European Actuarial Journal (2023) 13:1–53
https://doi.org/10.1007/s13385-023-00341-9

SURVEY PAPER

Modeling and pricing cyber insurance

Idiosyncratic, systematic, and systemic risks

Kerstin Awiszus[1,2] · Thomas Knispel[1,3] · Irina Penner[1,4] · Gregor Svindland[1,5] · Alexander Voß[1,5] · Stefan Weber[1,5]

Received: 18 August 2022 / Revised: 8 December 2022 / Accepted: 22 December 2022 /
Published online: 23 January 2023

Abstract
The paper provides a comprehensive overview of modeling and pricing cyber insurance and includes clear and easily understandable explanations of the underlying mathematical concepts. We distinguish three main types of cyber risks: idiosyncratic, systematic, and systemic cyber risks. While for idiosyncratic and systematic cyber risks, classical actuarial and financial mathematics appear to be well-suited, systemic cyber risks require more sophisticated approaches that capture both network and strategic interactions. In the context of pricing cyber insurance policies, issues of interdependence arise for both systematic and systemic cyber risks; classical actuarial valuation needs to be extended to include more complex methods, such as concepts of risk-neutral valuation and (set-valued) monetary risk measures.

✉ Stefan Weber
stefan.weber@insurance.uni-hannover.de

Kerstin Awiszus
kerstin.awiszus@hs-hannover.de

Thomas Knispel
thomas.knispel@insurance.uni-hannover.de

Irina Penner
Irina.Penner@HTW-Berlin.de

Gregor Svindland
gregor.svindland@insurance.uni-hannover.de

Alexander Voß
alexander.voss@insurance.uni-hannover.de

1 House of Insurance, Leibniz Universität Hannover, Hanover, Germany

2 Hochschule Hannover, Hanover, Germany

3 Berlin School of Economics and Law, Berlin, Germany

4 HTW, University of Applied Sciences, Berlin, Germany

5 Institute of Actuarial and Financial Mathematics, Leibniz Universität Hannover, Hanover, Germany

Keywords Cyber Risks · Cyber Insurance · Idiosyncratic Risk · Systematic Risk · Systemic Risk

1 Introduction

Cyber risks constitute a major threat to companies worldwide.[1] In the last years, the estimated costs of cyber crime have continuously been increasing—from approximately USD 600 billion in 2018 to more than USD 1 trillion in 2020, cf. CSIS [25]. Consequently, the market for cyber insurance is experiencing strong growth, providing contracts that mitigate the increasing risk exposure—with significant potential ahead. However, cyber insurance differs from other lines of business in multiple ways that pose significant challenges to insurance companies offering cyber coverage:

- *Data* on cyber events and losses is scarce and typically not available in the desired amount or granularity.
- Cyber threats are evolving dynamically in a highly *non-stationary* cyber risk landscape.
- *Aggregate cyber risks* arise due to common IT architectures or complex interconnections that cannot easily be captured.
- The term 'cyber' risk itself comprises many *different types of risk* with different root causes and types of impact.

Insurance companies cannot solely rely on standard actuarial approaches when modeling and pricing cyber risks. Their traditional methods need to be complemented by novel and innovative techniques for both underwriting and quantitative risk management. The current paper provides the following main contributions:

(i) We present a *comprehensive overview of the state of the art of modeling and pricing cyber insurance*. In contrast to other surveys (see, e.g., [39]) that focus on a high-level review of the literature, we explain the underlying mathematical concepts and discuss their advantages and drawbacks.[2]

(ii) The second main contribution of the paper is a classification of cyber risks into three different types: *idiosyncratic, systematic, and systemic cyber risks*. While the distinction between idiosyncratic and systemic risks is common in the current cyber insurance literature (see, e.g., [116]), a further refinement is necessary. The three risk types can be described as follows:

   - **Idiosyncratic risks** refer to cyber risks at the level of individual policyholders that are independent from risks of others parties. This might, for example, be caused by internal errors within the company. Prototypical idiosyncratic risks are independent risks in large insurance pools that allow to apply classical actuarial techniques.

[1] For example, according to the annually published Allianz Risk Barometer (see, e.g., [2]), cyber risk ranges among the top three global business risks since 2016.

[2] Surveys that include detailed conceptual explanations are, e.g., Böhme and Schwartz [19], Marotta et al. [81], and Böhme et al.[12]. In contrast to our paper, these authors focus exclusively on game-theoretic models. We discuss this dimension in Section 3.3.

- **Systematic risks** are cyber risks that result from common vulnerabilities of entities affecting different firms at the same time, e.g., firms belonging to the same industry sector or region, or firms that utilize the same software, server, or computer system. These risks can be modeled via common risk factors. In classical actuarial and financial mathematics, systematic risks include financial market risks as well as stochastic fluctuations and evolutions of mortality rates within a population.
- **Systemic risks** are cyber risks caused by local or global contagion effects in interconnected systems or by strategic interaction. Examples are worm-type malware or supplier attacks. These risks are similar to important feedback mechanisms observed in financial crises, e.g., contagion in networks of counterparties or fire sales of stressed market participants in illiquid markets. Models include random processes with feedback, or locally and globally interacting processes. We will also include strategic interactions in this category which are studied in game theory.

Idiosyncratic and systematic cyber risks can be captured by classical approaches of actuarial and financial mathematics; systemic cyber risks require different methodologies such as epidemic network models which focus on the interconnectedness of the entities. We suggest pricing techniques that adequately incorporate interdependence for both systematic and systemic cyber risks by combining the concepts of risk-neutral valuation and risk measures.

The paper is structured as follows. Section 2 reviews classical actuarial approaches. We begin with an introduction to the frequency-severity approach in the context of cyber risk and discuss how to model both idiosyncratic and systematic risks in this framework. We explain how dependence is captured in such models. Systemic cyber risks are considered in Sect. 3. Three different modeling approaches for interconnectedness, contagion, and interaction between entities are discussed, with a special focus on their advantages and possible drawbacks. In Sect. 4, we describe pricing methods for cyber insurance contracts that are applicable in the face of idiosyncratic, systematic, and systemic risks. Section 5 discusses open questions for future research.

2 Classical actuarial approaches applied to cyber risks

The pricing of cyber insurance contracts as well as quantitative cyber risk management require sound models for the loss distributions, customized to the application purpose. While classical actuarial premium principles are essentially related to the expected claims amount (plus a safety loading), quantitative risk management particularly refers to extreme losses in the tail of the distribution and their quantification in terms of risk measures such as *Value at Risk* or *Average Value at Risk*, see Sect. 4.

In actuarial mathematics, a standard model for insurance losses—used across all lines of business— is the *frequency-severity approach*, also called *collective risk model*. For a certain time interval $[0, t]$, $t > 0$ (typically $t = 1$ year), a collective of policyholders causes a random number of claims $\mathcal{N}_t$ (*frequency*) with correspond-

ing random loss sizes $\mathcal{Y}_1, \mathcal{Y}_2, \ldots$ (*severity*) generating the total claim amount

$$\mathcal{S}_t = \sum_{j=1}^{\mathcal{N}_t} \mathcal{Y}_j, \quad t > 0.$$

Calculations within the frequency-severity approach typically rely on the following mathematical assumptions (see, e.g., [88]):

(C1) Claims occur at arrival times $0 \leq T_1 \leq T_2 \leq \ldots$. The number of claims in the time interval $[0, t]$, $t \geq 0$, is defined by

$$\mathcal{N}_t := \#\{j \geq 1 \mid T_j \leq t\},$$

i.e., $\mathcal{N} = (\mathcal{N}_t)_{t \geq 0}$ constitutes a counting process on $[0, \infty)$.

(C2) The jth claim arriving at time T_j causes the claim size $\mathcal{Y}_j$. It is assumed that the sequence $(\mathcal{Y}_j)_{j \geq 1}$ of claim sizes consists of independent and identically distributed random variables.

(C3) Claim sizes and claim numbers are assumed to be independent from each other.

In contrast to classical insurance risks, however, cyber risk is more challenging in different ways. In particular, the standard assumptions of the frequency-severity approach as well as classical statistical techniques[3] are no longer applicable:

- Claims *data* is not available in sufficient quantity or in the required granularity.
- Technology and cyber threats are evolving rapidly, i.e., the cyber environment is highly *non-stationary*.
- Cyber incidents[4] may affect different policyholders at the same time, i. e., the typical assumption of *independence* for insurance risks does not hold any longer. Moreover, there is—in contrast to natural catastrophe risk—no simple geographical delimitation of dependent risks.

Nonetheless, the frequency-severity approach can be customized to account for cyber risk—at least as a first approximation and for certain types of non-systemic cyber risks, which can be subdivided into idiosyncratic and systematic risks (as defined in Sect. 1). In the frequency-severity approaches presented below, we explicitly distinguish between techniques suitable for modeling idiosyncratic or systematic incidents. In the context of cyber insurance, however, a third class of risks can be identified, namely systemic risks, i.e., cyber risks resulting from contagion between interconnected entities. Proper modeling of such risks goes beyond the classical framework of actuarial modeling and requires appropriate models for networks, (cyber) disease spread, and strategic interaction. Hence, we discuss the modeling of systemic cyber risks separately in Sect. 3, while the pricing for all types of cyber risks is discussed in Sect. 4.

[3] For details on statistical techniques in classical actuarial models, see Sects. 2.1.3 and 2.2.3.

[4] According to NIST [91], a cyber incident can be defined as: "Actions taken through the use of an information system or network that result in an actual or potentially adverse effect on an information system, network, and/or the information residing therein." We will also use the term cyber attacks interchangeably in this paper.

To present frequency-severity approaches in the context of cyber risk in a unified and practically applicable way, we use the following notation and definitions. We consider an insurer's portfolio of n policyholders (firms) exposed to the considered type of cyber risk incidents. Each firm admits an individual risk profile characterized by a vector of covariates, e.g., *industry sector*, *size*, *IT security level*, which are elicitable, for example, via a questionnaire or from public information. Using the covariates, the insurer's portfolio is decomposed into homogeneous groups, labeled $\{1, \ldots, K\}$, with covariates vector x^k for group k. We denote by $n_k, k = 1, \ldots, K$, the number of firms in group k, i.e., $n_1 + \ldots + n_K = n$. For pricing purposes, these homogeneous groups can be viewed as tariff cells, i.e., the insurance firm should charge all firms[5] within group k the same premium π_k. In particular, if n_k is large, then the premium of the idiosyncratic cyber risk can be derived from the law of large numbers as the expected claims amount per firm of group k plus a suitable safety loading to avoid ruin in the long run.

Both idiosyncratic and systematic incidents can be grouped into different cyber risk categories, labeled $\{1, \ldots, C\}$. Categories may include, for example, *data breach*, *fraud*, and *business interruption*. Two exemplary actuarial classification approaches are sketched and discussed in Appendix A. Cyber risk is modeled per risk category $c \in \{1, \ldots, C\}$ and per group $k \in \{1, \ldots, K\}$. A pair $m := (c, k)$ is called a *cyber risk module*. The total number of modules $C \cdot K$ is a trade-off between homogeneity and availability of data for statistical estimation.

Within this framework, we model the losses for an insurance company – for each cyber risk module as well as on an aggregate level. For this purpose, we first focus on frequency-severity based approaches to modeling cyber risks in the spirit of the classical collective risk model. Second, we add dependence to our cyber risk model in order to capture accumulation risks. Note that appropriate dependence modeling is particularly important for calculating capital requirements in quantitative risk management, since the underlying risk measures refer to events in the extreme tail of the loss distribution.

2.1 Frequency and severity

A frequency-severity model may be applied on the level of each cyber risk module $m = (c, k)$. For simplicity, we describe the losses per risk category of individual firms by a collective risk model. This can be justified as follows: Since all firms in any group are (approximately) homogeneous, they will be charged the same premium for any given risk category. From the point of view of the insurance company, only aggregate losses are relevant, i.e., an artificial allocation of losses to individual companies for pricing purposes will produce the correct implications. We thus describe the losses per risk category *at the level of any individual firm* by a collective risk model with the same severity as the corresponding module, but with a suitably reduced frequency.

[5] For simplicity, we assume that the firms within a group possess the same types of exposures that are of the same size. This can be generalized by introducing suitable volume measures that characterize the size of exposures.

For a firm i in group k and a fixed risk category c, i.e., a cyber risk module $m = (c, k)$, we consider the frequency and severity model $(\mathcal{N}_t^{m,i}, (\mathcal{Y}_j^{m,i})_{j\geq 1})$. Then the total claim amount of firm i up to time t can easily be obtained by summing up:

$$\mathcal{S}_t^{m,i} = \sum_{j=1}^{\mathcal{N}_t^{m,i}} \mathcal{Y}_j^{m,i}.$$

In mathematical terms, all quantities correspond to random variables on a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{P}$ plays the role of the statistical measure that models the relative frequency with which events occur.

As outlined in the introduction of this section, one of the most common assumptions in the frequency-severity model is assumption (C3), i. e., claim numbers and sizes are independent of each other. This assumption facilitates and simplifies many calculations regarding the compound total claim amount process. In particular, the expected total claim amount and its variance follow from *Wald's formulas:*

$$\begin{aligned}\mathbb{E}[\mathcal{S}_t^{m,i}] &= \mathbb{E}[\mathcal{N}_t^{m,i}] \cdot \mathbb{E}[\mathcal{Y}_1^{m,i}],\\ \mathrm{Var}(\mathcal{S}_t^{m,i}) &= \mathbb{E}[\mathcal{N}_t^{m,i}]\mathrm{Var}(\mathcal{Y}_1^{m,i}) + \mathrm{Var}(\mathcal{N}_t^{m,i})(\mathbb{E}[\mathcal{Y}_1^{m,i}])^2.\end{aligned}$$

However, the independence assumption may not always be reasonable—e.g., if hidden factors influence both frequency and severity: Sun et al. [106] detect a positive nonlinear dependence between frequency and severity in hacking breach risks at the firm level. A firm with a strong cyber self protection is expected to experience both fewer and weaker hacking attacks than companies with weak self protection mechanisms. In mathematical terms, the authors capture this dependence between frequency and severity by the Gumbel copula, see also Sect. 2.2.

2.1.1 Frequency

Let $\mathcal{N}_t^{m,i}$ denote the number of incidents in module $m = (c, k)$ until time t that are allocated to a firm i in group k, and let $(\mathcal{N}_t^{m,i})_{t\geq 0}$ denote the corresponding counting process. At the aggregate level,

$$\mathcal{N}_t^{m,agg} := \sum_{i=1}^{n_k} \mathcal{N}_t^{m,i} \text{ and } \mathcal{N}_t^{(c)} := \sum_{k=1}^{K} \mathcal{N}_t^{m,agg}, \quad t \geq 0,$$

will count the total number of incidents per module $m = (c, k)$ and the total number of incidents per cyber risk category c, respectively.

Poisson Process

A simple counting process for incidents—reflecting non-stationarity of cyber risk—is a *time-inhomogeneous* Poisson process with intensity function λ^m per firm for cyber risk module m.

Definition 2.1 (Time-inhomogeneous Poisson process) A counting process $(\mathcal{N}_t)_{t\geq 0}$ is called a time-inhomogeneous Poisson process on $(\Omega, \mathcal{F}, \mathbb{P})$ with locally integrable rate (or intensity) function $\lambda : [0, \infty) \to [0, \infty)$ if:

1. $\mathcal{N}_0 = 0$,
2. the process has independent increments,
3. for any time interval $(s, t]$, the number of incidents is Poisson distributed with mean $\int_s^t \lambda(u)\, du$, i.e.,

$$\mathcal{N}_t - \mathcal{N}_s \sim \text{Poiss}\left(\int_s^t \lambda(u)\, du\right).$$

Unless the intensity function is constant, the increments of a time-inhomogeneous Poisson process are *non-stationary*. The cumulative rate function $\int_0^t \lambda(u)\, du$ corresponds to the expected number of incidents up to time t.

Zeller and Scherer [116] adopt this approach for idiosyncratic incidents. For each policyholder i of group k and module $m = (c, k)$, the number of idiosyncratic incidents $(\mathcal{N}_t^{m,i})_{t\geq 0}$ is assumed to follow a time-inhomogeneous Poisson process with intensity $\lambda^m = \lambda^{(c,k)}$. Clearly, for each cyber risk category c, the intensity at the level of an individual firm i depends on the covariates x^k of group k (but not on the individual policyholder i), and Zeller and Scherer [116] propose a *generalized additive model*

$$\lambda^{(c,k)}(t) = \exp(f^c(x^k) + g^c(t))$$

to estimate the intensity rates.[6] In particular, similarities and deviations of the risk profiles of the K groups—expressed in terms of the covariate vectors x^k, $k = 1, \ldots, K$—are reflected by the intensity functions $\lambda^{(c,k)}$.

Since *idiosyncratic* incidents are independent across firms, the total number of incidents $\mathcal{N}_t^{m,agg}$, $t \geq 0$, per module $m = (c, k)$ as well as the total number of incidents $\mathcal{N}_t^{(c)}$, $t \geq 0$, per cyber risk category c, respectively, are again time-inhomogeneous Poisson processes with respective intensities

$$\lambda^{m,agg}(t) = n_k \lambda^{(c,k)}(t), \quad \lambda^{(c)}(t) = \sum_{k=1}^{K} n_k \lambda^{(c,k)}(t), \quad t \geq 0. \tag{1}$$

More delicate, however, is the case of *systematic* cyber risk incidents. In particular, frequency distributions of different policyholders might be subject to dependencies due to joint underlying cyber risk factors $\mathcal{R}^1, \ldots, \mathcal{R}^d$, representing, for example, the random discovery of exploits in commonly used software, improvements in cyber security, or the technological progress of tools for cyber attacks.

Cox Process

Such dependencies between counting processes can be captured in the context of *Cox processes*, also called *doubly stochastic Poisson processes*, extending the notion of a time-inhomogeneous Poisson process to a random intensity.

[6] The auxiliary function f additively maps the covariates, while g captures the time dependence.

Definition 2.2 (Cox process) A Cox process $(\mathcal{N}_t)_{t\geq 0}$ is a counting process described by a random intensity process $(\lambda_t)_{t\geq 0}$ such that conditional on the specific realization $t \mapsto \lambda_t(\omega)$, $\omega \in \Omega$, the process $(\mathcal{N}_t)_{t\geq 0}$ is a time-inhomogeneous Poisson process with intensity $t \mapsto \lambda(t) = \lambda_t(\omega)$.

A reasonable assumption could be that the intensity is a function of the current state of random cyber risk factors, i.e., for an $\mathbb{R}^d$-valued stochastic process $\mathcal{R}_t = (\mathcal{R}_t^1, \dots, \mathcal{R}_t^d)$, $t \geq 0$, of cyber risk factors and a function $\lambda : \mathbb{R}^d \to [0, \infty)$, the intensity process is defined as

$$\lambda_t(\omega) = \lambda(\mathcal{R}_t(\omega)), \quad t \geq 0,\ \omega \in \Omega.$$

More generally, the intensity process could be modeled as a function of the whole history of cyber risk factors, i.e.,

$$\lambda_t(\omega) = \lambda(\mathcal{R}_u(\omega) : u \leq t), \quad t \geq 0,\ \omega \in \Omega.$$

In summary, in the case of systematic cyber risk, a reasonable model for the number of incidents $\mathcal{N}_t^{m,i}$ up to time t allocated to policyholder i in group k for module $m = (c, k)$ could be to assume that $(\mathcal{N}_t^{m,i})_{t\geq 0}$ follows a Cox process with intensity process $\lambda_t^m = \lambda^m(\mathcal{R}_t)$, $t \geq 0$, defined in terms of a suitable function $\lambda^m : \mathbb{R}^d \to [0, \infty)$, such that conditional on the cyber risk factors $t \mapsto \mathcal{R}_t(\omega) = (\mathcal{R}_t^1(\omega), \dots, \mathcal{R}_t^d(\omega))$ the counting processes $(\mathcal{N}_t^{m,i})_{t\geq 0}$, $m = (c, k)$, $c = 1, \dots, C$, $k = 1, \dots, K$, are independent time-inhomogeneous Poisson processes. In particular, conditional independence implies that—conditional on the specific realization $t \mapsto \lambda_t^m(\omega)$—the total number of incidents $\mathcal{N}_t^{m,agg}$, $t \geq 0$, per module $m = (c, k)$ and the total number of incidents $\mathcal{N}_t^{(c)}$, $t \geq 0$, per cyber risk category c are again time-inhomogeneous Poisson processes with intensities

$$\lambda_t^{m,agg} = n_k \lambda_t^{(c,k)}, \quad \lambda_t^{(c)} = \sum_{k=1}^{K} n_k \lambda_t^{(c,k)}, \quad t \geq 0,$$

in analogy to (1).

In contrast to the time-inhomogeneous Poisson process, the increments of a Cox process $(\mathcal{N}_t)_{t\geq 0}$ are in general no longer independent, but subject to autocorrelation. More precisely, for any $s < t \leq u < v$, the tower property of conditional expectation implies

$$\operatorname{Cov}(\mathcal{N}_t - \mathcal{N}_s, \mathcal{N}_v - \mathcal{N}_u) = \operatorname{Cov}\left(\int_s^t \lambda_z \, dz, \int_u^v \lambda_z \, dz \right),$$

i.e., the autocorrelation depends on the random intensity process. The statistical analysis of Bessy–Roland et al. [7] yields empirical evidence for autocorrelation in the number of attacks, and thus provides an additional rationale for Cox processes when modeling claims frequency. The specification of an intensity process that reproduces the empirically observed autocorrelation appears to be challenging.

2.1.2 Severity

Every claim occurring in the frequency-severity model triggers a loss size that is modeled as a random variable. We let $\mathcal{Y}_j^{m,i}$ denote the claim size of the jth event allocated to firm i for module $m = (c, k)$ and assume that $(\mathcal{Y}_j^{m,i})_{j\geq 1}, i = 1, \ldots, n_k$, is a collection of non-negative independent[7] and identically distributed random variables. One among many different possible approaches is to assume that the key governing parameter for the choice of the claim size distribution is the incident category c; characteristics of group k then determine distributional details, e.g., parameter values.

Due to the limited availability of loss data, empirical research on cyber risk severity distributions has mostly focused on the category of data breaches. For this category, open source data bases, such as the Privacy Rights Clearinghouse Chronology of Data Breaches, are available and regularly updated. Data breach severities are found to follow strongly heavy-tailed distributions such as power-law (see, e.g., [80]), log-normal (see, e.g., [37]) or generalized Pareto distributions (GPD) (see, e.g., [112] or [106]). For cyber risk categories different from data breaches, less data is publicly available. Consequently, fewer papers have appeared that empirically analyze the respective severity distributions.

An exception is Dacorogna et al. [29] who study a non-public database of the *French Gendarmerie Nationale* on cyber complaints and describe a process for cleaning the data. Their analysis suggests that losses are heavy-tailed. Dacorogna et al. [30] refine the analysis and provide a tool for classifying attacks based on the fatness of the tail. Another promising direction are studies based on data on operational risk such as Biener et al. [9] or Eling and Wirfs [41]. These approaches offer the benefit of being able to analyze all categories of cyber incidents simultaneously. In particular, Eling and Wirfs [41] detect distributional differences between small and large claim sizes for all considered cyber incident categories. The authors propose a *composite distribution approach*, where excess losses over a threshold are modeled using a GPD and the remaining smaller losses are modeled using a simple parametric distribution such as a gamma or log-normal distribution. In general, composite distribution approaches constitute a flexible modeling tool to take the empirically observed distributional differences between body and tail of severity distributions adequately into account. A composite distribution approach can be formalized as follows.

For each module m, we choose a threshold θ^m distinguishing small from large cyber claims. Small and large claims, i.e., the body and tail of the severity distribution, are then modeled separately: The i.i.d. claim sizes follow a composite distribution with density

$$f_{\mathcal{Y}_j^m}(y) := \begin{cases} C_1^m \cdot f_{\text{small}}^m(y), & \text{if } -\infty < y \leq \theta^m, \\ C_2^m \cdot f_{\text{large}}^m(y), & \text{if } \theta^m < y < \infty, \end{cases}$$

[7] Cyber event claim sizes in a certain time interval may not always be independent, e.g., due to commonly used cyber security measures. The resulting dependence structures could be captured by alternatively imposing *conditional* independence assumptions given a set of joint underlying risk factors—similar to the conceptual idea underlying Cox processes that we already discussed above.

where f^m_{small}, f^m_{large} are probability density functions modeling the sizes of small and large claims in module m, respectively, and C^m_1, C^m_2 are normalizing constants that are additionally constrained by continuity conditions at the threshold θ^m. Depending on the characteristics of the module m, different choices for f^m_{small}, f^m_{large} may be suitable. Examples include

- **Small Claims:** PERT, Normal, Gamma, Log-Normal, GPD, Kernel Distribution
- **Large Claims:** GPD

The composite distribution approach is well-suited for modeling non-life insurance severity distributions in general, and cyber risks in particular.[8] As discussed here, the methodology is independent of time, i.e., it provides only a snapshot of the current cyber environment. In the light of the fast-evolving, non-stationary cyber landscape, the suitability of the model must, however, be regularly validated and updated. For further details and discussions, we refer the interested reader to the excellent summaries provided by Zeller and Scherer [116], Sect. 2.1, or Eling [39], in particular Tables 4 and 6, and to Cooray and Ananda [23] for an application of composite distributions in a non-cyber specific context.

2.1.3 On calibration and application

In general, frequency-severity models are well-understood, easy to implement and to calibrate if a sufficient amount of data is available. They are also straightforward to explain, for example, to an executive board of an insurance company; this is partly due to their prevalence in actuarial modeling. For frequency modeling, intensities can, e.g., be fit to data using generalized additive models (as in Zeller and Scherer [116] and described above), maximum- or marginal likelihood, or Bayesian methods. Cox processes are generally more difficult to estimate – the choice of a calibration method critically depends on the law of the underlying common risk factor processes.[9]

For the statistical analysis of the severity, there exist well-known estimation techniques including maximum-likelihood, see, e.g., Maillart and Sornette [80] or Edwards et al. [37] for applications in a cyber severity context, or the peaks-over-threshold method for fitting a GPD to the tail of a distribution, see, e.g., McNeil et al. [85] and Embrechts et al. [42]. For a general review on methods for the parameter estimation of GPDs, including maximum-likelihood, the method of moments, the probability weighted moments method, and Bayesian approaches, see also de Zea Bermudez and Kotz [32] and de Zea Bermudez and Kotz [33].

The practical application of frequency-severity models to cyber risk is challenging, in particular due to the limited amount of available data and its insufficient quality. Moreover, Poisson and Cox processes do not capture the systemic interaction between different (groups of) policyholders; see also Reinhart [98] for a discussion of the

[8] Sun et al. [106] also suggest a composite distribution approach for modeling malicious hacking data breach risk. The tail of their distributions follows a GPD, and the distribution body is modeled using a non-parametric kernel distribution. Due to both its suitability and flexibility, a similar approach is also incorporated in the cyber risk model of Zeller and Scherer [116].

[9] For details on the statistical estimation of point processes and theoretical background see, e.g., Daley and Vere-Jones [31].

frequency-severity model presented by Zeller and Scherer [116]. An alternative are Hawkes processes that incorporate systemic self-excitation into frequency models, see Sect. 3.1. Like Cox processes, Hawkes processes are able to capture autocorrelation observable in the data.

2.2 Dependence modeling

The distribution of the total claim amount per module and at the portfolio level is affected by the underlying dependence structures. For cyber risk, dependencies may be present at different levels including:

- dependence between frequency distributions or between severity distributions of different policyholders in the same homogeneous group (e.g., due to the random evolution of common cyber security measures and cyber threats over time),
- dependence between frequency and severity—in contrast to the classical framework of frequency-severity models (e.g., due to unobservable random factors within a tariff class such as heterogeneous levels of cyber self protection).

One approach to deal with the first type of dependencies are Cox processes as described in Sect. 2.1.1. In this section, we review further approaches to model dependence in the context of cyber risk that have been proposed in the literature.

2.2.1 Common risk factors

Common risk factors capture dependence for systematic risks; the factors are random quantities to which all risks are jointly exposed. Common risk factors appear in static as well as in dynamic models and have been widely used in the cyber risk modeling literature. For example, they are key elements of the cyber risk models proposed by Böhme [10], Böhme and Kataria [11] and Zeller and Scherer [116]. Cox processes, as introduced in Sect. 2.1.1, are an example of dynamic factor models.

Böhme [10] captures dependence using one common risk factor in a static model. The factor represents a common vulnerability in a portfolio of n individual risks. The connection between individual risks and the latent risk factor is studied on the basis of linear correlation.[10] Common risk factors also appear in the cyber risk model of [116]. The authors use marked point processes with two-dimensional marks: the first component describes the strength of an attack, and the second component represents the subset of companies affected. Dependence among firms occurs due to the restriction of incidents to certain industry sectors which is modeled via a common risk factor. The paper suggests a conceptual framework, but does not yet calibrate the model to real data.

[10] Linear correlation, defined as $\rho(X, Y) = \mathrm{Cov}(X, Y)/\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)} \in [-1, 1]$, captures a possible linear relationship between the random variables X and Y. The maximum and minimum values of 1 and -1 are not always attainable. While often used to impose ad-hoc dependence assumptions in practice (in a cyber context, see, e.g., [10], [11]), linear correlation suffers from many well-known fallacies, see, e.g., [85] for a detailed discussion.

2.2.2 Copulas

In actuarial applications, copulas are a standard tool that fully characterizes the dependence structure of the components of finite-dimensional random vectors. A d-dimensional *copula* $\mathcal{C} : [0, 1]^d \to [0, 1]$ is the distribution function of a d-dimensional random vector with uniform one-dimensional marginal distributions.

Theorem 2.1 (Sklar's Theorem)

1. *For any d-dimensional distribution function F with margins $F_1, \ldots, F_d$ there exists a copula $\mathcal{C}$ with*

$$F(x_1, \ldots, x_d) = \mathcal{C}(F_1(x_1), \ldots, F_d(x_d)) \quad \text{for all } x_1, \ldots, x_d \in [-\infty, \infty]. \quad (2)$$

   If all F_i are continuous, then $\mathcal{C}$ is unique.
2. *Conversely, for a given copula $\mathcal{C}$ and given one-dimensional distribution functions $F_1, \ldots, F_d$, the function F in (2) is a d-dimensional distribution function with copula $\mathcal{C}$ and marginal distribution functions $F_1, \ldots, F_d$.*

Property 1 states that a copula extracts the dependence structure of a random vector from its multivariate distribution, while property 2 provides a flexible construction principle of multivariate models by combing marginal distributions and copulas to multivariate distributions. Prominent examples of copulas are:

- **Gaussian copula:** Letting Φ^{-1} be the quantile function of the standard normal distribution and Φ_Σ the joint cumulative distribution function of a multivariate normal distribution with covariance matrix Σ, the corresponding Gaussian copula is given by

$$\mathcal{C}^{\text{Ga}}_\Sigma(u_1, \ldots, u_d) = \Phi_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)) \quad ((u_1, \ldots, u_d) \in [0, 1]).$$

- **t-copula:** Let $t_{\nu,\Sigma}$ signify the distribution function of a d-dimensional t-distribution $t_d(\nu, 0, \Sigma)$ for a given correlation matrix Σ and with ν degrees of freedom, and let t_ν denote the distribution function of a univariate t-distribution with ν degrees of freedom. The corresponding t-copula takes the form

$$\mathcal{C}^t_{\nu,\Sigma}(u_1, \ldots, u_d) = t_{\nu,\Sigma}(t_\nu^{-1}((u_1), \ldots, t_\nu^{-1}(u_d)) \quad ((u_1, \ldots, u_d) \in [0, 1]).$$

  Like the Gaussian copula, the t-copula is an implicit copula that is extracted from a given parametric multi-variate distribution.
- **Archimedean copulas**: Explicit copulas are constructed from given functions; the prime example are Archimedean copulas. We consider a suitable continuous function $\psi : [0, \infty) \to [0, 1]$ with $\psi(0) = 1$, $\lim_{x\to\infty} \psi(x) = 0$, and ψ strictly decreasing on $[0, \psi^{-1}(0)]$, where ψ^{-1} denotes its generalized inverse. The Archimedean copula with generator ψ is given by

$$\mathcal{C}^{\text{Ar}}_\psi(u_1, \ldots, u_d) = \psi^{-1}(\psi(u_1) + \ldots + \psi(u_d)) \quad ((u_1, \ldots, u_d) \in [0, 1]).$$

A special case is the *Gumbel copula* for $\psi_\theta(s) = (-\ln(s))^\theta$, $\theta \in [1, \infty)$ that is applied in the cyber model of Sun et al. [106].

2.2.3 On calibration and application

Common risk factor models are able to capture dependence from bottom-up and are widely used in economics. From a practical perspective, they are particularly useful when a modeler is confident that random outcomes are influenced by common external factors. In Cox processes, described in Sect. 2.1.1, the common factors enter the model via the intensity. Their estimation depends on the specific choice of the distribution of the underlying risk factors. For the class of *linear* factor models, a large amount of statistical estimation methods exist. Important techniques are time series regression, cross-sectional regression (at each time point), and principal component analysis, see, e.g., McNeil et al. [85] and the references therein.

Another approach are copulas; these are theoretically able to represent every form of static dependence. They can be viewed as a top-down approach that imposes a dependence structure without modeling the underlying mechanisms, as contrasted with factor models that can be interpreted as a bottom-up approach. Copulas have already been used in the literature on cyber risk. Herath and Herath [62] model the loss distribution *at a single firm* using a copula that captures the dependence structure between the number of affected computers of the firm and the overall severity of the loss. In Böhme and Kataria [11] dependence *between different firms* is captured using a t-copula with a given linear correlation coefficient.

Another example is an application of copulas in a modified collective risk model in which the standard independence assumption is relaxed. For the incident category c of hacking data breaches, Sun et al. [106] observe upper tail dependence between frequency and severity. This may be caused by hidden factors such as the degree of cyber self protection. They propose to model this dependence for any firm i in module m up to time t via a Gumbel copula.

Eling and Jung [40] and Liu et al. [79] apply vine copulas in the context of data breaches. Vine copulas are very flexible, and their calibration is quite tractable, since high-dimensional dependence structures are decomposed into components of lower dimension. For detailed information on vine copulas we refer to Czado [26], Czado and Nagler [27], and an online collection of material on vine copulas, see TU Munich, Statistics Research Group [107].

In general, the choice of a suitable copula estimation method depends on the structure of the chosen copula model: parametric, semiparametric or nonparametric. A good survey on various methods is Hofert et al. [65]. In a fully parametric model, both the copula and the marginal distributions are completely characterized by (vector) parameters. The maximum likelihood (ML) method can be applied to the dependence and the marginal part either jointly or sequentially. The sequential approach is often referred to as the method of inference functions for margins (IFM), see, e.g., the surveys in Choroś et al. [22], Sect. 2.1, or McNeil et al. [85], Sect. 7.6. Semiparametric approaches typically still involve a parametric copula model, but a nonparametric model for the marginals. Here, classically, the marginal distributions are estimated via their empirical distribution functions. Estimation of the full model can then be

performed using a maximum-pseudo likelihood approach, in which the nonparametric marginal estimators are inserted, see the seminal paper of Genest et al. [53]. This approach is considered to be more robust than the parametric ML and IFM methods in many practical applications, see Kim et al. [70], unless substantial information is available on a parametric class to which the margins belong to. Nonparametric copula models may be estimated on the basis of different variations of nonparametric marginal and joint distribution function estimates, see, e.g., the seminal paper of Deheuvels [34] using empirical distribution functions or Chen and Huang [21] (and the references therein) for kernel-based estimators of the copula (or copula density).

3 Systemic cyber risks

Systemic risk generally refers to the possibility that distortions in a system may spread across many entities and be augmented due to local or global feedback effects. This is in contrast to systematic risk that introduces dependence via exogenous factors. Systemic risk refers to the internal mechanism of a system in which the behavior of the various entities has a sequential impact. It is often associated with a cascading risk propagation such that

> "in case of an adverse local shock (infection) to a system of interconnected entities, a substantial part of the system, or even the whole system, finally becomes infected due to contagion effects."[11]

As a consequence of the 2008 financial crisis, systemic risk was intensively studied in systems of interdependent financial institutions, see, e.g., Staum [105]. This concept is also important in the context of cyber risk, since agents and organizations in cyber systems are interconnected, for example within IT networks or via business contacts.[12] The relevance of systemic cyber threats has been emphasized by leading regulatory and macroprudential institutions, cf. WEF [110] and ESRB [46]. Examples of contagious threats include the WannaCry and NotPetya cyber attacks where the corresponding malware spread through networks of interconnected IT devices and firms, causing tremendous losses to cyber systems worldwide.[13]

Modeling systemic cyber risks requires models of feedback effects, local and global interaction, as well as strategic interaction. We describe three concrete methodological approaches (see Fig. 1): Firstly, self-excitation of cyber incidents can be captured by *Hawkes processes* on an aggregate level (Sect. 3.1); in this respect, Hawkes processes can be interpreted as a top-down approach. Secondly, *epidemic network models* (Sect. 3.2) capture the interconnectedness and cascading propagation of risks; this bottom-up approach may focus on local connections, but can also capture global interaction via aggregate, mean-field quantities. Both approaches can be viewed as mechanistic interaction models in which rational or strategic behavior of agents is typically not mirrored. This is the focus of the third approach, *game-theoretic models*

[11] See [35].

[12] See, e.g., the discussion in Sect. 2 of Welburn and Strong [111].

[13] For further information and a detailed risk analysis see [46].

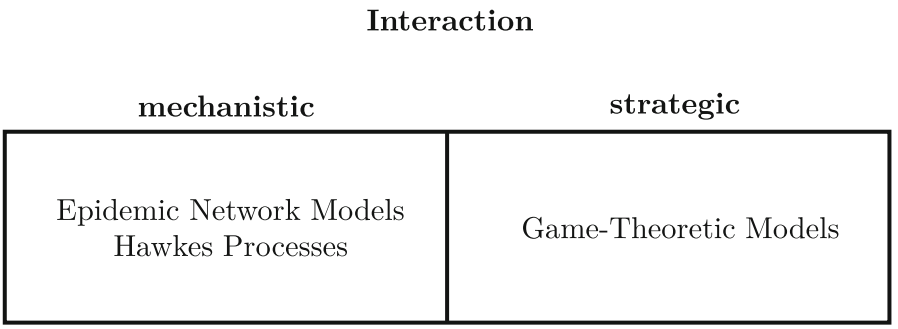

Fig. 1 Interaction in models of systemic cyber risks

(Sect. 3.3). These study explicitly the *strategic* interaction of interconnected entities, usually under strongly simplified connectivity assumptions; notions of equilibria typically characterize the solutions.

3.1 Hawkes processes

Systematic dependence of cyber incidents can be modeled by Cox processes; these permit to capture empirical features such as the autocorrelation of cyber attacks. Cox processes focus on common factors, but they do not model contagion in interconnected systems. An alternative are *Hawkes processes*, self-exciting processes, that mirror feedback effects, a specific form of *systemic* cyber risk; they also capture the stylized fact of autocorrelation of the number of events.

Definition 3.1 (Hawkes process) A one-dimensional Hawkes process $(\mathcal{N}_t)_{t\geq 0}$ is a point process with jump times $T_1, T_2, \ldots$ and with random intensity $t \mapsto \lambda_t$, given by

$$\lambda_t = \mu(t) + \sum_{T_n \leq t} \varphi(t - T_n) = \mu(t) + \int_{[0,t)} \varphi(t-u)\, d\mathcal{N}_u,$$

where $\mu(\cdot)$ is a baseline intensity of jumps, and where φ is the excitation function or kernel function resp. which expresses the positive influence of past incidents at time T_n on the current value of the intensity.

From a conceptual point of view, Hawkes processes allow to capture—besides autocorrelation of the number of cyber risk incidents—excitation effects, by coupling the arrival rate of events with the number of past incidents. In particular, this allows modeling systemic incidents that affect a very large number of counterparties at the same time, e.g., the spread of worm-type malware.

Self-excitation of cyber incidents for each policyholder as well as the excitation between policyholders of different groups can be modeled by a multivariate Hawkes model. More precisely, for all cyber risk modules $m = (c, k)$ and for any policyholder

i of group k, the intensity of the counting process $(\mathcal{N}_t^{m,i})_{t\geq 0}$ takes the form

$$\lambda_t^{(c,k,i)} = \mu^{(c,k)}(t) + \sum_{l=1}^{K} \sum_{j=1}^{n_l} \sum_{T_n^{(c,l,j)} \leq t} \varphi_{i,j}^{c,k,l}(t - T_n^{(c,l,j)}),$$

where

- $t \mapsto \mu^{(c,k)}(t)$ is the deterministic base intensity function, depending on the cyber risk module $m = (c, k)$ only,
- $t \mapsto \varphi_{i,j}^{c,k,l}(t)$ are self- and mutually-exciting maps (called kernels), depending on both the cyber risk module $m = (c, k)$, the other group l and the individual policyholders i, j,
- and $T_n^{(c,l,j)}$, $n \in \mathbb{N}$, are the claims arrival times of policyholder j in group l with respect to the cyber risk category c.

In this multivariate Hawkes model, the kernels $\varphi_{i,i}^{c,k,k}$ describe the self-excitation for policyholder i of group k, while the $\varphi_{i,j}^{c,k,l}$ for different policyholders $i \neq j$ model contagion between policyholders and across groups.

3.1.1 On calibration and application

Using suitable parametric functions for both the baseline intensity and the kernels of Hawkes processes can in principle be estimated by maximum-likelihood methods—provided that data is available in the desired amount and granularity. Data availability is, of course, still a major challenge in cyber insurance. Model calibration and statistical parameter estimates in a cyber context are, e.g., presented in Bessy-Roland et al. [7] focusing on data breaches. Further, Hawkes processes are also used in an empirical study of cyber risk contagion in Baldwin et al. [4]. In the context of financial data, maximum-likelihood methods and graphical goodness-of-fit are, e.g., discussed in Embrechts et al. [43]. Da Fonseca and Zaatour [28] develop an estimation by the method of moments which is fast compared to likelihood estimation. A general discussion including Bayesian estimation is presented in Daley and Vere-Jones [31], see also Giesecke [54], Errais et al. [45], and Aït-Sahalia et al. [1].

Since Hawkes processes can be easily incorporated with a classical actuarial frequency model for systemic cyber risk, they can be integrated into the standard collective risk model if complemented by an appropriate severity modeling approach. In principle, the severities of systemic events could be chosen as described in Sect. 2.1.2 for idiosyncratic and systematic events. Due to the limited amount of data and uncertainty about the possible impact of future systemic cyber incidents, accurate modeling of systemic severities is extremely challenging in practice.

Hawkes processes take a top-down approach to modeling systemic cyber risk and neglect the specific infection processes that underlie risk contagion in interconnected systems. Important aspects of risk amplification and possible accumulation scenarios may not be adequately captured. This is the main attractive feature of epidemic network models; their disadvantage is their increased complexity.

3.2 Epidemic network models

Interconnectedness constitutes a key characteristic of cyber systems. Systemic cyber risks may spread and amplify in networks of interconnected companies, economic actors, or financial institutions. Cyber network models for contagious risk propagation consist of the following three key components:

1. A **network** (also called *graph*) whose nodes represent components or agents. These entities could be individual corporations, subsystems of computers, or single devices. The edges of the network correspond to possible transition channels, e.g., IT connections or exchange of data/computer code, see Sect. 3.2.1;
2. A **spread process** on the network that models the propagation of a contagious cyber risk, like the spread of a computer virus, a Trojan, or ransomware,[14] see Sect. 3.2.2;
3. A **loss model** which determines the severity of cyber events and the monetary impact on different agents in the network, see Sect. 3.2.3.

3.2.1 Networks

Definition 3.2 (Network) A *network*[15] (or *graph*) G is an ordered pair of sets $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} \neq \emptyset$ is a countable set of N elements, called *nodes* (or *vertices*), and $\mathcal{E}$ is a set of pairs (i, j), $i, j \in \mathcal{V}$, of different nodes, called *edges* (or *links*). If all edges in $\mathcal{E}$ are unordered, formally, $(i, j) \in \mathcal{E} \Rightarrow (j, i) \in \mathcal{E}$, then G is called an *undirected network*. Otherwise, the network G is called *directed*.

The network structure is encoded in its *adjacency matrix* $A = (a_{ij})_{i,j\in\{1,\dots,N\}} \in \{0, 1\}^{N\times N}$, which is defined by its entries

$$a_{ij} := \begin{cases} 1, & \text{if } (i, j) \in \mathcal{E} \\ 0, & \text{if } (i, j) \notin \mathcal{E}. \end{cases}$$

By definition, G is undirected if and only if A is symmetric. Examples of undirected network topologies with $N = 8$ nodes are depicted in Fig. 2.

In applied network analysis, the exact network structure is often unknown. In this case, random network models enable sampling from a class of networks with given fixed topological characteristics (such as the overall number of nodes).[16]

In the cyber insurance literature, network models are mainly applied to study risk contagion, e.g., modeling the propagation of malware in IT networks of interconnected firms or devices. In addition to an underlying network, an appropriate model of the contagion process that captures epidemic spread is needed.

[14] Moreover, contagion can also be interpreted in a broader sense, e.g., considering the propagation of business interruptions or the breakdown of supply chains as the consequence of cyber attacks on single entities.

[15] This definition refers to unweighted networks. In the context of weighted networks, the notion of undirected networks refers to the symmetry of the weight matrices.

[16] Two commonly used models are discussed in Appendix B.

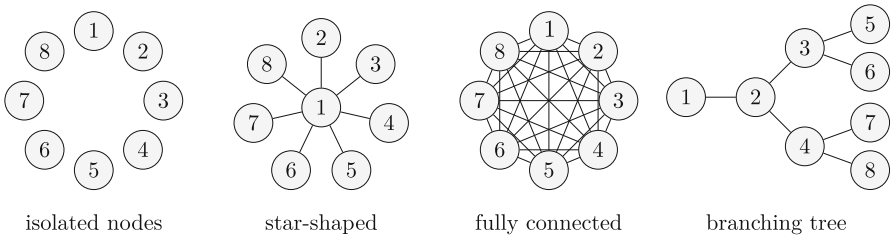

Fig. 2 Examples of network topologies with $N = 8$ nodes

3.2.2 Epidemic spread processes

Models of infectious disease spread dynamics have been studied extensively in mathematical biology and epidemiology, dating back at least to the seminal work of Kermack and McKendrick [68].[17] In this paper, we focus on epidemic *network* models for populations of entities.

At any point in time, each node is in a particular state, which may change over time as it interacts with other nodes. According to their state, individuals are divided into various *compartments*, e.g., individuals that are *susceptible* (S) to an infection, *infected* (I) individuals, or individuals who have *recovered* (R) from the infection. For a network of N nodes, the spread process at time t can be described by a *state vector*

$$X(t) = (X_1(t), \ldots, X_N(t)) \in E^N,$$

where E is the set of compartments. Both *Markov* and *non-Markov* processes have been considered in the context of epidemic spread processes.[18]

Markovian Spread Models

In Markovian spread models on networks, the evolution of the state vector $X(t)$ is described by a (in many cases: time-homogeneous) continuous-time Markov chain on the discrete state space E^N. The Markov models SIS *(Susceptible-Infected-Susceptible)* and SIR *(Susceptible-Infected-Recovered)* form a class of commonly used models for epidemic propagation in networks. They are distinguished by the presence (SIR) or absence (SIS) of immunity: Reinfection events are only possible in the SIS framework because in the SIR model recovered individuals acquire (permanent) immunity, i.e., the models are based on the two different sets of compartments $E = \{S, I\}$ and $E = \{S, I, R\}$.

In both models, a transition of X from one state in E^N to another is possible only if exactly one node changes its state X_i in E. State changes may occur through infection or recovery: It is assumed that each node may be infected by its infected neighbors, but can be cured independently of all other nodes in the network. Each node is endowed

[17] The models typically focus either on an epidemic spread within a population, as, e.g., in Kermack and McKendrick [68], or on the spread along paths of a predefined network; for a detailed overview, see, e.g., Pastor-Satorras et al. [96] and Kiss et al. [72].

[18] The Markov property captures that a process is "memoryless", i.e., that the conditional distribution of future values X_{t+s}, $s > 0$, of the process does only depend on the present value of the process X_t and not additionally on past values X_μ, $\mu < t$.

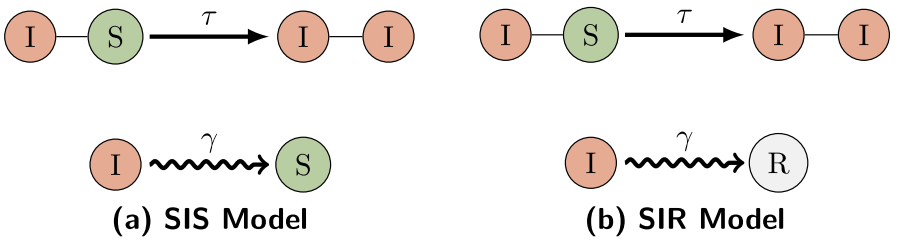

Fig. 3 Infection and recovery for the SIS and SIR network model

with an independent exponential clock and changes its state when the exponential clock rings. Letting $\tau > 0$ and $\gamma > 0$, the rates of these transitions are illustrated in Fig. 3 and given as follows ($i = 1, \ldots, N$):

$$\begin{aligned} X_i : S \to I \quad &\text{with rate} \quad \tau \sum_{j=1}^{N} a_{ij} \mathbb{1}_{\{X_j(t)=I\}} \\ X_i : I \to Z \quad &\text{with rate} \quad \gamma, \end{aligned} \tag{3}$$

where $Z = S$, for the SIS, and $Z = R$ for the SIR model, respectively.

The exponential transition times enable an intuitive stochastic simulation algorithm: the well-known *Gillespie algorithm*, first introduced in Gillespie [55] and Gillespie [56]; see Appendix C for details.

For practical purposes such as the pricing of cyber insurance contracts, we often do not need the full information provided by the Markov chain evolution, but only the dynamics of specific quantities such as moments or (infection) probabilities. Of particular interest are the dynamics of the state probabilities of individual nodes $\mathbb{P}(X_i(t) = x_i)$, $t \geq 0$. They can be derived from Kolmogorov's forward equation and written in general form as ($i = 1, \ldots, N$)

$$\frac{d\mathbb{P}(X_i(t) = x_i)}{dt} = \sum_{y : y_i = x_i} \sum_{z \neq y} [\mathbb{P}(X(t) = z) q_{zy} - \mathbb{P}(X(t) = y) q_{yz}], \tag{4}$$

where q_{zy} denotes the transition rate of the entire process X from $z \to y$. In natural sciences, this equation is also known under the term *master equation*. For the SIS and SIR models, using Bernoulli random variables $S_i(t) := \mathbb{1}_{\{X_i(t)=S\}}$, $I_i(t) := \mathbb{1}_{\{X_i(t)=I\}}$, and (for SIR) $R_i(t) := \mathbb{1}_{\{X_i(t)=R\}}$, the dynamics of state probabilities of individual nodes (4) can conveniently be written via moments:

- **SIS model:**[19] Since $E = \{I, S\}$, we have $S_i(t) = 1 - I_i(t)$, i.e., the evolution of X is fully described by the evolution of the vector $I(t) = (I_1(t), \ldots, I_N(t))$, and

[19] In the cyber insurance literature, the SIS Markov model was used by Fahrenwaldt et al. [47]. Also, a brief application was studied in Xu and Hua [114] with a modified ε-SIS model, originally proposed in Van Mieghem and Cator [86]. Here, an infectious threat for node i from outside the network is included with a rate ε_i.

the single node infection dynamics for $i = 1, \ldots, N$ are given by

$$\frac{d\mathbb{E}[I_i(t)]}{dt} = -\gamma\mathbb{E}[I_i(t)] + \tau\sum_{j=1}^{N} a_{ij}\mathbb{E}[I_j(t)] - \tau\sum_{j=1}^{N} a_{ij}\mathbb{E}[I_i(t)I_j(t)], \tag{5}$$

since $\mathbb{P}(X_i(t) = I) = \mathbb{P}(I_i(t) = 1) = \mathbb{E}[I_i(t))]$. This system of N equations is not closed as second order moments $\mathbb{E}[I_i(t)I_j(t)]$, i.e., second order infection probabilities, appear.

- **SIR model:** The dynamics of the recovery Bernoulli random variable $R_i(t)$ result from the dynamics of $I_i(t)$ and $S_i(t)$ due to $\mathbb{E}[R_i(t)] = 1 - \mathbb{E}[S_i(t)] - \mathbb{E}[I_i(t)]$. Equation (4) corresponds to:

$$\begin{aligned}\frac{d\mathbb{E}[S_i(t)]}{dt} &= -\tau\sum_{j=1}^{N} a_{ij}\mathbb{E}[S_i(t)I_j(t)],\\ \frac{d\mathbb{E}[I_i(t)]}{dt} &= \tau\sum_{j=1}^{N} a_{ij}\mathbb{E}[S_i(t)I_j(t)] - \gamma\mathbb{E}[I_i(t)],\end{aligned} \tag{6}$$

for $i = 1, 2, \ldots, N$. Again, the system is not closed due to the presence of second order moments.

The main problem with systems (5) and (6) is the fact that they are *not closed*: They depend on second order moments, which, in turn, depend on third order moments, etc. For example, the fully closed SIS model yields $\sum_{i=1}^{N}\binom{N}{i} = 2^N - 1$ moment (i.e., infection probability) equations. Solving these systems exactly becomes intractable for networks of realistic size. To deal with this issue, the following two approximation approaches have been proposed:

1. **Monte Carlo simulation:** Monte Carlo simulation using the Gillespie algorithm (see Appendix C) constitutes a powerful tool to obtain various quantity estimates related to the evolution of the epidemic spread. In particular, this includes the state probability dynamics of individual nodes (4).[20]
2. **Moment closures:** If a set of nodes $J \subset \mathcal{V}$ is infected, this increases the probability of other nodes in the network (that are connected to the set J via an existing path) to become infected as well. Node states do not evolve independently and are to some extent *correlated*. To break the cascade of equations and to make ODE systems tractable, the moment closure approach consists in factorizing moments beyond a certain order k, substituting all higher-order moments. This is done by considering the exact moment equations up to this order k and *closing* the system by approximating moments of order $k + 1$ in terms of products of lower-order moments using a mean-field function. A detailed description of two different types of moment closures is provided in Appendix D. However, a major problem with

[20] Pseudocode and further explanations of the Gillespie algorithm applied to the SIS and SIR epidemic network models is, e.g., given in Appendix A.1.1 of Kiss et al. [72].

moment closures is that only little is known about rigorous error estimates.[21] This presents an important avenue for future research.

Non-Markovian Spread Models

Non-Markovian models possess conditional distributions that may depend on the past and on further random factors. In contrast to the Markovian setup, where transition times are necessarily exponential, non-Markovian models might allow additional flexibility to freely choose the distributions of infection and recovery times. In addition, dependence among the infection times may be included. This generality may improve the quality of a fit to real-world data. However, the extended generality in comparison to Markov models is typically associated with reduced tractability. For this reason, non-Markovian models are less commonly considered. In addition, a similar scope of flexibility can also be achieved within the class of Markovian models by extending the dimension of the state space; but this comes again at the price of increased complexity and possibly reduced tractability.

A simple example of a non-Markovian model for the spread of cyber risks has been proposed by Xu and Hua [114]. The model does not include immunity, i.e., the underlying compartment set is the same as for the Markovian SIS model. The considered waiting times in the model are:

- The individual **recovery times** T_i^{recov} of *infected* nodes.
- For nodes i which are in the *susceptible* state, two different types of infections are considered, *internal infections* from within the network and *external infections* coming from outside:
  1. **Internal infection times:** Let the random variable $K_i(t) = \sum_{j=1}^{N} a_{ij} I_j(t)$ denote the number of infected neighbors of node i at time t. Infectious transmissions to node i are given with waiting times $T_{i_1}, \ldots, T_{i_{K_i}}$. These times share the same marginal distribution F_i. Their underlying dependence structure is captured by a prespecified copula.
  2. **External infection times:** A random variable T_i^{out} with distribution G_i models the arrival time of threats from outside the network to node i. T_i^{out} is assumed to be independent of times $T_{i_1}, \ldots, T_{i_{K_i}}$.

To simulate the process, the waiting times for all nodes are generated according to their current state (i.e., recovery times for all infected nodes, and internal and external infection times for all susceptible nodes). The minimum of these waiting times determines the next event (infection or recovery). After this change, all quantities are recomputed and the process is repeated until a prespecified stopping criterion is met.[22]

Finally, note that a Markovian SIS model with outside infections[23] can be obtained as a special case by choosing exponentially distributed infection and recovery times and assuming independence between all waiting times.

[21] This problem has also been highlighted in the epidemic literature, see, e.g., Kiss et al. [72], p.115.

[22] Pseudocode for stochastic simulations is provided in Algorithm 1 of Xu and Hua [114].

[23] To be precise, the so-called ε-SIS model, originally proposed in Van Mieghem and Cator [86], arises.

3.2.3 Loss models

Given the underlying network, and the epidemic spread process X on it, the third and final ingredient of a cyber risk network model is given by a suitable loss model $Y_{i,j}$ for each node $i = 1, \ldots, N$, where j describes the number of loss events. In the existing literature, loss models are kept rather simple as the focus lies on modeling the cyber-epidemic spread. We give two examples:

1. In Fahrenwaldt et al. [47], cyber attacks are launched in a two-step procedure: First, using a random process, times of attacks on the entire network (loss events) $t_1, t_2, \ldots$ are generated. Second, for each node i, a possible random loss $L_{i,j}$ is modeled, where j describes the index of the corresponding attack time. The loss, however, only materializes if node i is infected at the attack time. This is captured by the loss model

$$Y_{i,j} = L_{i,j} \cdot \mathbb{1}_{X_i(t_j)=I}, \quad i = 1, \ldots N, \quad j = 1, 2, \ldots.$$

2. In Xu and Hua [114], the loss model $Y_{i,j}$ is given by

$$Y_{i,j} = \eta_i(D_{i,j}) + C_i(T_{i,j}^{recov}), \quad i = 1, \ldots N, \quad j = 1, 2, \ldots$$

with a legal cost function η_i, the number $D_{i,j}$ of data damaged in the infection j, and a cost function C_i depending on the recovery time $T_{i,j}^{recov}$ of node i for infection event j. Here, the recovery time $T_{i,j}^{recov}$ for each event j is obtained from the infection dynamics, while the data loss sizes $D_{i,j}$ are assumed to follow a beta distribution.

3.2.4 On calibration and application

Epidemic network models in the cyber insurance literature mostly focus on a general assessment of the underlying structure of systemic cyber risks: aspects of risk contagion and propagation are characterized in a *qualitative* sense. For example, Fahrenwaldt et al. [47] study the effect of homogeneous, star-shaped, and clustered topologies on the resulting overall insurance losses in regular networks, demonstrating the strong impact of the network topology on risk propagation. Further, epidemic network models could also be applied to identify critical initial infection locations or critical network links that may augment cyber losses. The models are thus particularly useful for counterfactual simulations and have not yet been calibrated to real-world data.

More applications of epidemic network models to cyber risk contagion can be found in the engineering literature. However, these works do not study risk emergence on a global level. Instead, they analyze cyber risks which are building from the microstructure of interconnected IT devices in *local* environments. For example, Powell [97] focuses on local IT authentication procedures, where the corresponding vectors of lateral movements within a network can be interpreted as edges of a directed mathematical graph. Possible attack vectors are evaluated using classical metrics from

network theory and epidemic spreading models of SIR type. More technical and IT-related aspects of cyber security issues in smart grids, i.e., networked power systems for energy production, distribution, and consumption, are surveyed and discussed in Wang and Lu [108].

However, for the *quantitative* assessment of systemic cyber risk from a regulatory or actuarial perspective, contagion among different companies needs to be studied on a *global* scale. A major challenge for accurate modeling is the estimation of the exact network structure and the epidemic parameters of past and future incidents—particularly due to data limitations and the speed of technological evolution. In Appendix E, we provide a brief overview and classification of existing estimation approaches for epidemic network models that are not necessarily related to cyber; in our view, however, it is conceivable that such approaches could also be implemented and further developed in a cyber context in future cyber risk research.

To overcome the estimation challenge, *top-down approaches* have been proposed in the literature. In Hillairet and Lopez [63], the impact of massive global-scale cyber-incidents, like the WannaCry scenario, on insurance losses and assistance services is determined. While network contagion is implicitly considered, it is not modeled within an actual network framework; instead, the authors choose the original population-based SIR model of Kermack and McKendrick [68] which describes deterministic dynamics of the total numbers of susceptible, infected, and recovered individuals within the global population of IT devices. The corresponding ODE system is given by

$$\begin{aligned}\frac{dS(t)}{dt} &= -\tau S(t)I(t)\\ \frac{dI(t)}{dt} &= \tau S(t)I(t) - \gamma I(t)\\ \frac{dR(t)}{dt} &= \gamma I(t)\end{aligned}$$

with constant global population size $N = S(t) + I(t) + R(t)$. Parameters N, τ, and γ are estimated from data of the WannaCry cyber incident.

Given this global spread, the focus of the paper lies on the *stochastic* evolution of the insurer's *local* portfolio consisting of $n << N$ policyholders and their corresponding losses. The influence of the global cyber epidemic on the local portfolio is captured by the hazard rate $\lambda_{T_i^{infec}}$ of the policyholders' infection times T_i^{infec}:

$$\lambda_{T_i^{infec}}(t) := \lim_{dt\to 0^+} \frac{1}{dt}\mathbb{P}(T_i^{infec} \in [t, t+dt] \mid T_i^{infec} \geq t) = \tau I(t),$$

i.e., the local hazard rates are assumed to be proportional to the number of infected individuals in the global population.

Most recently, this model has further been extended by replacing the homogeneous global population model with a network scenario of interconnected industry sectors, see Hillairet et al. [64]. The underlying directed and weighted network structure is derived from OECD data that measures the economic flow between different industries,

and this data is interpreted as a reasonable estimate of the digital dependence between these sectors. Contagion between sectors is modeled using a deterministic multi-group SIR model for the total numbers of susceptible, infected, and recovered companies in the sectors. Due to the scarcity of data currently available, such top-down approaches present promising avenues for risk management and actuarial modeling.

Additionally, future research should analyze the implementation of more realistic loss models, that, e.g., contain different types of cyber events and capture their characteristic severity distributions (see also the discussion on classical frequency-severity approaches in Sect. 2.1.2). This would further strengthen the applicability of network models in practice.

3.3 Game-theoretic models and strategic interaction effects

In addition to contagion due to the interconnectedness of entities in cyber networks, potentially different objectives of the actors and their strategic interaction constitute a key characteristic of systemic cyber risk. The risk exposure of individuals is often interdependent, since it is influenced by the behavior of other actors. *Game theory* provides a suitable framework to study this dimension in the cyber ecosystem.

In the first part of this section, we briefly review and provide a short mathematical introduction to game theoretic approaches applied to study cyber risk and cyber insurance (Sect. 3.3.1). For an exhaustive review of the corresponding literature, we refer to the surveys Böhme and Schartz [19], Böhme et al. [12], and Marotta et al. [81]. We will adopt the notation from Marotta et al. [81]. Sect. 3.3.2 evaluates the considered models.

3.3.1 Game theoretic modeling approaches

The majority of game theoretic contributions focuses on self protection of interdependent actors in the presence or the absence of cyber insurance. A key question is whether and under which conditions cyber insurance provides incentives for self protection and improves global IT security. In this section, we present[24] the main ideas and characteristics of such models.

Three Different Types of Actors in the Game

We consider three types of strategic players with different objectives: potential buyers of insurance (for simplicity, called agents), insurance companies, and the regulator.

1. **Agents** are the potential cyber insurance policyholders. To capture interdependence, most models assume that agents form a network. Agent i aims to maximize her expected utility

$$\max \mathbb{E}[U_i(W_i)],$$

where

- U_i denotes the utility function of agent i. Various types of utility functions are considered in the literature; most of them satisfy the classical von-Neumann-Morgenstern axioms. While some papers, such as Naghizadeh and Liu [90],

[24] We refer to Marotta et al. [81] for an in-depth overview of the topic.

Pal [93], and Pal et al. [94], allow for *heterogeneous preferences*, the majority of models assumes *homogeneous preferences*, i.e., $U_i = U$ across all agents.

- W_i is the financial position of agent i at the end of the insurance period. The value W_i depends on whether the agent has bought an insurance contract or not, on her investment C_i in cyber security, and on potential losses L_i in case the agent is affected by a cyber attack.

The agent's **self protection level** x_i is a *crucial model component* when studying interdependence.[25] Most of the existing literature falls into either of the following two distinct categories: Some assume that only two security states are possible, secured or not, with the corresponding constant cost C or 0. Others propose a continuous scale of security levels, e.g., $x_i \in [0, 1]$. The value of x_i affects

- *the cost of self protection* C_i:
  For a continuous spectrum of security levels, i.e., $x_i \in [0, 1]$, $C_i = C(x_i)$ is typically assumed to be an increasing convex function of x_i, reflecting that user costs rapidly increase when improving security.
- *agent* i*'s probability of becoming infected* $p_i := \mathbb{P}(I_i = 1)$:
  Obviously, this probability depends on the individual security level x_i of the agent i, but—due to interdependence—it may also be influenced by the individual security levels of other network participants.

Within this framework, agent i's expected utility can be computed

(a) **without insurance:**

$$\mathbb{E}[U_i(W_i)] = (1 - p_i) \cdot U_i(W_i^0 - C_i) + p_i \cdot U_i(W_i^0 - L_i - C_i)$$

(b) **with insurance:**

$$\mathbb{E}[U_i(W_i)] = (1 - p_i) \cdot U_i(W_i^0 - \pi_i - C_i) + p_i \cdot U_i(W_i^0 - L_i - C_i - \pi_i + \hat{L}_i)$$

where

- W_i^0 denotes the initial wealth of agent i.
- π_i is the insurance premium of agent i set by the insurer. This premium depends on the type of insurance market; we will discuss different models below.
- L_i is the potential loss of agent i that is governed by a binary distribution: only two possible scenarios are considered. Either the agent experiences a cyber attack with a *fixed loss size*, or she is not attacked which corresponds to no loss. This particular setting excludes the possibility of different types of cyber attacks. Multiple attacks are also not considered.[26] The majority of game theoretic models relies on the assumption of constant homogeneous losses for all agents, i.e., $L_i \equiv L$.

[25] Only few papers, e.g. Böhme [10], Böhme and Kataria [11], Johnson et al. [67] and Johnson et al. [66], do not include self protection in the model.

[26] We will discuss the scope of the existing models in Sect. 3.3.2.

- $\hat{L}_i$ is the cover in case of loss which is specified in the insurance contract. Most papers assume full coverage, i.e., $\hat{L}_i = L_i$, but some consider alternatively partial coverage, e.g., in order to mitigate the impact of information asymmetries, cf. Mazzoccoli and Naldi [84], Pal [93], Pal et al. [94].

2. **Insurance Companies**: The insurer sets the cyber insurance premiums and specifies the insurance cover $\hat{L}_i$. Insurance premiums depend on the market structure:
   - **Competitive market**: This is the prevailing model in the literature. The profits of the insurers are zero in this case; customers pay fair premiums. Competitive markets are a boundary case that almost surely leads to the insurer's ruin in the long run.
   - **Monopolistic market / Representative insurer**: Another extreme is a market with only one insurance company. In these models, the impact of a monopoly can be studied. An alternative consists in studying objective functions that are different from the insurer's profit. This situation is mostly studied in the context of regulation: The insurer represents a regulatory authority and is not aiming for profit maximization, but focuses on the wealth distribution in order to incentivize a certain standard of IT protection.[27]
   - **Immature market/Oligopoly**: Instead of a monopoly, imperfect competition is studied with multiple insurers that may earn profits. The increments between the fair price and the insurance premium is determined by the markets structure.[28]
3. **Regulator:** Market inefficiencies and a lack of cyber security may be mitigated by regulatory policies. *Regulatory instruments* include mandatory insurance, fines and rebates, liability for contagion, etc. The choice of policies and their impact can be studied[29] by introducing a third party, the regulator. The objective of the regulator is to maximize a *social welfare function*. This could, for example, be chosen as the sum of the expected utilities of the agents

$$\sum_i \mathbb{E}[U_i(W_i)].$$

Interdependent Self Protection in IT Networks

The strategic interaction of the three types of players introduced above is modeled as a game. The agents form an interconnected network and optimize their expected utility. Their individual security level and the amount of cyber insurance coverage serve as their controls. The insurance companies are provider of risk management solutions. In some models, a regulator is included as a third party with the aim to improve welfare, e.g., by implementing standards of protection in cyber systems.

[27] Market models of this type are studied in Naghizadeh and Liu [90] with a zero-profit insurer. Profits are still possible in Pal [93], Pal et al. [94] and Pal et al. [95], and maximized in Khalili et al. [69].

[28] Immature markets are considered, e.g., in Martinelli et al. [82], Martinelli and Yautsiukhin [83], Ogut et al. [92].

[29] The effects of such regulatory instruments were, e.g., studied in Bolot and Lelarge [15], Naghizadeh and Liu [90], Pal [93], Pal et al. [94].

The network topologies are, typically, quite stylized to guarantee tractability. For example, two-agent models are considered in Ogut et al. [92]. Most papers investigate complete graphs, e.g., Ogut et al. [92], Schwartz and Sastry [100] and Pal et al. [94]. Bolot and Lelarge [15] and Yang and Lui [115], in contrast, investigate networks with degree heterogeneity, but restrict their analysis to Erdős-Rényi random graphs.

Agents are interdependent in the network, since the infection probability p_i depends on the local security level x_i and levels of the other nodes $y_i := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N)$ (or at least of i's neighbors). In some cases, p_i is assumed to depend on an overall network security level as well.[30] However, in contrast to the models from Sect. 3.2, attacks do not result from a dynamic contagion process; instead, the infection is assumed to be *static* and the values p_i are derived from *ad hoc schemes*. The most common one[31] assumes a continuous spectrum of security levels and computes p_i as the complementary probability of the case that neither a direct nor an indirect attack occurs:

$$\begin{aligned} p_i(x_i, y_i) &= 1 - (1 - p_i^{dir})(1 - p_i^{cont}) \\ &= 1 - (1 - \psi_i(x_i)) \times \prod_{j \neq i} (1 - h_{i,j}\psi_j(x_j)) \end{aligned}$$

where

- $p_i^{dir} = \psi_i(x_i)$ denotes the probability of direct infection of i through threats from outside the network. It is interpreted as a function of the individual security level x_i.
- $p_i^{cont} = 1 - \prod_{j \neq i}(1 - h_{i,j}\psi_j(x_j))$ is the probability for node i to become infected through contagion. The probability for i to be infected via node j is given by $h_{i,j}$, i.e., $h_{i,j} \neq 0$ only if i and j are adjacent. This is where the underlying network topology comes into play.

In the absence of information asymmetries, the game between agents and the insurer(s) involves three perspectives:[32]

1. A legal framework is set by the regulator (if a regulator is present).
2. Agents specify their levels of self protection and insurance protection and select from the available contract types to maximize their expected profits.
3. Insurance companies compute the corresponding contract details, i.e., premiums π_i and indemnities $\hat{L}_i$. In absence of information asymmetries between agents and the insurer(s), the protection levels of policyholders can be observed by the insurer and are reflected by the contract.

The model may be augmented to incorporate information asymmetries:

[30] This is the case in Shetty et al. [103], Shetty et al. [102] and Schwartz and Sastry [100].

[31] An alternative approach using a simplified two-state scenario of security investments is analyzed in Bolot and Lelarge [13], Bolot and Lelarge [14], and Yang and Lui [115]. Infection probabilities are derived from a recursive branching process.

[32] Variations of the game design are possible; e.g., in Laszka et al. [77] the authors use a signaling game instead of a game model to study the adverse selection problem, allowing insurers to audit the agents' security. A similar game is considered in Khalili et al. [69] who introduce a pre-screening procedure.

- **Moral hazard**: A dishonest policyholder may behave in a way that increases the risk, if the insurer cannot properly monitor the policyholder's behavior. In the game, this is represented by the possibility for agents to modify their self protection[33] levels.
- **Adverse selection**: Agents with larger risks have a higher demand for insurance than safer ones. The degree of the policyholders' risk tolerances cannot be observed by the insurer. The self protection levels of policyholders is not precisely known by the insurer when the contract details are computed.

In most papers, cyber insurance is not associated with additional incentives to enhance self protection. In contrast, agents may prefer to buy insurance instead of investments in self protection, i.e., from a welfare perspective, they *underinvest* in security. These observations may be interpreted as an indication that *regulatory interventions* are necessary, such as fines and rebates, mandatory cyber insurance, or minimal investment levels.[34]

3.3.2 On calibration and application

Many questions remain to be answered in future research, since the existing game theoretic models of cyber insurance and cyber security are oversimplified—thus, not yet fully applicable to real-world data:

- **Simplified network topologies**: In the vast majority of the discussed literature, networks are assumed to be homogeneous. However, agents are typically heterogeneous in reality which substantially alters the cyber ecosystem. Network contagion and cyber loss accumulation are highly sensitive to the topological network arrangement; for example, important determinants are the presence (or the absence) of central hub nodes or clustering effects, see, e.g., Fahrenwaldt et al. [47]. For appropriate risk measurement and management these aspects need to be taken into account explicitly.
- **Static contagion**: A key feature of cyber risk in networks is the systemic amplification of disturbances. From the insurer's perspective, the contagion dynamics will clearly influence tail risks; an example are catastrophic incidents that affect a large fraction of its portfolio. Such events may be critical in terms of the insurer's solvency. An understanding of cyber losses and an evaluation of countermeasures requires dynamic models of contagion processes.

[33] Some authors distinguish between *self insurance*—a reduction in the size of a loss—and *self protection*—a reduction in the probability of a loss, as suggested by Ehrlich and Becker [38]. While such a distinction may be intuitive in models with simple loss distributions or in frequency-severity models, it is sometimes more appropriate to model loss exposure by random variables or distributions and analyze action on that level. Safety measures often influence loss sizes and probabilities together. How useful the distinction of Ehrlich and Becker [38] is, depends on the modeling framework chosen and the particular application. The term *self protection* in this paper refers to any activity to reduce physical risk—including both the size of losses and their probabilities.

[34] The effect of fines and rebates was studied in papers [15], [90], [93], and [94]. In the presence of information asymmetries, fines and rebates cannot easily be applied. An alternative regulatory instrument are requirements on minimal investment levels for IT security. However, Shetty et al. [103], Shetty et al. [102] and Schwartz et al. [101] argue against such requirements.

- **Constant losses**: In all considered game-theoretic models, the agent's losses are assumed to be constant, i.e., modeled as binary random variables. However, in reality we observe that the severity of instances varies substantially due to the heterogeneity of cyber events, ranging from mild losses (e.g. malfunctioning of email accounts) to very large losses (e.g. attacks on production facilities, or breakdowns of systems).

Cyber insurance and instruments to control cyber risk depend on the structures of networks, the dynamics of epidemic spread processes, as well as loss models—and vice versa. These feedback loops need to be properly incorporated in future research. Key ingredients of systemic cyber risks—the interconnectedness captured by epidemic network models, and strategic interaction described in game-theoretic models—must be combined.

4 Pricing cyber insurance

Cyber risk comprises both *non-systemic risk*, further subdivided into *idiosyncratic* and *systematic cyber risk*, cf. Sect. 2, and *systemic risk*, cf. Sect. 3. Classical actuarial pricing, however, relies on the *principle of pooling*, and it is thus applicable for idiosyncratic cyber risks only. For systematic and systemic cyber risk, the appropriate pricing of insurance contracts requires more sophisticated concepts and techniques. A discussion of current industry practice for pricing cyber risks can be found in Romanosky et al. [99]. However, the described approaches do not yet cover the full complexity of cyber risk such that further (scientific) efforts are necessary. In this section, we explain and suggest suitable pricing techniques[35] tailored to the three different components of cyber risk.

4.1 Pricing of non-systemic cyber risks

In non-life insurance, contracts are usually signed for 1 year. At renewal time, the insurer may adjust premium charges as well as terms and conditions, while the policyholder can decide whether or not to continue the contract. Premium calculation thus typically refers to loss distributions on a one-year time horizon. In this section, we adopt this market convention and consider premiums payable annually in advance.[36]

In this chapter, we are concerned with a general pricing approach, and we do not restrict ourselves to frequency-severity models. We do, however, adopt some of the notations presented earlier. As introduced in Sect. 2, losses and associated premiums

[35] Another challenge is the insurability of (systemic) cyber risks. Many insurers report that they limit their exposure to this line of business due to a lack of data and models. At the same time, a comparison of rough estimates of supply and potential demand reveals a significant gap in cyber insurance. This is detrimental to agents who are exposed to the risk and do not receive insurance coverage. But insurance companies are also missing out on potentially large business opportunities. In addition to the problems with data and models, however, there is also a fundamental question about the insurability of cyber risk in light of systemic risk. A structured pricing methodology can provide a realistic assessment.

[36] For simplicity, we assume that interest rates are zero, or alternatively that insurance claims are already discounted.

are considered in the granularity of cyber risk categories $c \in \{1, \ldots, C\}$ and homogeneous groups $k \in \{1, \ldots, K\}$ of policyholders. Each pair $m = (c, k)$ is called a cyber risk module. In terms of a modular system, the premium per risk category serves as a component for the overall premium. Homogeneous groups—specified for example in terms of covariates—correspond to tariff cells, i.e., any policyholder in group k should pay the same premium $\pi^{m,\text{non-sys}}$ per risk category c. We denote by n_k the number of policyholders in group k and assume that volumes and distributions of risks within a group are identical. Although adopting the previously introduced notation, we do not necessarily consider a frequency-severity approach, but discuss pricing strategies that may also be applied in a more general framework. The methodology is inspired by Wüthrich et al. [113] and Knispel et al. [74].

To decouple the pricing of idiosyncratic and systematic cyber losses in the absence of systemic risk, one possible approach is to construct a decomposition of the total non-systemic claims amount $\mathcal{S}_1^{m,\text{non-sys}}$ on a 1-year time horizon. This decomposition takes the form

$$\mathcal{S}_1^{m,\text{non-sys}} = \mathcal{S}_1^{m,\text{idio}} + \mathcal{S}_1^{m,\text{systematic}} \tag{7}$$

where the total systematic claims equal $\mathcal{S}_1^{m,\text{systematic}}$ and the term $\mathcal{S}_1^{m,\text{idio}}$ denotes the total idiosyncratic fluctuations around the systematic claims. We explain below how a premium can be computed per risk group. Finally, a smoothing algorithm might be helpful in order to avoid structural breaks between the premiums of risk groups with similar covariates. The terms $\mathcal{S}_1^{m,\text{idio}}$ and $\mathcal{S}_1^{m,\text{systematic}}$ are unique only up to a constant that may be subtracted of one of the terms and added to the other.

In order to obtain a decomposition (7), we consider the σ-algebra $\mathcal{F}$ that encodes the systematic information. This is, for example, the information that is generated by observing the underlying exogenous stochastic factors. The full information at the time horizon of one year is jointly generated by the σ-algebra $\mathcal{F}$ and idiosyncratic fluctuations, also called technical risks, sometimes explicitly encoded by another σ-algebra $\mathcal{T}$. A decomposition (7) can be obtained by setting

$$\begin{aligned} \mathcal{S}_1^{m,\text{systematic}} &= \mathbb{E}_{\mathbb{P}}\left[\mathcal{S}_1^{m,\text{non-sys}} | \mathcal{F}\right] - \text{const}, \\ \mathcal{S}_1^{m,\text{idio}} &= \mathcal{S}_1^{m,\text{non-sys}} - \mathcal{S}_1^{m,\text{systematic}} + \text{const}, \end{aligned}$$

where conditional expectations are computed under the statistical measure $\mathbb{P}$ and where the constant const can freely be chosen. If $\mathcal{S}_1^{m,\text{non-sys}}$ is square-integrable and const $= 0$, Eq. (7) corresponds to an orthogonal decomposition in the Hilbert space of square-integrable random variables. An adjustment of the constant might be desirable for allocating the total premium for non-systemic risk to its two components.

Pricing Idiosyncratic Risk

As a special case, we consider the case that idiosyncratic cyber risks in a portfolio of individual claims are conditionally independent given $\mathcal{F}$. For homogeneous groups of policyholders, defined in terms of covariates vectors x^k, $k \in \{1, \ldots, K\}$, this type of risk is thus subject to pooling of risk, and hence a conditional version of classical actuarial pricing is still applicable. A valuation based on $\mathcal{F}$-conditional means with

respect to the statistical or real-world measure $\mathbb{P}$ is mathematically justified by a conditional version of the strong law of large numbers.

More precisely, for each firm i with cyber risk module $m = (c, k)$, annual losses[37] $\hat{\mathcal{S}}_1^{m,i}$, $i \in \mathbb{N}$, are identically distributed given $\mathcal{F}$, and we suppose that the increments

$$\varepsilon^{m,i} := \hat{\mathcal{S}}_1^{m,i} - \mathbb{E}_{\mathbb{P}}[\hat{\mathcal{S}}_1^{m,i}|\mathcal{F}], \quad i \in \mathbb{N},$$

are conditionally independent given $\mathcal{F}$. Then the average claims amount tends asymptotically[38] to the conditionally expected claims amount per policyholder:

$$\lim_{n_k \uparrow \infty} \frac{1}{n_k} \sum_{i=1}^{n_k} \hat{\mathcal{S}}_1^{m,i} = \mathbb{E}_{\mathbb{P}}[\hat{\mathcal{S}}_1^{m,1}|\mathcal{F}] \quad \mathbb{P}\text{-a.s.}$$

When setting const $= 0$, this is exactly the systematic claims amount for any firm i according to decomposition (7), suggesting that any premiums per policyholder for idiosyncratic cyber fluctuations should—for a large number of policyholders n_k in group k—be equal to zero and only[39] the systematic part should be priced. This is analogous to the *net risk premium* in the unconditional setting. But the net risk premium is known to be insufficient! Indeed, in a multi-period model, ruin theory states that ruin of the insurer occurs—no matter of the initial capital—in the long run $\mathbb{P}$-a.s. if only the net risk premium is charged, see, e.g., Mikosch [88] and the references therein.

A related result already holds in the one-period setting: Suppose that the premium charged from each firm admits the perfect replication of the systematic part $\mathbb{E}_{\mathbb{P}}[\hat{\mathcal{S}}_1^{m,1}|\mathcal{F}]$ (e.g., in a complete financial market). Letting the number of policyholders n_k tend to infinity, the $\mathcal{F}$-conditional one-period loss probability

$$\mathbb{P}\left(n_k \mathbb{E}_{\mathbb{P}}[\hat{\mathcal{S}}_1^{m,1}|\mathcal{F}] - \sum_{i=1}^{n_k} \hat{\mathcal{S}}_1^{m,i} < 0 \,\middle|\, \mathcal{F} \right) = \mathbb{P}\left(\sum_{i=1}^{n_k} \frac{\hat{\mathcal{S}}_1^{m,i} - \mathbb{E}_{\mathbb{P}}[\hat{\mathcal{S}}_1^{m,1}|\mathcal{F}]]}{\sqrt{\mathrm{Var}_{\mathbb{P}}[\hat{\mathcal{S}}_1^{m,1}|\mathcal{F}]}} > 0 \,\middle|\, \mathcal{F} \right)$$

converges to 50%, due to the central limit theorem. To stay on the safe side, a safety loading is necessary in addition to the net risk premium.

In our specific construction, the idiosyncratic part $\mathcal{S}_1^{m,\mathrm{idio}}$ for firm i in (7) equals $\varepsilon^{m,i}$ possessing expectation zero; but in alternative decompositions, a non-zero expectation corresponding to a non-zero net risk premium for the idiosyncratic part would also be admissible. This is an issue of premium allocation between idiosyncratic and systematic cyber risk only, but does not affect the total premium for non-systemic cyber risk, if cash-additive premium principles are used to price idiosyncratic risks.

Classical actuarial premium principles provide explicit safety loadings in a transparent manner, based on the first two moments of the loss distribution:

[37] Since we are not assuming a frequency-severity model as in Sect. 2.1, we introduce a slightly different notation to indicate that we are not generally referring to the specific setting discussed in Sect. 2.1.

[38] We recall that n_k denotes the number of policyholders in group k.

[39] This relies on the specific choice const $= 0$. When setting const $= 0$, idiosyncratic fluctuations will be both positive and negative. From a technical point of view, this does not cause any complications. However, if one needs to require that $\mathcal{S}_1^{m,\mathrm{idio}} \geq 0$, the constant should be suitably adjusted.

- **Variance principle:** $\pi^{m,\text{idio}} = \mathbb{E}[\mathcal{S}_1^{m,\text{idio}}] + a\,\text{Var}(\mathcal{S}_1^{m,\text{idio}})$ with $a > 0$,
- **Standard deviation principle:** $\pi^{m,\text{idio}} = \mathbb{E}[\mathcal{S}_1^{m,\text{idio}}] + a\sqrt{\text{Var}(\mathcal{S}_1^{m,\text{idio}})}$ with $a > 0$.

The safety loading parameter a can be chosen for each cyber risk module $m = (c, k)$ separately, for example, depending on the specific loss distribution and the number of contracts n_k in tariff cell k. In addition to these simple explicit premium principles, the safety loading can be imposed implicitly, e.g., in terms of convex principles of premium calculation including the well-known *exponential principle* or *Wang's premium principle* as special cases, cf. Example 4.1.

Pricing Systematic Risk

Systematic cyber incidents affect different firms at the same time—in contrast to idiosyncratic cyber incidents. Thus, (perfect) pooling of risk is no longer applicable and classical actuarial valuation has to be replaced by a more complex analysis. In the insurance context, systematic risk is not limited to cyber risk only, but also plays a prominent role in the valuation of unit-linked policies or the calculation of the market-consistent embedded value (MCEV) of an insurance portfolio, whereby idiosyncratic actuarial risk and systematic financial market risks must be evaluated together. For an overview on financial pricing methods in insurance we refer to Bauer et al. [6].

In this section, we propose a valuation of systematic cyber risk in terms of modern financial mathematics, combining the principle of *risk-neutral valuation* with the theory of *monetary risk measures*, see Knispel et al. [74] for a similar discussion related to the *Market-Consistent Embedded Value* (MCEV) of insurance portfolios. This comprehensive approach requires two different probability measures: While the assessment of risk in terms of a monetary risk measure is based on the real-world measure $\mathbb{P}$ that models the relative frequency with which events occur, valuation of contingent claims in financial mathematics relies on a technical measure $\mathbb{Q}$, called *risk-neutral measure* or *martingale measure*. In concrete application, tractable models may be obtained by assuming that the systematic one-year losses $\mathcal{S}_1^{m,\text{systematic}}$ in Eq. (7) is triggered by common risk factors to which all policyholders are jointly exposed. The total claim amount may be viewed as a contingent claim, depending on the evolution of these common factors.

Generally speaking, contingent claims are contracts between two or more parties which determine future payments to be exchanged between the parties conditional or contingent on future events. Formally, a contingent claim with payoff at terminal time $t = 1$ is described as a random variable. In financial mathematics, the valuation of contingent claims relies on a financial market model on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$ with a number, say $d + 1$, of liquidly traded primary products with price processes $(P_t^0)_{t\in[0,1]}, (P_t^1)_{t\in[0,1]}, \ldots, (P_t^d)_{t\in[0,1]}$. The underlying price processes can be modeled either as stochastic processes in discrete time or in continuous time. The asset '0' plays the role of a numéraire, i.e., it is used for discounting purposes. A contingent claim H_1 maturing at time $t = 1$ is called replicable or hedgeable if there exists a self-financing trading strategy[40] $\vartheta = (\vartheta_t^0, \vartheta_t^1, \ldots, \vartheta_t^d)_{t\in[0,1]}$

[40] Intuitively, the self-financing condition means that the portfolio is always rearranged such that on the one hand no additional capital is required and on the other hand no capital is withdrawn.

(specifying the number of shares ϑ_t^i of primary products in the portfolio at time t) whose terminal wealth V_1^ϑ coincides with the payoff H_1 for almost all scenarios. In absence of arbitrage, the price H_0 of a replicable contingent claim H_1 is unique and equals the *cost of perfect replication*. The calculation of this price can, however, be decoupled from the calculation of the replication strategy itself by the *principle of risk-neutral valuation*. Formally, risk-neutral valuation resembles the classical actuarial valuation in the sense that prices are computed as expectation of future discounted payments. The real-world measure $\mathbb{P}$ must, however, be replaced by a technical probability measure $\mathbb{Q}$, called *risk-neutral measure* or *martingale measure*. The latter name is motivated by the fact that discounted prices $(P_t^i/P_t^0)_{t\in[0,1]}$, $i = 0, 1, \ldots, d$, must be martingales with respect to $\mathbb{Q}$, i.e., the current discounted price at some time t is the best prognosis of the expected discounted price at a future date $s > t$ given the available information $\mathcal{F}_t$:

$$\mathbb{E}_{\mathbb{Q}}\left[\frac{P_s^i}{P_s^0}\middle|\mathcal{F}_t\right] = \frac{P_t^i}{P_t^0} \quad \text{for } 0 \le t < s \le 1, i = 0, 1, \ldots, d.$$

The risk-neutral valuation formula transfers this martingale property to the discounted prices of contingent claims. In particular,

$$H_0 = P_0^0 \mathbb{E}_{\mathbb{Q}}\left[\frac{H_1}{P_1^0}\right],$$

i.e., the cost of replication can be obtained as expectation of the discounted payoff with respect to any arbitrary (equivalent) martingale measure $\mathbb{Q}$.[41]

Markets are, however, typically incomplete[42] in the sense that not every contingent claim can be replicated in terms of liquidly traded primary products. In particular, contingent claims arising from cyber risks cannot be hedged perfectly in the financial market. For non-replicable contingent claims, risk-neutral valuation is still applicable, but now provides—depending on a whole class of martingale measures—an interval of prices which are consistent with the absence of arbitrage. Our evaluation of non-replicable contingent claims, however, is based on monetary risk measures and capital requirements.

Let us denote by $\mathcal{X}$ the set of financial positions with maturity $t = 1$ whose risk needs to be assessed. Mathematically, the family $\mathcal{X}$ is a vector space of real-valued mappings X_1 on $(\Omega, \mathcal{F}, \mathbb{P})$ that contains the constants. By sign-convention, negative values of X_1 correspond to debt or losses, i.e., the claims amount $\mathcal{S}_1^{m,\text{systematic}}$ corresponds to the financial position $X_1 = -\mathcal{S}_1^{m,\text{systematic}}$. A monetary risk measure[43] $\rho : \mathcal{X} \to \mathbb{R}$ quantifies the risk of a contingent claim that cannot be priced by the cost

[41] The martingale measure $\mathbb{Q}$ is said to be equivalent to the underlying real-world measure $\mathbb{P}$ if both measures have the same null sets, i.e., for any $A \in \mathcal{F}$ we have $\mathbb{Q}[A] = 0$ if and only if $\mathbb{P}[A] = 0$. Conceptually, this means that market events that are not relevant with respect to the real-world measure also play no role for the evaluation of contingent claims under $\mathbb{Q}$ and vice versa.

[42] In absence of arbitrage, incomplete financial market models are characterized by the existence of a whole class of equivalent martingale measures.

[43] For a rigorous introduction to the theory of risk measures we refer to Föllmer and Schied [51].

of perfect replication. Intuitively, a monetary risk measure can be viewed as a capital requirement: $\rho(X_1)$ is the minimal capital that has to be added to the position X_1 to make it acceptable, e.g., from the perspective of a financial supervisory authority, a rating agency, or the board of management. To capture the idea that homogeneous risks are assessed in the same way, we assume that ρ is distribution-based, i.e., $\rho(X) = \rho(Y)$ whenever the distributions of X and Y under $\mathbb{P}$ are equal. Prominent examples of distribution-based monetary risk measures are *Value at Risk* (VaR) and *Average Value at Risk* (AVaR).[44]

Combining these two approaches, an *algorithm* for the calculation of the premium $\pi^{m,\text{systematic}}$ can be summarized as follows (see also [74]):

1. Consider a decomposition of the financial position $-\mathcal{S}_1^{m,\text{systematic}} = H_1^m + R_1^m$, where
   - H_1^m is a replicable contingent claim with respect to the underlying market model,
   - and R_1^m denotes the residual term.
2. Calculate the premium $\pi^{m,\text{systematic}} = -H_0^m + \rho(R_1^m)$, where
   - H_0^m equals the cost of perfect replication of H_1^m, and the insurance needs to charge its customer the amount $-H_0^m$ for setting up[45] the offsetting position;
   - $\rho(R_1^m)$ is the premium for R_1^m.

The decomposition and the premium derived from it may not be unique. From the insurer's perspective, the goal of the decomposition into the summands (H_1^m, R_1^m) is the minimization of the theoretical premium $\pi^{m,\text{systematic}} = -H_0^m + \rho(R_1^m)$ which provides a lower bound[46] for the actual premium charge. The minimization problem apparently involves a trade-off between the cost of replication and the risk of the residual. In practice, it might be reasonable to impose constraints on the decomposition such as upper bounds for $-H_0^m$ and $\rho(R_1^m)$, respectively. Indeed, since the risk of the hedgeable part H_1^m can be completely eliminated for the price $-H_0^m$, the specification of a bound $\rho(R_1^m) \leq \rho_{\max}$ would already control the overall risk of the systematic cyber losses $\mathcal{S}_1^{m,\text{systematic}}$, in accordance with the company's risk strategy. Conversely, if the insurer's risk budget has not yet been exhausted, it might be helpful to limit

[44] For a financial position X_1, its Value at Risk at level $\lambda \in (0, 1)$ is the smallest monetary amount that needs to be added to X_1 such that the probability of a loss does not exceed λ:

$$\text{VaR}_\lambda(X_1) = \inf\{m \in \mathbb{R} | \mathbb{P}[X_1 + m < 0] \leq \lambda\}.$$

In particular, $\text{VaR}_\lambda(X_1) = -q_{X_1}^+(\lambda)$, where $q_{X_1}^+$ is the upper quantile function of X_1. The Average Value at Risk, also called Expected Shortfall, at level $\lambda \in (0, 1]$ is defined by

$$\text{AVaR}_\lambda(X_1) = \frac{1}{\lambda} \int_0^\lambda \text{VaR}_\alpha(X_1)\, d\alpha.$$

[45] Observe that H_0^m is typically negative, thus $-H_0^m$ positive.

[46] This lower bound could also directly be computed via a modified risk measure that is constructed according to Föllmer and Schied [50], see also Chapter 4.8 in Föllmer and Schied [51].

the hedging expenses by a bound $-H_0^m \leq \bar{H}_{\max}$ and to accept the remaining risk $\rho(R_1^m)$.[47]

This concept can be applied for each cyber risk module standalone, but provides additional benefits at portfolio level. If the underlying risk measure ρ is even subadditive (and thus provides incentives for diversification), then the lower bound for the actual premium charge can be further reduced. More precisely, for any given decomposition $-\mathbb{E}_{\mathbb{P}}[\hat{\mathcal{S}}_1^{m,i}|\mathcal{F}] = H_1^{m,i} + R_1^{m,i}$ of the systematic term in Eq. (7) per cyber risk module $m = (c,k)$ and policyholder $i = 1,\ldots,n_k$ in group k into a hedgeable part $H_1^{m,i}$ and a residual summand $R_1^{m,i}$, the risk of the residual term of the aggregated systematic risk satisfies

$$\rho\left(\sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k} R_1^{m,i}\right) \leq \sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k} \rho(R_1^{m,i}),$$

while the costs of perfect replication are additive. Thus, the total premium required at portfolio level is in fact lower than the aggregated premiums:

$$-\sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k} H_0^{m,i} + \rho\left(\sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k} R_1^{m,i}\right) \leq \sum_{c=1}^{C}\sum_{k=1}^{K}\sum_{i=1}^{n_k}(-H_0^{m,i} + \rho(R_1^{m,i})).$$

The diversification effect can be allocated—according to the insurer's business strategy—to the cyber risk modules, yielding a reduction of the lower bound for the actual premium charge per module.

4.2 Pricing of systemic cyber risks

Systemic risk is an important issue in cyber insurance. If entities are interconnected, risks may spread and amplify in cyber networks. In addition, this process depends on investments in cyber security and self protection of the agents in the network. Insurance premiums may, in turn, influence investment decisions and thereby modify the safety of the system, cf. Sect. 3. How to deal with complex cyber systems and the computation of systemic cyber insurance premiums is a topic of current research.

In this section, we develop some new ideas and introduce a preliminary, stylized approach that builds a bridge between the pricing of cyber insurance contracts and systemic risk measures. We consider N interconnected insurance customers in a cyber network that are also subject to idiosyncratic and systematic risk. For simplicity, we suppose that there exists only a single insurance company that offers J types of contracts. There are two dates, $t = 0$ and $t = 1$. The initial prices of the insurance contracts, represented by a matrix $\pi = (\pi_{i,j})_{i,j} \in \mathbb{R}^{N\times J}$, are $\pi_{i,j}$ where $i = 1,2,\ldots,N$ denotes the insurance customer and $j = 1,2,\ldots,J$ the contract type. Each customer

[47] To ensure the existence of a decomposition under constraints, the bounds on risk $\rho_{\max}$ and hedging expenses $\bar{H}_{\max}$ must satisfy the lower bounds $\rho_{\max} \geq \inf\{\rho(R_1^m)|\exists\,(H_1^m, R_1^m) : -\mathcal{S}_1^{m,\text{systematic}} = H_1^m + R_1^m\}$ and $\bar{H}_{\max} \geq \inf\{-H_0^m|\exists\,(H_1^m, R_1^m) : -\mathcal{S}_1^{m,\text{systematic}} = H_1^m + R_1^m\}$, respectively.

i chooses a contract type j_i from this menu and is charged a premium π_{i,j_i}. Customers decide simultaneously about insurance contracts and their investments into cyber security resulting in random losses $Y^\Pi = (Y_i^\Pi)_{i=1,2,\dots,N}$ at date $t = 1$, the end of the considered period.

In this setting, we discuss the notion of *systemic premium principles*. Suppose that—excluding the considered cyber business—the random net asset value of the insurance firm at date $t = 1$ is given by $\tilde{E}$. Including the cyber contracts, the net asset value[48] of the insurance firm is

$$E^\Pi = \tilde{E} + \sum_{i=1}^N \pi_{i,j_i} - \sum_{i=1}^N Y_i^\Pi. \tag{8}$$

The computation of the net asset value implicitly considers network effects that influence losses and the underlying investment decisions of the insurance customers, i.e., the systemic risk inherent in the network.

Systemic premium principles[49] refer to the family of premium matrices Π that are consistent with solvency requirements or risk limits and admissibility requirements of the insurance company. These can, for example, be formalized in terms of two acceptance sets[50] $\mathcal{A}^E$ and $\mathcal{A}^Y$ of monetary risk measures. The solvency condition or risk limit is satisfied, if $E^\Pi \in \mathcal{A}^E$. An admissibility requirement is that the stand-alone business is viable, i.e.,

$$\sum_{i=1}^N \pi_{i,j_i} - \sum_{i=1}^N Y_i^\Pi \in \mathcal{A}^Y. \tag{9}$$

Conditions (8) and (9) characterize the systemic premiums, i.e., the family $\mathbb{M}^\Pi$ of admissible premium matrices Π.

Idiosyncratic risk and systematic risk can also be priced within this framework. Idiosyncratic risk can be priced by classical actuarial premium principles. This was discussed in Sect. 4.1. Many premium principles correspond to monetary risk measures that can be encoded by acceptance sets, leading to a framework that is consistent with our suggested approach for pricing systemic risk. The same applies to the residual part of systematic risks that is not replicated. If the insurance firm has access to a financial market that is itself not exposed to systemic risk, it may use this market to partially hedge its exposure. In the absence of systemic risk, this was outlined in Sect. 4.1. In the current setting, the impact on insurance pricing of trading in financial markets can be included by adjusting the acceptance sets $\mathcal{A}^E$ and $\mathcal{A}^Y$ according to Föllmer and Schied [50], see also Chapter 4.8 in Föllmer and Schied [51].

Example 4.1 Solvency regulation varies in different regions of the world. Solvency II in the European Union and the Swiss Solvency Test in Switzerland are based[51] on

[48] The interest rate over the considered period is set to 0 in this example.

[49] The suggested concept of systemic premium principles parallels the notion of systemic risk measures of Feinstein et al. [48], see also Biagini et al. [8].

[50] See [52] and [51] for reviews on monetary risk measures.

[51] To be more precise, the implementation of the regulatory rules are based on Mean-VaR and Mean-AVaR. Details are, e.g., discussed in Weber [109], Hamm et al. [60].

the risk measures VaR and AVaR, respectively. These risk measures would define the acceptance set $\mathcal{A}^E$ in our setting.

The acceptance set $\mathcal{A}^Y$, in contrast, corresponds to a classical premium principle. Indeed, important actuarial premium principles are based on convex risk measures ρ (defined w. r. t. financial positions) and their acceptance sets $\mathcal{A}$, respectively,[52] by choosing

$$\rho(-L) = \inf\{\pi \in \mathbb{R} | \pi - L \in \mathcal{A}\}$$

as a premium for a loss position $L \in \mathcal{X} \subseteq L^0_+(\Omega, \mathcal{F})$.[53] Examples[54] of risk measures ρ corresponding to well-known actuarial premium principles are:

- **The family of entropic risk measures:**

$$\rho_\gamma(X) := \sup_{\mathbb{Q} \in \mathcal{M}_1} \{\mathbb{E}_\mathbb{Q}[-X] - \frac{1}{\gamma} H(\mathbb{Q}|\mathbb{P})\}, \quad \gamma \in (0, \infty).$$

  Here, $\mathcal{M}_1$ is the set of all probability measures on $(\Omega, \mathcal{F})$, and

$$H(\mathbb{Q}|\mathbb{P}) := \begin{cases} \mathbb{E}_\mathbb{Q}[\log \frac{d\mathbb{Q}}{d\mathbb{P}}], & \text{if } \mathbb{Q} \ll \mathbb{P}, \\ \infty, & \text{else,} \end{cases}$$

  is the relative entropy of $\mathbb{Q}$ with respect to a reference measure $\mathbb{P}$, for example, the real-world measure. Using a variational principle for the relative entropy, the entropic risk measure ρ_γ takes the explicit form $\rho_\gamma(X) = \frac{1}{\gamma} \log \mathbb{E}_\mathbb{P}[\exp(-\gamma X)]$ and thus corresponds to the exponential premium principle for the claims amount $Y = -X$. Note that $\rho_\gamma(X)$ is increasing in γ and satisfies

$$\lim_{\gamma \downarrow 0} \rho_\gamma(X) = \mathbb{E}_\mathbb{P}[-X] \quad \text{and} \quad \lim_{\gamma \uparrow \infty} \rho_\gamma(X) = \operatorname{ess\,sup}(-X),$$

  i.e., the limiting cases are the negative expected value $\rho(X) = \mathbb{E}_\mathbb{P}[-X]$ (net risk premium) as a lower bound and the maximum loss as an upper bound for premium charges.
- **Distortion risk measures:** For any increasing function $\psi : [0, 1] \to [0, 1]$ with $\psi(0) = 0$ and $\psi(1) = 1$ the map $c^\psi(A) := \psi(\mathbb{P}(A))$, $A \in \mathcal{F}$, is called a distortion of a probability measure $\mathbb{P}$. The Choquet integral

$$\rho^\psi(X) := \int (-X) dc^\psi = \int_0^\infty c^\psi[-X > x]\, dx + \int_{-\infty}^0 (c^\psi[-X > x] - 1)\, dx$$

[52] A monetary risk measure can be recovered from its acceptance set $\mathcal{A}$ via $\rho(X) = \inf\{m \in \mathbb{R} | X + m \in \mathcal{A}\}$, i.e., $\rho(X)$ is the smallest capital amount that has to be added such that the financial position becomes acceptable, see, e.g., Föllmer and Schied [51].

[53] For details, see Section 8 in Föllmer and Knispel [49] and the references therein.

[54] We include these examples only for the purpose of illustrating the tractability of the suggested approach that may build on well-known premium principles. Of course, the acceptance set of any other monetary risk measure can also be used. A decision should be made on the basis of what properties are desired.

defines a distortion risk measure, a comonotonic risk measure. If the distortion function is concave, the distortion risk measure corresponds to Wang's premium principle

$$\rho^{\psi}(X) = \int_0^{\infty} \psi(\mathbb{P}(-X > x))\, dx > \int_0^{\infty} \mathbb{P}(-X > x)\, dx = \mathbb{E}_{\mathbb{P}}[-X]$$

that guarantees a non-negative loading for any loss position $Y = -X \geq 0$. In particular, the limiting case $\psi = \mathrm{id}$ again corresponds to the negative expected value which provides a lower bound for the actuarial premium.

If we introduce a weak partial order $\leq$ on the space of real-valued $(N \times J)$-matrices $\mathbb{R}^{N\times J}$ by component-wise $\leq$-comparison, the smallest admissible premiums $\bar{\mathbb{M}}^{\Pi}$ in the family $\mathbb{M}^{\Pi}$ of admissible premium matrices may be characterized. Although we are dealing only with one insurance firm in our specific construction, the heuristic argument of competitiveness might motivate to focus on premiums in $\bar{\mathbb{M}}^{\Pi}$ only. Typically, the admissible premium allocations will not be unique.

A remaining question is the choice of a specific premium allocation. Further criteria or objectives need to be specified for this purpose. We briefly discuss three options:

- **Competition:** The heuristic argument of competitiveness might also be used to argue that total premium payments should be as small as possible. This would lead to those allocations where $\sum_{i=1}^{N} \pi_{i,j_i}$ is minimal.
- **Competitive segments:** If some insurance customers are more price-sensitive and more important than other, one might introduce positive weights v_i, $i = 1, 2, \ldots, N$, and focus on allocations with minimal $\sum_{i=1}^{N} v_i \pi_{i,j_i}$.
- **Performance optimization:** If insurance customers were willing to accept any premium allocation in $\bar{\mathbb{M}}^{\Pi}$, one could formulate an objective function of the insurance company that determines specific premium allocations. This could be a utility functional or a performance ratio such as RoRaC.

A detailed analysis of systemic premium principles in specific models and their statistical and algorithmic implementation are challenging and important questions for future research.

5 Conclusion and future research

In this paper, we provided a comprehensive overview of the current literature on modeling and pricing cyber insurance. For this purpose, we introduced a basic distinction between three different types of cyber risks: *idiosyncratic*, *systematic* and *systemic* cyber risks. Models for both *non-systemic* risk types were discussed within the classical actuarial framework of collective risk models. The (separate) discussion of modeling *systemic* cyber risks then focused on risk contagion among network users in interconnected environments as well as on their strategic interaction effects. Finally, we presented concepts for an appropriate pricing of cyber insurance contracts that crucially relies on the specific characteristics of each of the three risk types.

For both practitioners and academic researchers, modeling and pricing cyber insurance constitutes a relatively new topic. Due to its relevance, the area of research is growing rapidly, but modeling and pricing approaches are still in their infancy. In the introduction, we highlighted four important challenges: *data*, *non-stationarity*, *aggregate cyber risks*, and *different types of risk*.

Classical actuarial approaches rely heavily on claims *data*. To date, for cyber insurance this data is sparse and often inaccessible in the actuarial context due to confidentiality issues. As more data becomes available, different modeling approaches could be more easily tested and evaluated. Therefore, the development of (freely accessible) data collections for cyber risks is an important topic for future research. This requires collaboration between researchers, insurance companies, IT firms, and regulators. However, due to the evolutionary nature of cyber risks and their *non-stationarity*, the evaluation of data needs to be adaptive, and the relevance of historical information will most likely decrease over time. For this reason, it is important to combine expert opinions supported by advanced modeling techniques with the statistical evaluation of data.

Our systematic differentiation of risk types—idiosyncratic, systematic and systemic—structures the development of models and the evaluation of data. This facilitates an appropriate consideration of *different types of risks*. We advocate a pluralism of models that provides multiple perspectives in an evolving environment where issues of data availability and data quality remain unresolved. *Aggregate cyber risks* represent an important challenge that needs to be addressed at the systematic and systemic level. In this regard, we have identified the following promising avenues for future research:

- *Data on contagion.* Epidemic network and contagion models require a special kind of data, namely connectivity data for designing realistic network topologies and epidemic spread data for determining epidemic parameter values. The non-stationarity of the cyber environment remains a challenge that must be addressed in this area as well.
- *Networks and contagion processes.* The analysis of large-scale network models and epidemic processes is a difficult task. Developing and improving models and assessment methods is an important research task. Monte Carlo simulations are computationally intensive, and moment closures in mean-field approaches lack estimates of the resulting approximation error. Implementation procedures need to be refined and validated. In addition, realistic loss models are needed for assessing contagious cyber risks.
- *Top-down approaches.* To capture contagion effects in digital networks without rendering the models impossible to handle, a number of top-down approaches has already been developed. These employ population-based models that omit the detailed structure of the underlying networks and processes. However, network topology, e.g., centrality or cluster effects, has a significant role to play. Existing models should therefore be improved via more elaborate refinements that bridge the gap between bottom-up network modeling and population-based top-down approaches.
- *Dynamic strategic interaction.* The analysis of strategic interaction effects in cyber models has focused almost exclusively on static frameworks. Such an oversimpli-

fication may be inappropriate in environments where systemic spread processes are present. Studying the strategic interaction of network participants in dynamic environments could improve our understanding of the effects of cyber insurance contracts on policyholder behavior and vice versa.

- *Multi-layer networks.* Both manufacturing industries and financial operations are now highly dependent on digital technology. Cyber attacks on critical infrastructure pose a systemic threat to modern societies. Such hierarchical systems are characterized by a high degree of interdependence. To understand cyber risks in these structures, analyzing multilayered networks offers a promising approach.
- *Pricing systemic cyber risks.* In Sect. 4.1, we outline an approach for pricing systemic cyber risks. It integrates classical valuation concepts and systemic risk measures as a basis for systemic premium principles. Future research must extend the theoretical methodology and apply systemic premium principles in specific models. In addition, statistical and algorithmic techniques need to be developed.

The list of research tasks that we provide here is not exhaustive. Many further challenges exist. Addressing them will contribute to a more resilient and safer cyber landscape in the future. The evolutionary nature of cyber risk, however, precludes all challenges from ever being finally resolved.

Acknowledgements We are grateful for useful comments of the reviewers and the editor.

Funding Open Access funding enabled and organized by Projekt DEAL.

Appendix A Classification of cyber risks

In this section, we present two exemplary classification approaches of cyber risks from an actuarial perspective: CRO Forum [24] and Zeller and Scherer [116]. For general cyber classification approaches without a specific focus on insurance, we refer to the information security literature, see, e.g., Harry and Gallagher [61] and the references therein.

CRO Forum [24]

suggests a classification by manifold factors summarized in Table 1. However, due to its granularity, it does not seem very suitable for modeling purposes. Indeed, the classification rather intents to provide a " starting point for discussion" ([24], p. 24) on a unifying framework for data-gathering purposes.

Zeller and Scherer [116]

provides a more model-oriented classification of cyber incidents, see Table 2. The paper distinguishes between idiosyncratic and systemic incidents. However, the latter

Table 1 CRO Forum [24] Classification Overview

Cyber incident	Event type	Root causes	Actor	Impact type
1 System malfunctions/misuse	Operational Risk Categories	A People	1 Nation states	Business interruption
2 Data confidentiality		B External causes	2 Organized criminals	Data and software loss
3 Data integrity/availability		C Processes	3 Hackers	Theft or fraud
4 Malicious activity		D System	4 Hacktivists	Cyber ransom and extortion
			5 Insiders	Breach of privacy
				Reputational damage
				Regulatory & legal defense costs
				Fines and penalties
				Physical asset damage
				and many more, in total: 22 categories

Table 2 Zeller and Scherer [116] Classification Examples (see Table 2 therein)

	Idiosyncratic incidents Targeted attack	Individual failure	Systemic events Untargeted attack	Mass failure
Data Breach (DB)	Targeted data theft	Individual unintended data disclosure	Data theft through widespread malware / phishing	Unintended data disclosure at cloud service provider
Business Interruption (BI)	Targeted (D)DoS / Ransomware attack	Disruption of IT system or process through accidental malfunction	Widespread ransomware attack	Cloud service outage disrupting business services
Fraud / General (FR)	CEO fraud through targeted (spear-)phishing attack	Accidental compromise of database by employee	Widespread ransomware attack or social engineering fraud	Accidental compromise of data stored at cloud service provider

category should, in our view, be further divided into *systematic* and *systemic* incidents, see the discussion in Sect. 2.

Appendix B Random network models

Two standard classes of undirected random networks are Erdős–Rényi networks and scale-free networks:

- **Erdős–Rényi networks.** The simplest random network model was introduced by Erdős–Rényi [44]: The Erdős–Rényi network $G_p(N)$ is constructed from a set of N nodes in which each of the possible $N(N-1)/2$ edges is present independently with equal probability p. The resulting degree distribution, i.e., the distribution of the number of neighbors of any node in the network, is binomial, since the probability to create a node of degree k (i.e., with k neighbors) $\mathbb{P}(k)$ is equal to the probability that this node is connected to exactly k other nodes and not connected to the remaining $N-1-k$ nodes of the network:

$$\mathbb{P}(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}.$$

  For large N and in the limit of constant average degree $(N-1)p \approx Np =: c$, the binomial distribution can be approximated by a Poisson distribution

$$\mathbb{P}(k) = e^{-c} \frac{c^k}{k!}.$$

- **Scale-free networks.** Empirical analysis in various research areas suggests that real-world networks exhibit much more heterogeneous degrees than Poisson distributions would suggest. Often a hierarchy of nodes is observable—with a few nodes of high degree (called *hubs*), and a vast majority of less connected nodes having a relatively low degree. Typically, the degree distribution is approximately *scale-free*, i.e., we have

$$\mathbb{P}(k) \approx ak^{-\lambda}, \qquad a > 0, \quad \lambda > 0.$$

  A special case with $\lambda = 3$ is given by the *Barabási-Albert model* where a growing network is generated following a preferential attachment rule, see [5] for details.

Appendix C Gillespie Algorithm

Algorithm C.1 (Gillespie). *Input:* Initial state of the system $x_0 \in E^N$ and initial time $t_0 \geq 0$.

1. *(Initialization)* Set the current state $x \to x_0$ and current time $t \to t_0$.

2. *(Rate Calculation)* For the current state of the system x, calculate the sum of rates for all possible transitions $q_x = \sum_{i=1}^{N} q_{x_i}$, where q_{x_i} denotes the rate for a state change of node i according to (3).
3. *(Generate Next Event Time)* Sample the next event time t_{new} from an exponential distribution with parameter q_x.
4. *(Choose Next Event)* Sample the node i_{new} at which the next transition occurs: Each node $i = 1, \ldots, N$ is chosen with probability q_{x_i}/q_x.
   Change the state $x_{i_{\text{new}}} \to y_{i_{\text{new}}}$ according to (3).
5. Set $t \to t + t_{\text{new}}$, $x \to (x_1, \ldots, x_{i_{\text{new}}-1}, y_{i_{\text{new}}}, x_{i_{\text{new}}+1}, \ldots, x_N)$ and return to Step 2 until a prespecified stopping criterion is met.

Appendix D Moment closures

This section provides details on moment closures as a measure to solve the Markovian master equation problems (5) and (6).

For node i, we let B_i be a representative of the Bernoulli random variables I_i, S_i, or R_i at a certain time t. The product $B_{j_1} \cdot \ldots \cdot B_{j_{k+1}}$, with pairwise different and ordered indices $j_1 < \ldots < j_{k+1} \leq N$, is denoted by B_J, $J = \{j_1, \ldots, j_{k+1}\}$. For example, B_J with $J = \{j_1, j_2, j_3\}$ may denote a triple $I_{j_1} S_{j_2} I_{j_3}$, or $S_{j_1} S_{j_2} I_{j_3}$, etc.

A moment closure now approximates the moment $\mathbb{E}[B_J]$ by

$$\mathbb{E}[B_J] \approx H(\mathbb{E}[B_{J_1}], \ldots, \mathbb{E}[B_{J_m}]), \quad J_1, \ldots J_m \subset J, \quad |J_1|, \ldots, |J_m| \leq k.$$

Assuming that the single variables B_i are independent leads to the simplest possible moment closure, the *first order independent approximation*, also known as NIMFA in the epidemic literature[55]. It is given by

$$\mathbb{E}[B_i B_j] = \mathbb{E}[B_i]\mathbb{E}[B_j] + \text{Cov}(B_i, B_j) \approx \mathbb{E}[B_i]\mathbb{E}[B_j].$$

Under this assumption, the full SIS and SIR dynamics are given by the ODE systems of Eqs. (5) and (6), respectively, when replacing second-order moments with the corresponding product of means. The resulting systems can easily be analyzed by standard techniques from ODE theory.[56]

However, in certain network structures, the first order independent approximation may yield a large approximation error, see, e.g., Fahrenwaldt et al. [47]. Hence, more complex approaches to moment closures have been derived. Examples include:

1. **Split closures:** These closures are considered by Fahrenwaldt et al. [47]. The main idea of split closures consists in splitting a set J of $k+1$ nodes into two disjoint and non-empty subsets J_1, J_2 of order $\leq k$:

$$H(\mathbb{E}[B_{J_1}], \mathbb{E}[B_{J_2}]) = F(\mathbb{E}[B_{J_1}]) \cdot F(\mathbb{E}[B_{J_2}]),$$

[55] NIMFA is short for "N-intertwined mean-field approximation", see [87] for details.

[56] For the SIS model, the linear stability condition $R_0 = \frac{\tau}{\gamma} < \frac{1}{\hat{\mu}}$ for the infection-free state of the network can be obtained, where $\hat{\mu}$ denotes the spectral radius, i.e., the largest absolute eigenvalue, of the adjacency matrix A.

with $J_1 \cap J_2 = \emptyset, \quad J_1 \cup J_2 = J, \quad |J_1|, |J_2| \leq k$ and a *mean-field function* $F : [0, 1] \to [0, 1]$. Different mean-field functions lead to different approximations, e.g.:

- **Independent approximation:** Using the identity as mean-field function, $F(y) = y$, the factorization of the moment of order $k+1$ is done as if the split components were independent:

$$\mathbb{E}[B_J] \approx \mathbb{E}[B_{J_1}]\mathbb{E}[B_{J_2}].$$

  For the special case $k = 1$, this equals the first order independent approximation derived above.
  In the SIS model, since

$$\mathbb{E}[I_J] = \mathbb{E}[I_{J_1}]\mathbb{E}[I_{J_2}] + \mathrm{Cov}(I_{J_1}, I_{J_2})$$

  and $\mathrm{Cov}(I_{I_1}, I_{I_2}) \geq 0$, cf. Cator and Mieghem [20], the independent approximation leads to an *upper bound* of infection probabilities.
- **Hilbert approximation:** The space of square-integrable random variables forms a Hilbert space with scalar product $\langle Y, Z\rangle := \mathbb{E}[Y \cdot Z]$ and corresponding norm $\|Y\| := \sqrt{\langle Y, Y\rangle} = \sqrt{\mathbb{E}[Y^2]}$. For $Y, Z \in L^2$, the scalar product defines an angle ϕ between the elements:

$$\langle Y, Z\rangle = \|Y\| \cdot \|Z\| \cdot \cos\phi. \tag{D1}$$

  Hence, taking the mean-field function $F(y) = \sqrt{y}$, and using the idempotence of Bernoulli random variables, a moment of order $k+1$ can be split into:

$$\mathbb{E}[B_J] \approx \sqrt{\mathbb{E}[B_{J_1}]}\sqrt{\mathbb{E}[B_{J_2}]}.$$

  Due to (D1), the resulting approximation error is low, if the angle ϕ is close to $0°$ or $180°$. This observation may be used to determine an optimal split (J_1, J_2) of a given index set J.
  In the SIS model, the Cauchy-Schwarz inequality implies that the first order Hilbert approximation leads to a *lower bound* of infection probabilities.

To apply these approximations, an appropriate partition scheme (J_1, J_2) for index sets J of order $k+1$ needs to be found. For the SIS model, a first optimal split procedure for both approximation types is suggested in Fahrenwaldt et al. [47], Algorithm 3.13.

2. **Kirkwood closures:** These closures constitute the main approach used in the epidemic literature. The underlying idea originates from statistical physics, precisely from the so-called BBGKY hierarchy, which describes the evolution dynamics of an interacting N-particle system, originally proposed by Kirkwood [71]: Considering a set $J \subset \mathcal{V}$ of $k+1$ nodes and the corresponding moment $\mathbb{E}[B_J]$, we only take correlations into account which are stemming from infectious transmissions *over paths of length at most* $k-1$, i.e., passing through a maximum of k nodes.

This idea reflects the original statistical physics approach that particle states may be assumed to be independent, if their distance exceeds a certain critical threshold. Now, assuming the independence of node states which are sufficiently far apart, the Kirkwood approximation estimates the moment $\mathbb{E}[B_J]$ of Bernoulli random variables with $J = \{j_1, \ldots, j_{k+1}\}$ through

$$H(\mathbb{E}[B_{J_1^1}], \ldots, \mathbb{E}[B_{J_{m_1}^1}], \ldots, \mathbb{E}[B_{J_1^k}], \ldots, \mathbb{E}[B_{J_{m_k}^k}]) = \prod_{i=1}^{k} \prod_{\ell=1}^{m_i} \mathbb{E}[B_{J_\ell^i}]^{(-1)^{k-i}},$$

where $J_\ell^i \subset J$ denotes a subset of size i, $i \leq k$, and $\ell \in \{1, \ldots, m_i\}$, i.e., m_i denotes the number of such subsets. A detailed derivation can be found, e.g, in Sect. V of Singer [104]. The Kirkwood approximation can be interpreted as generalization of the following scheme:
For $k = 1$, states of any two nodes are assumed to be independent, i.e., the approximation equals the first order independent approximation described above. For $k = 2$, we obtain a so-called **pair-based model**. Here, the system is closed on the level of triplets and *correlations over edges* are considered. In this case, the closure reads

$$\mathbb{E}[B_{j_1} B_{j_2} B_{j_3}] = \frac{\mathbb{E}[B_{j_1} B_{j_2}]\mathbb{E}[B_{j_1} B_{j_3}]\mathbb{E}[B_{j_2} B_{j_3}]}{\mathbb{E}[B_{j_1}]\mathbb{E}[B_{j_2}]\mathbb{E}[B_{j_3}]}.$$

Two different cases for the node triplet $\{j_1, j_2, j_3\}$ may be considered: For *closed* triplets, i.e., triplets in which edges exist between all pairs of nodes (triangles), node states are pairwise correlated, and hence, the closure cannot be reduced. In contrast, for an *open* triplet only consisting of edges (j_1, j_2) and (j_2, j_3), the states of nodes j_1 and j_3 are assumed to be independent, and therefore, the closure may be reduced to

$$\mathbb{E}[B_{j_1} B_{j_2} B_{j_3}] = \frac{\mathbb{E}[B_{j_1} B_{j_2}]\mathbb{E}[B_{j_2} B_{j_3}]}{\mathbb{E}[B_{j_2}]}.$$

This equals the *exact result* for $\mathbb{E}[B_{j_1} B_{j_2} B_{j_3}]$ *under the independence assumption*, using Bayes' theorem.
Thus, in the SIR Markov model, *exact closures* can be derived when considering *cut-vertices* i, i.e., nodes connecting two otherwise disconnected subgraphs G_1 and G_2 of the network: If i is in the susceptible state of the SIR model, the infection has not yet passed through this node. Hence, the infection processes in the subgraphs G_1 and G_2, that are connected solely through i, are independent of each other, see, e.g., Kiss et al. [73]. This result in particular applies to tree graphs, where, by definition, all nodes with degree higher than one are cut-vertices and all triplets are open with $B_{j_2} = S_{j_2}$. For tree networks, the SIR pair-based model is thus exact.

Appendix E Estimation of (cyber) epidemic network models

Statistical estimation relies first on an underlying statistical model that specifies a range of probabilistic mechanisms that might have generated the data, and second on the observable data. Both components, the model framework and the available data, define the statistical challenge that needs to be addressed. We briefly review some work that focuses on inference for SIS, SIR, or related models. In all cases, the resulting propagation process is simply denoted by $(X(t))_{t\geq 0}$, although we consider different models. The specification of the interaction between entities in the underlying probabilistic mechanisms in the statistical model can be interpreted as a graph G in this framework. The graph G may simply be encoded by an adjacency matrix in some models; in other, heterogeneous models, it might be described as a weighted graph corresponding to a matrix with entries different from 0 and 1 that encodes the interaction in the underlying statistical model.

In the context of statistical inference, some parameters of the interaction dynamics are unknown, such as overall infection and recovery rates, but in some problems the interaction graph G might still be known a priori, while in others the graph must be inferred on the basis of available data. We classify the estimation approaches for (cyber) epidemic network models of SIS-, SIR- or related type roughly in the following way:

First, we distinguish if on the level of the underlying statistical model the interaction graph G is known; second, on the level of the data, we differentiate two situations, i.e., the realization of the infection process X might be directly observable or, alternatively, only some related data might be observable, while the spread process itself is hidden. We refer to Sects. 3.2.1 and 3.2.2 for background on spread models. In this section, we provide a brief review of some papers that belong to the following possible categories:

−− G unknown & $(X(t))_{\geq 0}$ not directly observable:
In a cyber epidemic network context, Antonio et al. [3] propose a graph mining approach in a generalized SIS network model (in which infection rates are heterogeneous and self-infection is possible) where the process X is not directly observable, but only auxiliary communication data is available. A filtering mechanism is applied that deletes low-weighted edges below a minimum weight threshold. However, the model is not yet calibrated with real-world data. For readers interested in more general (inverse) problems, we refer to the book by Kolaczyk [75].

++ G known & $(X(t))_{t\geq 0}$ observable:
If G is known and the realization of X is observable, the statistical problem boils down to inference of the epidemic parameters τ and γ of the epidemic spread model in the case of a Markovian SIR network model. This is discussed in Sect. 6.1 of Britton [16]; the estimation can be implemented using a maximum-likelihood approach.

+(−) G known & only terminal individual information on $(X(t))_{t\geq 0}$ observable:
Britton [16], Sect. 6.1., discusses the case that G possesses fully connected subgraphs in the case of a Markovian SIR network model, i.e., a so-called household structure. In this case, a maximum-likelihood approach to esti-

mate the epidemic parameters is still feasible, if only the realization of X at a final date is observable, but not its whole evolution. However, the author emphasizes that, without making this specific structural assumption, epidemic parameter estimation is not straight-forward for arbitrary known graphs, if e.g. only the realization of the final number of infections is observable. One approach to overcome this difficulty could thus be to gather some additional time-dependent spread data.

$(+)(+)$ G unknown, but network model class fixed & only individual recovery information on $(X(t))_{t\geq 0}$ observable:
Another example are random graph models. For example, Britton [17] develop a Bayesian approach in a Markovian SIR model to estimate the epidemic parameters τ and γ jointly with the connection probability p in Erdős-Rényi type networks, if the spread process $(X(t))_{t\geq 0}$ or only the individual recovery processes $(R_i(t))_{t\geq 0} = (\mathbb{1}_{X_i(t)=R})_{t\geq 0}$, $i = 1, \ldots, N$, are observable (see also [59] for a generalization to the SEIR model and Gamma-distributed infection arrival times). Samples from the posterior distribution can be generated using MCMC methods.

$(-)(-)$ G unknown, but set of possible network model classes fixed & only aggregate infection information on $(X(t))_{t\geq 0}$ observable:
Often the *individual* time-dependent spread data is not observable, but only the evolution of the *aggregate* number of infections over time is known. To overcome this issue, for example, in a Markovian SIS framework, Lauro et al. [78] suggest a so-called birth-death process approximation (see also [117] for an extension of this approach to the question of forecasting an ongoing epidemic). Such birth-death processes keep track of the number of infected nodes at population level and thus present an approximation of the original Markovian spread processes in a reduced dimension. Lauro et al. [78] provide a Bayesian estimation framework in which the epidemic and network parameters for certain random network classes can jointly be estimated; in particular, the method is able to identify the most likely network class out of a regular, Erdős–Rényi, or Barabási-Albert model.

$-+$ G unknown & $(X(t))_{\geq 0}$ observable:
In the previously discussed approaches with an observable epidemic process, the network G is (partially) known—at least it belongs to a set of random network classes. How can one proceed if no prior information is available on the network on the level of the statistical model, but the realization of the infection spread process is observable? One suggestion is a cascade approach in (possibly non-Markovian) SI-models (also: activation/information diffusion models). The goal is to infer the network structure under the assumption that the cascades of infections are fully observable using a likelihood approach. The proposed methods in the literature mostly differ in their assumptions on the spread process. For example, Myers and Leskovec [89] and Gomez-Rodriguez et al. [58] assume homogeneous parametric infection arrival distributions (see also [57] for dynamically evolving networks), while Du et al. [36] do not impose distributional assumptions, but propose a kernel estimation method to estimate the network structure.

Our overview is not meant to be exhaustive, but intends to highlight different perspectives possibly implied by the structure of a specific application. We also refer to the surveys Brugere et al. [18] or Kolaczyk and Csárdi [76]. The current literature on (epidemic) network estimation is fragmented with each approach tackling only a specific problem at a time. A unifying methodology does not yet exist—and is maybe also not realistic to expect. Many questions remain open, see, e.g., the discussion in Britton [16]. Statistics for (cyber) epidemic network models will most likely continue to be a very active field of research in the future.

References

1. Aït-Sahalia Y, Cacho-Diaz J, Laeven RJ (2015) Modeling financial contagion using mutually exciting jump processes. J. Financial Econ. 117(3):585–606
2. Allianz (2022) Allianz risk barometer. Technical report, Allianz Global Corporate & Specialty
3. Antonio Y, Indratno SW, Simanjuntak R (2021) Cyber insurance ratemaking: a graph mining approach. Risks 9(12). https://doi.org/10.3390/risks9120224
4. Baldwin A, Gheyas I, Ioannidis C, Pym D, Williams J (2017) Contagion in cyber security attacks. J Oper Res Soc 68:780–791. https://doi.org/10.1057/jors.2016.37
5. Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286:509–512
6. Bauer D, Phillips RD, Zanjani GH (2013) Financial pricing of insurance. In: Dionne G (ed) Handbook of insurance. Springer, New York, pp 627–645
7. Bessy-Roland Y, Boumezoued A, Hillairet C (2021) Multivariate Hawkes process for cyber insurance. Ann Actuarial Sci 15:14–39. https://doi.org/10.1017/S1748499520000093
8. Biagini F, Fouque J, Frittelli M, Meyer-Brandis T (2019) A unified approach to systemic risk measures via acceptance sets. Math Finance 29(1):329–367
9. Biener C, Eling M, Wirfs JH (2015) Insurability of cyber risk: an empirical analysis. Geneva Papers Risk Insurance-Issues Practice 40(1):131–158
10. Böhme R (2005) Cyber-insurance revisited. In WEIS
11. Böhme R, Kataria G (2006) Models and measures for correlation in cyber-insurance. In WEIS 2:3
12. Böhme R, Laube S, Riek M (2018) A fundamental approach to cyber risk analysis. Variance
13. Bolot J, Lelarge M (2008a) A local mean field analysis of security investments in networks. In NetEcon '08: Proceedings of the 3rd international workshop on Economics of networked systems, pp. 25–30
14. Bolot J, Lelarge M (2008) Network externalities and the deployment of security features and protocols in the internet. In Proc, ACM Sigmetrics, Annapolis, MD
15. Bolot J, Lelarge M (2009) Economic incentives to increase security in the internet: The case for insurance. Proceedings of the 28th Conference on Computer Communications, Rio de Janeiro, Brazil: 1494–1502
16. Britton T (2020) Epidemic models on social networks—with inference. Stat. Neerlandica 74(3):222–241
17. Britton T, O'Neill PD (2002) Bayesian inference for stochastic epidemics in populations with random social structure. Scandinavian J Stat 29(3):375–390
18. Brugere I, Gallagher B, Berger-Wolf, TY (2018) apr. Network structure inference, a survey: Motivations, methods, and applications. ACM Comput. Surv. 51(2)
19. Böhme R, Schwartz G (2010) Modeling cyber-insurance: Towards a unifying framework. WEIS
20. Cator E, Mieghem PV (2014) Nodal infection in Markovian susceptible-infected-susceptible and susceptible-infected-removed epidemics on networks are non-negatively correlated. Phys. Rev E 89(5). https://doi.org/10.1103/physreve.89.052802
21. Chen SX, Huang TM (2007) Nonparametric estimation of copula functions for dependence modelling. Canadian J Stat. La Revue Canadienne de Statistique 35(2): 265–282. https://doi.org/10.1002/cjs.5550350205
22. Choroś B, Ibragimov R, Permiakova E (2010) Copula estimation. In: Jaworski P, Durante F, Härdle WK, Rychlik T (eds) Copula theory and its applications, Berlin, Heidelberg. Springer, Berlin Heidelberg, pp 77–91

23. Cooray K, Ananda MM (2005) Modeling actuarial data with a composite lognormal-pareto model. Scandinavian Actuarial J 2005(5):321–334
24. Forum CRO (2016) Cro forum concept paper on a proposed categorisation methodology for cyber risk. Technical report, CRO Forum
25. CSIS (2020) The hidden costs of cybercrime. Technical report, Center for Strategic and International Studies (CSIS) in partnership with McAfee
26. Czado C (2019) Analyzing dependent data with vine copulas. Lecture Notes in Statistics, Springer
27. Czado C, Nagler T (2022) Vine copula based modeling. Ann Rev Stat Appl 9(1):453–477
28. Da Fonseca J, Zaatour R (2014) Hawkes process: Fast calibration, application to trade clustering, and diffusive limit. Journal of Futures Markets 34(6):548–579
29. Dacorogna M, Kratz M (2020) Moving from uncertainty to risk: the case of cyber risk. In Cybersecurity in Humanities and Social Sciences. Res Methods Approach pp 123–152. WILEY
30. Dacorogna M, Debbabi N, Kratz M (2022) Building up cyber resilience by better grasping cyber risk via a new algorithm for modelling heavy-tailed data. ESSEC Working Paper
31. Daley DJ, Vere-Jones D (2003) An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods. Springer
32. de Zea Bermudez P, Kotz S (2010) Parameter estimation of the generalized Pareto distribution-part i. J Stat Plann Inference 140(6):1353–1373
33. de Zea Bermudez P, Kotz S (2010) Parameter estimation of the generalized Pareto distribution-part ii. J Stat Plan Inference 140(6):1374–1388
34. Deheuvels P (1979) La fonction de dépendance empirique et ses propriétés. un test non paramétrique d'indépendance. Bulletins de l'Académie Royale de Belgique 65(1): 274–292
35. Detering N, Mayer-Brandis T, Panagiotou K, Ritter D (2019) Systemic risk in networks. In Network Science. An Aerial View, pp. 59–77. Springer
36. Du N, Song L, Yuan M, Smola A (2012) Learning networks of heterogeneous influence. In: Pereira F, Burges C, Bottou L, Weinberger K (eds) Advances in neural information processing systems, vol 25. Curran Associates Inc
37. Edwards B, Hofmeyr S, Forrest S (2016) Hype and heavy tails: a closer look at data breaches. J Cybersecur 2(1):3–14
38. Ehrlich I, Becker GS (1972) Market insurance, self-insurance, and self-protection. J Political Econ 80(4):623–648
39. Eling M (2020) Cyber risk research in business and actuarial science. Euro Actuarial J: 1–31
40. Eling M, Jung K (2018) Copula approaches for modeling cross-sectional dependence of data breach losses. Insur Math Econ 82:167–180
41. Eling M, Wirfs J (2019) What are the actual costs of cyber risk events? Euro J Oper Res 272(3):1109–1119
42. Embrechts P, Klüppelberg C, Mikosch T (2013) Modelling extremal events: for insurance and finance. Stochastic Modelling and Applied Probability, Springer, Berlin Heidelberg
43. Embrechts P, Liniger T, Lin L (2011) Multivariate Hawkes processes: an application to financial data. J Appl Probability 48A:367–378
44. Erdős P, Rényi A (1959) On random graphs I. Publicationes Mathematicae Debrecen 6:290–297
45. Errais E, Giesecke K, Goldberg LR (2010) Affine point processes and portfolio credit risk. SIAM J Financial Math 1(1):642–665. https://doi.org/10.1137/090771272
46. ESRB ed. (2020) Systemic Cyber Risk. European Systemic Risk Board
47. Fahrenwaldt MA, Weber S, Weske K (2018) Pricing of cyber insurance contracts in a network model. ASTIN Bulletin 48(3):1175–1218
48. Feinstein Z, Rudloff B, Weber S (2017) Measures of systemic risk. SIAM J Financial Math 8(1):672–708
49. Föllmer H, Knispel T (2013) Convex risk measures: Basic facts, law-invariance and beyond, asymptotics for large portfolios. In Handbook of the Fundamentals of Financial Decision Making, Part II, Eds. L.C. MacLean and W.T. Ziemba. World Scientific
50. Föllmer H, Schied A (2002) Convex measures of risk and trading constraints. Finance Stochastics 6:429–447
51. Föllmer H, Schied A (2016) Stochastic Finance: An Introduction in Discrete Time (4 ed.). Walter de Gruyter
52. Föllmer H, Weber S (2015) The axiomatic approach to risk measurement for capital determination. Ann Rev Financial Econ 7:301–337

53. Genest C, Ghoudi K, Rivest LP (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika 82(3):543–552
54. Giesecke K (2008) Portfolio credit risk: Top-down vs. bottom-up approaches, In Frontiers in Quantitative Finance: Volatility and Credit Risk Modeling, ed. Cont, R., Chapter 10. Wiley
55. Gillespie DT (1976) A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J Comput Phys 22(4):403–434. https://doi.org/10.1016/0021-9991(76)90041-3
56. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81(25):2340–2361. https://doi.org/10.1021/j100540a008
57. Gomez-Rodriguez M, Leskovec J, Balduzzi D, Schölkopf B (2014) Uncovering the structure and temporal dynamics of information propagation. Netw Sci 2(1):26–65. https://doi.org/10.1017/nws.2014.3
58. Gomez-Rodriguez M, Leskovec J, Krause A (2012 feb). Inferring networks of diffusion and influence. ACM Trans. Knowl. Discov. Data 5(4)
59. Groendyke C, Welch D, Hunter DR (2011) Bayesian inference for contact networks given epidemic data. Scandinavian J Stati 38(3):600–616
60. Hamm AM, Knispel T, Weber S (2020) Optimal risk sharing in insurance networks. Euro Actuarial J 10(1):203–234
61. Harry C, Gallagher N (2018) Classifying cyber events: a proposed taxonomy. J Inform Warf 17(3):17–31
62. Herath H, Herath T (2011) Copula-based actuarial model for pricing cyber-insurance policies. Insurance Markets Companies 2(1):7–20
63. Hillairet C, Lopez O (2021) Propagation of cyber incidents in an insurance portfolio: counting processes combined with compartmental epidemiological models. Scandinavian Actuarial J: 1–24
64. Hillairet C, Lopez O, d'Oultremont L, Spoorenberg B (2022) Cyber-contagion model with network structure applied to insurance. Insur Math Econ 107:88–101. https://doi.org/10.1016/j.insmatheco.2022.08.002
65. Hofert M, Kojadinovic I, Mächler M, Yan J (2018) Elements of copula modeling with R. Springer
66. Johnson B, Laszka A, Grossklags J (2014a) The complexity of estimating systematic risk in networks. In Proceedings of the 27th IEEE Computer Security Foundations Symposium. CSF
67. Johnson B, Laszka A, Grossklags J (2014b) How many down? In Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security, pp. 495–500. ACM
68. Kermack W, McKendrick A (1927) A contribution to the mathematical theory of epidemics. Proc R Soc Lond. Ser A 115: 700–721
69. Khalili MM, Naghizadeh P, Liu M (2017) Designing cyber insurance policies: mitigating moral hazard through security pre-screening. In GAMENETS 2017:63–73
70. Kim G, Silvapulle MJ, Silvapulle P (2007) Comparison of semiparametric and parametric methods for estimating copulas. Comput Stat Data Anal 51(6):2836–2850. https://doi.org/10.1016/j.csda.2006.10.009
71. Kirkwood JG (1935) Statistical mechanics of fluid mixtures. J Chem Phys 3:300–313
72. Kiss IZ, Miller JC, Simon PL (2017) Mathematics of Epidemics on Networks, Volume 46 of Interdisciplinary Applied Mathematics. Springer
73. Kiss IZ, Morris CG, Sélley F, Simon PL, Wilkinson RR (2015) Exact deterministic representation of Markovian SIR epidemics on networks with and without loops. J Math Biol 70:437–464. https://doi.org/10.1007/s00285-014-0772-0
74. Knispel T, Stahl G, Weber S (2011) From the equivalence principle to market consistent valuation. Jahresbericht der Deutschen Mathematiker-Vereinigung 113(3):139–172
75. Kolaczyk ED (2009) Statistical Analysis of Network Data. Methods and Models. Springer Ser. Stat. New York, NY: Springer
76. Kolaczyk ED, Csárdi G (2020) Statistical Analysis of Network Data with R (Second Edition ed.). Springer
77. Laszka A, Panaousis E, Grossklags J (2018) Cyber-insurance as a signaling game: Self-reporting and external security audits. In: Proceedings of the 9th Conference on Decision and Game Theory for Security, pp. 508–520
78. Lauro FD, Croix JC, Dashti M, Berthouze L, Kiss IZ (2020) Network inference from population-level observation of epidemics. Sci Rep 10(1):1–14

79. Liu J, Li J, Daly K (2022) Bayesian vine copulas for modelling dependence in data breach losses. Ann Actuarial Sci: 1-24
80. Maillart T, Sornette D (2010) Heavy-tailed distribution of cyber-risks. Euro Phys J B 75(3):357–364
81. Marotta A, Martinelli F, Nanni S, Orlando A, Yautsiukhin A (2017) Cyber-insurance survey. Comput Sci Rev
82. Martinelli F, Orlando A, Uuganbayar G, Yautsiukhin A (2017) Preventing the drop in security investments for non-competitive cyber-insurance market. In: International Conference on Risks and Security of Internet and Systems, pp. 159–174
83. Martinelli F, Yautsiukhin A (2016) Security by insurance for services. In: Proceedings of the 1st International Workshop on Cyber Resilience Economics
84. Mazzoccoli A, Naldi M (2020) Robustness of optimal investment decisions in mixed insurance/investment cyber risk management. Risk analysis 40:550–564. https://doi.org/10.1111/risa.13416
85. McNeil AJ, Frey R, Embrechts P (2015) Quantitative risk management: concepts, techniques and tools-revised edition. Princeton University Press
86. Van Mieghem P, Cator E (2012) Epidemics in networks with nodal self-infection and the epidemic threshold. Phys Rev E 86(1). https://doi.org/10.1103/physreve.86.016116
87. Van Mieghem P, Omic J, Kooij R (2009) Virus spread in networks. IEEE/ACM Trans Netw 17:1–14. https://doi.org/10.1109/TNET.2008.925623
88. Mikosch T (2004) Non-Life Insurance Mathematics: An Introduction With Stochastic Processes. Number Bd. 13 in Non-life insurance mathematics: an introduction with stochastic processes. Springer
89. Myers S, Leskovec J (2010) On the convexity of latent social network inference. In: Lafferty J, Williams C, Shawe-Taylor J, Zemel R, Culotta A (eds) Advances in neural information processing systems, vol 23. Curran Associates Inc
90. Naghizadeh P, Liu M (2014) Voluntary participation in cyber-insurance markets. In: Proceedings of the 2014 Annual Workshop on Economics in Information Security
91. NIST (2022) Glossary of the national institute of standards and technology. https://csrc.nist.gov/glossary. Accessed: 2022-05-27
92. Ogut H, Menon N, Raghunathan S (2005) Cyber insurance and it security investment. In: Proceedings of the 4th Workshop on the Economics of Information Security
93. Pal R (2012) Cyber-insurance in internet security: A dig into the information asymmetry problem. CoRR abs/1202.0884. arXiv:1202.0884
94. Pal R, Golubchik L, Psounis K, Hui P (2014) Will cyber insurance improve network security? a market analysis. In: Proceedings of the 2014 INFOCOM, IEEE
95. Pal R, Golubchik L, Psounis K, Hui P (2019) Security pricing as enabler of cyber-insurance: a first look at differentiated pricing markets. IEEE Trans Dependable Secure Comput 16:358–372. https://doi.org/10.1109/tdsc.2017.2684801
96. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A (2015) Epidemic processes in complex networks. Rev Modern Phys
97. Powell BA (2020) The epidemiology of lateral movement: exposures and countermeasures with network contagion models. J Cyber Secur Technol 4:67–105. https://doi.org/10.1080/23742917.2019.1627702
98. Reinhart J (2022) Discussion on 'A comprehensive model for cyber risk based on marked point processes and its applications to insurance'(Zeller, Scherer). Euro Actuarial J 12:87–88
99. Romanosky S, Ablon L, Kuehn A, Jones T (2019) Content analysis of cyber insurance policies: How do carriers price cyber risk? J Cybersecur: 1–19
100. Schwartz G, Sastry S (2014) Cyber-insurance framework for large scale interdependent networks. In: Proceedings of the 3rd international conference on High confidence networked systems, pp. 145–154
101. Schwartz G, Shetty N, Walrand J (2013) Why cyber-insurance contracts fail to reflect cyber-risks. In 51st Annual Allerton Conference on Communication, Control, and Computing, Allerton 2013, Allerton Park & Retreat Center, Monticello, IL, USA, October 2-4, 2013, pp. 781–787. IEEE
102. Shetty N, Schwartz G, Felegyhazi M, Walrand J (2010) Competitive cyber insurance and Internet Security, pp. 229–247. Springer, US
103. Shetty N, Schwartz G, Walrand J (2010) Can competitive insurers improve network security? In Acquisti, A., Smith, S., Sadeghi, A.-R. (Eds.): Proceedings of the 3rd International Conference on Trust and Trustworthy Computing, in: Lecture Notes in Computer Science, vol. 6101, Springer, Berlin, Heidelberg, pp. 308–322

104. Singer A (2004) Maximum entropy formulation of the Kirkwood superposition approximation. J Chem Phys 121(8):3657–66
105. Staum J (2013) Counterparty contagion in context: contributions to systemic risk. In: Fouque J-P, Langsam J (eds) Handbook on Systemic Risk. Cambridge University Press, pp 512–544
106. Sun H, Xu M, Zhao P (2020) Modeling malicious hacking data breach risks. North Am Actuarial J. https://doi.org/10.1080/10920277.2020.1752255
107. TU Munich, Statistics Research Group. n.d. Vine copula models. https://www.math.cit.tum.de/math/forschung/gruppen/statistics/vine-copula-models/. Accessed: 2022-12-27
108. Wang W, Lu Z (2013) Cyber security in the smart grid: survey and challenges. Comput Netw 57(5):1344–1371
109. Weber S (2018) Solvency II, or how to sweep the downside risk under the carpet. Insur Math Econ 82: 191–200
110. WEF ed. (2016) Understanding systemic cyber risk. World Economic Forum
111. Welburn JW, Strong AM (2021) Systemic cyber risk and aggregate impacts. Risk Analysis
112. Wheatley S, Maillart T, Sornette D (2016) The extreme risk of personal data breaches and the erosion of privacy. Euro Phys J B 89(1):1–12
113. Wüthrich MV, Bühlmann H, Furrer H (2010) Market-consistent actuarial valuation. Springer, Berlin, Heidelberg
114. Xu M, Hua L (2019) Cybersecurity insurance: modeling and pricing. North Am Actuarial J 23(2):220–249
115. Yang Z, Lui J (2014) Security adoption and influence of cyber-insurance markets in heterogeneous networks. Performance Eval 74:1–17
116. Zeller G, Scherer M (2022) A comprehensive model for cyber risk based on marked point processes and its application to insurance. Euro Actuarial J 12:33–85
117. Zerenner T, Di Lauro F, Dashti M, Berthouze L, Kiss I (2022) Probabilistic predictions of SIS epidemics on networks based on population-level observations. Math Biosci. https://doi.org/10.1016/j.mbs.2022.108854