:765–783, 2013.

[28] Duncan J Watts. A Simple Model of global cascades on random networks. *Proc. Natl. Acad. Sci. U.S.A.*, 99:5766, 2002.