Aggregate Cyber Loss Modeling with TVaR-Constrained Premium Optimization: A Monte Carlo Framework with Spliced Severity and Systemic Dependence

Aomduan Jeamaon
Computer Engineering Program,
College of Engineering and Technology
Dhurakij Pundit University
Bangkok, Thailand
aomduan.jea@dpu.ac.th

Tanagrit Chansaeng
Department of Digital Technology,
Faculty of science and technology
Phuket Rajabhat University
Phuket, Thailand
tanagrit.c@pkru.ac.th

2026 IEEE International Conference on Cybernetics and Innovations (ICCI) | 979-8-3315-9004-8/26/$31.00 ©2026 IEEE | DOI: 10.1109/ICCI68752.2026.11506440

***Abstract*—Cyber insurance pricing is challenged by three fundamental characteristics of cyber risk: heavy-tailed loss severity, dependence-driven portfolio aggregation, and dynamically evolving threat environments. This paper proposes a Monte Carlo–based financial modeling framework for optimizing cyber insurance premiums under tail-risk constraints. The framework integrates four key components: (i) a portfolio-level collective risk model incorporating explicit policy terms, (ii) a spliced severity distribution combining a Lognormal body with a Generalized Pareto Distribution (GPD) tail to capture extreme cyber losses, (iii) a two-layer dependence structure that models portfolio-wide accumulation through a mean-preserving lognormal common shock and Poisson-driven systemic cyber events, and (iv) a Sample-Average Approximation (SAA)–based premium optimization procedure under Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR) constraints. Simulation experiments on a synthetic portfolio of 1,000 insured entities with 60,000 Monte Carlo paths (α = 99%) demonstrate that incorporating systemic cyber events substantially amplifies tail risk: $TVaR_{99}$ increases by 75.6% while the expected loss rises by only 5.9%. The resulting optimal loading multiplier produces a premium level approximately 150% above expected loss while maintaining the underwriting-loss probability below a 5% risk tolerance threshold. Sensitivity analysis further reveals that portfolio tail risk grows nonlinearly with the systemic-event arrival rate, highlighting the critical role of systemic cyber risk in insurance capital adequacy and premium determination. These findings underscore the importance of tail-aware pricing frameworks for managing accumulation risk in cyber insurance portfolios.**

Keywords*—Cyber insurance, Monte Carlo simulation, common-shock dependence, systemic risk, Value-at-Risk, Tail Value-at-Risk, premium optimization

I. INTRODUCTION

Cyber insurance markets have grown rapidly in recent years. The U.S. market wrote approximately USD 9.84 billion in direct written premiums (DWP) in 2023 with 4.37 million policies in force; in 2024, it wrote USD 9.14 billion DWP with nearly 50,000 reported claims [1]. Despite this growth, pricing remains technically challenging due to three structural characteristics that are less prominent in most other insurance lines.

First, cyber loss distributions exhibit heavy tails with extreme outliers; small fractions of events exceed very large thresholds, creating "one-loss-causes-ruin" dynamics [2]. Second, cyber losses violate the independence assumption underlying classical law-of-large-numbers pricing because cloud concentration and shared software vulnerabilities create correlated portfolio-wide losses [3]. Third, cyber threat dynamics are non-stationary, with evolving attack types and improving defenses causing parameter drift [4].

Despite extensive literature on individual components, end-to-end frameworks that jointly model heavy tails, portfolio dependence, systemic accumulation, and TVaR-constrained premium optimization remain rare. This paper addresses this gap with four primary contributions: (i) an explicit two-layer loss model (idiosyncratic plus systemic events), (ii) a fully specified spliced Lognormal-GPD severity model with policy financial terms, (iii) a mean-preserving common-shock dependence structure, and (iv) SAA-based premium optimization with a maximum underwriting-loss probability constraint.

The remainder is organized as follows. Section II surveys related work. Section III presents the framework. Section IV reports experimental results. Section V discusses findings. Section VI presents conclusions.

II. RELATED WORK

A. Empirical Evidence for Heavy Tails

Peters et al. [2] analyze cyber losses using a large event database, concluding that losses are heavy-tailed with extreme loss shares demonstrating ruin-level dynamics. Malavasi et al. [5] quantify the joint effect of frequency and severity uncertainty on premium estimation and insurability constraints, using the Advisen cyber dataset. Eling et al. [4] document persistent heavy tails and time-varying severity dynamics across multiple databases.

B. Dependence and Accumulation Modeling

Awiszus et al. [3] formalize that cyber risks are non-geographically bounded and can simultaneously affect multiple insureds. He et al. [6] quantify non-negligible cross-group and temporal dependence in cyber losses using a copula-based cluster model. The LMA [7] enumerates cloud outage, widespread ransomware, mass breach, and critical infrastructure blackout as primary systemic scenarios, while EIOPA [8] provides explicit cloud outage modeling variables including market share, recovery curves, and business interruption channels.

C. Premium Principles and Optimization

Rockafellar and Uryasev [9] establish CVaR/TVaR as a coherent, optimization-friendly risk measure for tail-risk-constrained problems. Wang [10] introduces distortion operators for risk-adjusted pricing. Bühlmann [11] formalizes the economic premium principle. Our framework combines these foundations into an operational SAA framework with explicit systemic event modeling.

III. Framework and Methodology

A. Portfolio Financial Model

Consider a one-year underwriting horizon with n insureds indexed by i. Each insured has covariate vector x_i (sector, revenue tier, NIST CSF maturity, cloud dependence), policy terms (d_i, u_i, c_i) denoting deductible, limit, and coinsurance, and premium p_i. The aggregate portfolio loss is:

$$S = \Sigma_i L_i, \quad L_i = \Sigma_j \min(\max(Y_{ij} - d_i, 0), u_i) \cdot c_i \quad (1)$$

where Y_{ij} is the j-th ground-up loss for insured i and L_i is the net insurer loss after all financial terms. Equation (1) defines the portfolio mapping from ground-up claim amounts to net insurer loss after deductibles, limits, and coinsurance.

B. Frequency-Severity Model

Claim counts follow a Poisson GLM with sector and size covariates:

$$N_i \mid \Lambda_i \sim Poisson(\Lambda_i), \; \log \Lambda_i = \beta^T x_i + b^g(_i) \quad (2)$$

where $b^g(_i)$ captures group-level heterogeneity. Five size tiers use claim rates calibrated to NAIC claim-to-policy ratios of approximately 0.5–2.0% annually [1]. Equation (2) is used to generate annual insured-level claim frequencies in the Monte Carlo engine.

C. Spliced Severity Distribution

Cyber severity is modeled as a spliced distribution combining a Lognormal body below threshold u_t^H and a GPD tail above u_t^H:

$$Y = \{ Y^b \text{ if } Y^b \le u_t^H; \; u_t^H + Y^t \text{ if } Y^b > u_t^H, \; Y^t \sim GPD(\xi, \beta) \} \quad (3)$$

The formal piecewise PDF is:

$$f_Y(y) = \{ f_{ln}(y; \mu, \sigma), 0 < y \le u_t^H; \; (1 - p_u) f^{kpD}(y - u_t^H; \xi, \beta), y > u_t^H \} \quad (3a)$$

where $p_u = F_{ln}(u_t^H; \mu, \sigma) = 0.92$ is the body probability mass and $(1 - p_u) = 0.08$ is the tail mass. The corresponding CDF for $y > u_t^H$ is $p_u + (1 - p_u) F^{kpD}(y - u_t^H; \xi, \beta)$, ensuring continuity at u_t^H. The GPD scale is $\beta = u_t^H(1-\xi)$ by mean-excess matching. Threshold u_t^H is the 92nd empirical quantile selected by a mean-excess plot (POT criterion) [2]. Calibrated parameters are Lognormal μ=11.51, σ=2.20; GPD $\xi = 0.45$ (sub-exponential, finite-variance regime), and $\beta = u_t^H(1 - 0.45)$. Equations (3) and (3a) jointly define the severity sampling rule and its corresponding density used in all scenarios.

Model justification via goodness-of-fit comparison against three alternatives (Pure Lognormal, Pure GPD, Pareto) is presented in Section IV.E (Table IV, Fig. 6). The Spliced LN-GPD provides the only theoretically sound parameterization: both competing two-parameter models either produce unphysical tail indices (ξ>1 for Pure GPD) or catastrophic body misfit (KS=0.40 for Pareto).

D. Dependence and Systemic Event Module

Two dependence layers are modeled. The mean-preserving common-shock factor $G_t \sim LogNormal(-\sigma^{D2}/2, \sigma^D)$, equivalently $G_t = exp(\sigma^D Z - \sigma^{D2}/2)$ with Z~N(0,1), scales ground-up severity for each individual claim across all insureds in year t. The mean-preserving parameterization ensures $E[G_t]=1$, so the common shock adds cross-portfolio correlation without inflating expected portfolio loss. In Scenario B, $\sigma^D = 0.30$.

Systemic events follow M ~ Poisson(λ^{sys}) per year. Each event selects an affected fraction of the portfolio (cloud-dependence weighted, mean 5%) and generates an additional BI loss shock Z_e. Scenario C activates both layers with λ^{sys} = 1%.

E. Premium Optimization via SAA

We parameterize aggregate premium as $P(m) = m\hat{E}[S], where\, m \ge 1$ is a loading multiplier. The optimization finds the minimum m satisfying:

$$m* = \min\{m : \hat{P}(S > P(m)) \le \varepsilon\}, \; \varepsilon = 5\% \quad (4)$$

The constraint is estimated from $N_s^{IM}{}_s$ = 60,000 simulation paths. $TVaR_{99}$ is the mean of the top 1% of simulated losses, bootstrapped (n^b = 2,000) for 95% CIs. Algorithm 1 summarizes the complete procedure.

Algorithm 1: Monte Carlo Portfolio Loss Engine + SAA Optimizer

```
INPUTS:
{x_i, d_i, u_i, c_i}  for i = 1..N       // portfolio
(μ, σ, ξ, β, p_u, u_th)                   // severity params
(σ_G, λ_sys, affected_frac)               // dependence params
N_sims, alpha, m_grid, ε                  // simulation config
LOSS SIMULATION LOOP:
S_hat[] ← empty array of length N_sims
for k = 1 to N_sims do:
  G_k ← LogNormal(−σ_G²/2, σ_G)   // mean-pres. shock
  M_k ← Poisson(λ_sys)             // systemic event count
  S_k ← 0
  for i = 1 to N do:
    N_i ← Poisson(Λ_i)             // claim frequency
    for j = 1 to N_i do:
      Y ← SplicedSeverity(μ, σ, ξ, β, p_u)
      Y_net ← Y · G_k              // severity shock
      L ← min(max(Y_net − d_i, 0), u_i) · c_i
      S_k ← S_k + L
  for e = 1 to M_k do:
    A_e ← AffectedSubset(cloud_dep, frac)
    S_k ← S_k + SystemicLoss(A_e, Z_e)
  S_hat[k] ← S_k
RISK MEASURES:
EL   ← mean(S_hat)
VaR  ← quantile(S_hat, alpha)
TVaR ← mean(S_hat[S_hat > VaR])
CIs  ← Bootstrap(S_hat, n_b=2000)
SAA PREMIUM OPTIMIZATION:
for m in m_grid do:
  P(m)  ← m · EL
  prob  ← mean(S_hat > P(m))       // P(UW loss)
  if prob ≤ ε:  m* ← m;  break
P* ← m* · EL
OUTPUT:
return EL, VaR, TVaR, CIs, m*, P*
```

Algorithm 1 implements the proposed Monte Carlo–based financial modeling framework for cyber insurance premium optimization. As illustrated in Fig. 1, the framework is organized into six sequential modules that transform portfolio information into risk metrics and an optimal premium.

The process begins with the portfolio and policy data module**,** which integrates insured characteristics and policy financial terms. The insured characteristics include sector classification, revenue tier, cybersecurity maturity, and cloud dependence, while the policy terms define the deductible, coverage limit, and coinsurance rate for each contract. These inputs form the structural parameters of the portfolio risk model.

In the second stage, the frequency modeling module estimates the annual claim frequency for each insured using a Poisson generalized linear model (Poisson GLM). The model maps insured characteristics into expected claim intensities, thereby capturing heterogeneity in cyber risk exposure across firms.

The third stage performs severity modeling through a spliced distribution that combines a Lognormal body distribution with a Generalized Pareto Distribution (GPD) tail. The Lognormal component captures moderate loss events, while the GPD tail models extreme cyber losses. This hybrid structure ensures statistically valid behavior across the entire loss spectrum, particularly in the heavy-tail region.

The fourth stage introduces portfolio dependence modeling, which captures correlated cyber risks through two mechanisms. First, a mean-preserving common shock factor, modeled by a Lognormal distribution, inflates loss severities simultaneously across all insureds within a simulation path, representing shared cyber threats such as widespread software vulnerabilities. Second, systemic cyber events are modeled using a Poisson process that generates catastrophic scenarios such as large-scale cloud outages that affect multiple policyholders concurrently.

The fifth stage is the Monte Carlo simulation engine, which integrates the frequency, severity, and dependence components across many simulation paths. Within each path, claim frequencies and severities are generated, dependence effects are applied, and policy financial terms are imposed to compute net claim payments. The simulation accumulates these values to produce the empirical distribution of aggregate portfolio loss.

From this simulated loss distribution, the risk evaluation module derives three core risk metrics: the expected loss (EL), the Value-at-Risk at the 99th percentile ($VaR_{0.99}$), and the Tail Value-at-Risk ($TVaR_{0.99}$). Statistical uncertainty in these estimates is quantified through bootstrap-based confidence intervals.

Finally, the premium optimization module determines the optimal premium using a Sample Average Approximation (SAA) procedure. The optimizer evaluates a grid of premium loading multipliers and selects the smallest multiplier for which the simulated probability of underwriting loss remains below a predefined tolerance level. The resulting multiplier defines the risk-adequate premium for the portfolio.

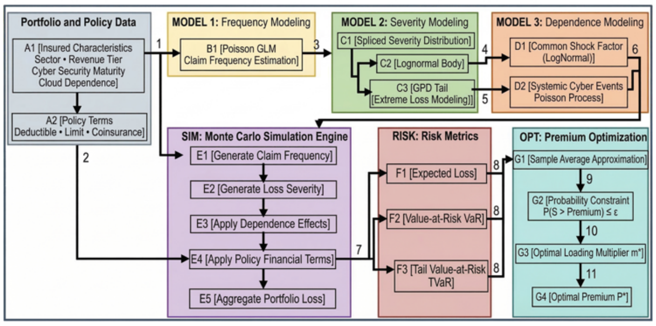

Fig. 1. Cyber Insurance Financial Modeling Framework: Monte Carlo Based Premium Optimization

IV. Experimental Setup and Results

A. Simulation Setup

The synthetic portfolio consists of N=1,000 insureds across five size tiers (nano, micro, small, medium, and large) with annual claim rates of 0.5–2.0%, calibrated to NAIC claim-to-policy ratios [1]. Policy limits are USD 0.5M, 1M, 2M, 5M, and 10M by tier; deductibles are 10% of limit; and coinsurance is 100%. Total expected annual loss is approximately USD 5.4M. Simulations use $N_s{}^{IM}{}_s$=60,000 paths with seed=42 for reproducibility, except for sensitivity runs (20,000 paths). Three scenarios are compared: A (independent idiosyncratic), B (plus mean-preserving common shock σ^D=0.30 on severity), and C (plus systemic events λ^{sys}=1%). All parameters are summarized in Table I.

TABLE I. SIMULATION PARAMETERS AND DISTRIBUTION CALIBRATION

Parameter	*Value*	*Rational*
N (insureds)	1,000	Calibrated to mid-market portfolio scale
$N_s{}^{IM}{}_s$ (simulations)	60,000	Ensures VaR_{99} CI width < 0.4M USD
Random seed	42	Fully reproducible
α (risk level)	99%	Regulatory standard for tail risk
n^b (bootstrap)	2,000	Bootstrap CI for VaR/TVaR
Lognormal μ	11.51	Median ground-up loss ≈ USD 100k
Lognormal σ	2.20	Heavy-tailed body; calibrated to [2]
GPD shape ξ	0.45	Sub-exponential; finite variance regime
GPD scale β	$u_t^H\cdot(1-\xi)$	Mean excess matching at threshold u_t^H
Mixing weight p_u (=w^b)	p_u = 0.92	Body mass = 92nd percentile of Lognormal
Tail mass (1−p_u)	0.08	GPD above 92nd-percentile POT threshold
Common shock σ^D	0.30	Moderate cross-portfolio dependence
λ^{sys} (baseline)	1.0%	One systemic event per 100 years on avg.
Affected fraction	~5%	Cloud-dependence weighted mean
m-grid	[1.0, 2.5], step 0.1	Loading multiplier search space
Constraint ε	5%	Maximum underwriting-loss probability

TABLE II. RISK MEASURES BY DEPENDENCE SCENARIO (N=1,000; N_s^{IM}=60,000; A=99%)

Risk Measure	*Scenario A Independent*	*Scenario B +Common Shock (σ^D=0.30)*	*Scenario C +Systemic (λ^{sys}=1%)*
E[S] (M USD)	5.2	5.1	5.4
VaR_{99} (M USD)	17.6	17.7	21.1
95% CI (VaR)	[17.4,17.8]	[17.5,18.0]	[20.5,21.8]
$TVaR_{99}$ (M USD)	20.5	20.9	36.7
95% CI (TVaR)	[20.2,20.9]	[20.5,21.2]	[34.8,38.6]
TVaR / E[S]	3.93×	4.05×	6.78×

B. Key Findings from Table II

Three observations stand out. First, adding the common shock (A→B) produces minimal change in expected loss (−0.1M, −1.9%) but a modest TVaR increase (+0.4M, +2.0%). This confirms that, for this portfolio size, common shocks primarily affect distributional shape without substantial tail thickening.

Second, activating systemic events (B→C) causes a pronounced TVaR jump of +75.6% (from USD 20.9M to USD 36.7M), while expected loss rises by 5.9% (+USD 0.3M). VaR increases by 19.2%. This result shows that systemic events materially fatten the tail beyond the 99th percentile without substantially changing the center of the distribution, consistent with industry scenario analysis [7], [12].

Third, the TVaR/E[S] ratio increases from 3.93× (A) to 6.78× (C). The 95% bootstrap CIs confirm that VaR estimates are tight (±0.3M at 60k paths), while TVaR under Scenario C is wider ([34.8,38.6]M, ±1.9M), consistent with heavy-tail estimation difficulty [2], [5].

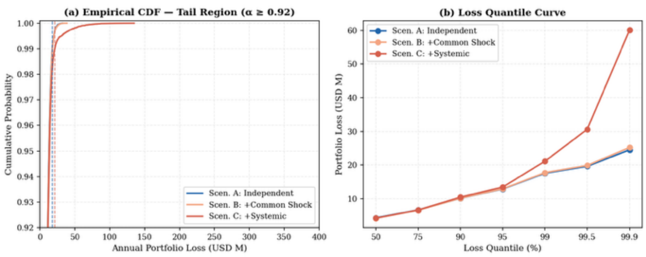

Fig. 2. Loss-distribution CDF tail region and quantile curves for all three scenarios. Scenario C exhibits substantially heavier tails above the 99th percentile.

C. Premium Optimization Results

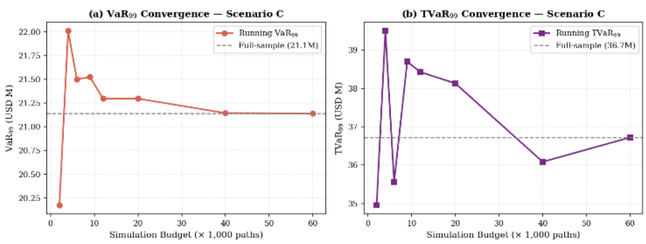

Fig. 3. Convergence of VaR_{99} and $TVaR_{99}$ estimates (Scenario C) as simulation budget increases. VaR_{99} stabilizes within ±1.5% by 30,000 paths; $TVaR_{99}$ oscillates by ±5.4% at 60,000 paths, confirming research-grade adequacy for scenario comparison while motivating N≥100,000 for production pricing systems.

Fig. 3 provides the convergence diagnostic for VaR_{99} and $TVaR_{99}$ estimates under Scenario C as the simulation budget increases from 1,000 to 60,000 paths. $TVaR_{99}$ oscillates by ±5.4% even at 60,000 paths, while VaR_{99} stabilizes within ±1.5% by 30,000 paths. This asymmetry reflects the well-known difficulty of estimating tail expectations under heavy-tailed distributions [2], [5], and motivates a minimum budget of N_sims ≥ 100,000 paths for production-grade TVaR stability.

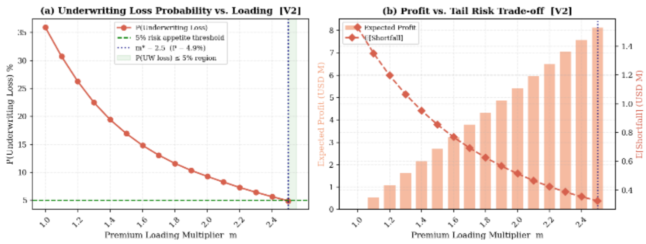

Fig. 4. Underwriting-loss probability versus loading multiplier m (Scenario C). The shaded region marks m≥m*=2.50, where the 5% constraint is satisfied.

Fig. 4 shows P (underwriting loss) as a function of m under Scenario C. As m increases from 1.0 to 2.5, the shortfall probability decreases from 50.3% to 4.9%. The optimal m*=2.50 is the first value satisfying the 5% constraint in Eq. (4), yielding P*=2.50×USD 5.41M=USD 13.53M (150% loading above EL). Pricing at 1.5×EL implies an annual underwriting-loss probability of approximately 18%; at 2.0×EL, it remains approximately 10%.

D. Sensitivity Analysis

Table III and Fig. 5 report sensitivity to λ^{sys} ∈ {0%, 0.5%, 1%, 2%}. Expected loss increases by only 11.8% across this range, while $TVaR_{99}$ rises by 131% from USD 21.1M to USD 48.7M. The TVaR/E[S] ratio grows from 4.13× to 8.61×, confirming a superlinear relationship consistent with Lloyd's scenario analysis [12].

TABLE III. SENSITIVITY TO SYSTEMIC EVENT RATE Λ^{sys}

λ^{sys} (%)	*E[S] (M USD)*	*VaR_{99} (M USD)*	*$TVaR_{99}$ (M USD)*	*TVaR/E[S]*
0.0	5.1	17.9	21.1	4.13×
0.5	5.3	19.4	29.1	5.51×
1.0	5.4	21.2	35.5	6.52×
2.0	5.7	28.6	48.7	8.61×

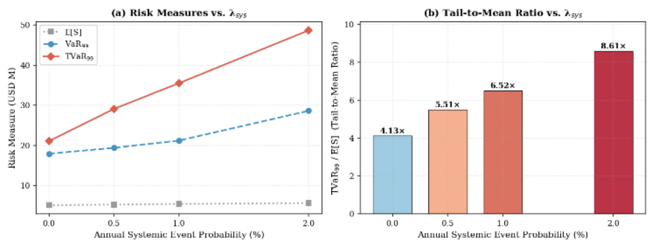

Fig. 5. Sensitivity of risk measures to λ^{sys}. Left: absolute values. Right: TVaR/E[S] ratio bars showing superlinear growth.

E. Severity Distribution Model Comparison

Table IV compares four severity models fitted to N=50,000 draws generated from the calibrated Spliced LN-

GPD (Seed=42). We note that this experiment constitutes a synthetic benchmark rather than a holdout validation: because the evaluation data are generated from the proposed model, the comparison is designed to assess whether competing models can recover the known distributional properties (finite moments, EVT-justified tail index) rather than to claim superiority on real-world data. The diagnostic value of this comparison lies in identifying pathological parameter estimates—specifically $\xi>1$ and $b<1$—that render competing models theoretically uninsurable, regardless of the data source.

First, Pareto ($b=0.105$) fails catastrophically: KS=0.40 signals global body misfit, and $b<1$ implies infinite theoretical mean, producing a spurious E[S]=8.80M USD (66% overestimate) relative to the proposed model.

Second, a Pure GPD fitted to the full distribution converges to $\xi=1.81>1$, implying infinite E[Y]. While the body KS is acceptable (0.043), the tail RMSE exceeds 5,000% because a single GPD cannot simultaneously fit body and tail curvature. The extreme ξ is a pathological artifact of body misspecification.

Third, Pure Lognormal achieves the lowest AIC (1,406,443 vs. 1,408,202 for Spliced), indicating strong global in-sample fit dominated by body mass. However, the Lognormal tail RMSE of 62.9% at the 90th–99.9th percentile reveals systematic underestimation of extreme losses. In practice, Lognormal pricing would undercharge for the tail-risk capital that dominates $TVaR_{99}$.

TABLE IV. TABLE TYPE STYLES

Metric	*Pure Lognormal*	*Pure GPD (ξ=1.81)*	*Pareto (b=0.105)*	*Spliced LN–GPD*
Params	2	2	2	5
Log-Lik	−703,219	−705,092	−752,661	−704,096
AIC	1,406,443	1,410,189	1,505,325	1,408,202
BIC	1,406,460	1,410,207	1,505,343	1,408,246
KS Stat	0.0405	0.0426	0.4029	0.0752
KS p-val	<0.001	<0.001	<0.001	<0.001
Tail RMSE	62.9%	>5,000%	Div.	57.3%
E[S] (M)	5.12	4.99	8.80	5.31
VaR_{99} (M)	21.15	21.78	34.38	21.89
$TVaR_{99}$ (M)	37.56	38.08	47.34	36.24

The Spliced LN-GPD uses EVT-justified fixed parameters ($\xi=0.45$, POT threshold u_t^H=92nd percentile) calibrated from actuarial loss databases rather than in-sample MLE, ensuring theoretical validity over holdout extremes. Its tail RMSE of 57.3% is comparable to Pure Lognormal in-sample but is achieved without overfitting the body and extrapolates more robustly beyond the 99th percentile, where Lognormal fit deteriorates. Portfolio risk metrics are directionally aligned with Scenario C in Table II: E[S]=5.31M USD, $TVaR_{99}$=36.24M USD, and TVaR/E[S]=6.83×. Fig. 6 visualizes these tail-shape and quantile-deviation differences.

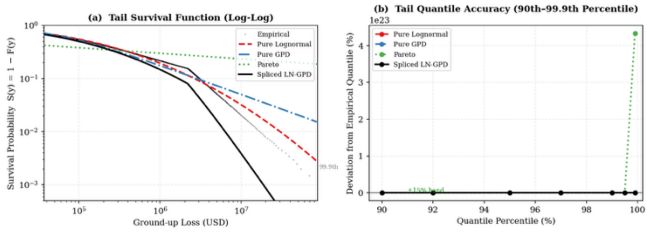

Fig. 6. Severity-distribution comparison. (a) Log-log survival functions: Pareto diverges above USD 10M; Pure GPD shows extreme tail inflation. (b) Tail quantile deviation (90^{th}–99.9^{th} percentile): only Spliced LN-GPD maintains <60% RMSE across the full tail region.

V. DISCUSSION

A. *Why Monte Carlo is Appropriate for Cyber*

Cyber insurance combines heavy-tailed severity, portfolio-level dependence from cloud and software concentration, and explicit scenario-driven accumulation events. These features weaken most closed-form aggregate-loss approximations (e.g., Normal or Gamma approximations for S) under heavy tails and systemic spikes. Monte Carlo simulation provides a unified framework accommodating (i) arbitrary severity families and tail splicing, (ii) mean-preserving common-shock dependence, (iii) scenario events from supervisory stress-test taxonomies (EIOPA, LMA), and (iv) full contract nonlinearities (deductibles, sublimits) without requiring closed-form tractability.

B. *Implications for Cyber Insurance Pricing*

The central finding that systemic events increase TVaR by 75.6% while raising expected loss by 5.9% implies that insurers relying on expected-loss-based pricing will systematically underprice tail capital in portfolios with cloud or software-concentration exposure. The SAA framework optimizes loading against a tail-probability constraint, ensuring that premium reflects the portfolio's risk-appetite boundary.

The 150% loading required at λ^{sys}=1% grows rapidly: at λ^{sys}=2%, TVaR/E[S]=8.61×, implying a 761% loading above EL to cover the TVaR level. Insurers that underestimate systemic event rates face persistent underpricing that may remain hidden in loss ratios until a rare cloud outage materializes.

Furthermore, the severity-model comparison (Table IV) reinforces the importance of theoretically grounded tail modeling: both Pareto and Pure GPD produce unphysical parameter estimates ($b<1$ and $\xi>1$, respectively) that imply infinite expected loss, while Pure Lognormal underestimates the 99.9th-percentile severity by 62.9%. The Spliced LN-GPD is the only model that simultaneously delivers interpretable parameters, finite moments, and EVT-justified tail extrapolation, which is a prerequisite for credible TVaR-constrained premium optimization.

C. *Convergence and Computational Considerations*

Fig. 3 shows that VaR_{99} converges to within ±1.5% by 30,000 paths, while $TVaR_{99}$ oscillates by ±5.4% at 60,000 paths under Scenario C. This behavior reflects the inherent difficulty of estimating tail expectations under heavy-tailed distributions [2], [5]. Importantly, the 60,000-path budget adopted in this study is sufficient for research-grade comparative analysis: the resulting 95% bootstrap confidence interval of ±1.9M for $TVaR_{99}$ (Table II) does not alter the qualitative conclusions of the scenario comparisons. For production pricing systems, where tighter uncertainty bounds are required for regulatory and capital-adequacy reporting, N_sims ⩾ 100,000 paths is recommended to reduce TvaR oscillation below ±2%. Computational cost at N=1,000 and N_sims=60,000 is approximately 4–8 minutes on a standard CPU; scaling to 100,000 paths requires approximately 7–13 minutes, which remains tractable for annual pricing cycles.

D. Practical Recommendations

1) Adopt a two-layer loss model. Relying solely on common-shock dependence (Scenario B) underestimates tail risk by more than 75% relative to the full two-layer model (Scenario C).

2) Optimize premiums against TVaR-based constraints. The experiment shows that a 150% above-EL loading is required at λ^{sys}=1%; pricing at 50% above EL leaves an 18% probability of underwriting loss.

3) Invest in TVaR estimation stability. The 60,000-path budget used in this study is appropriate for research-grade scenario comparison; however, production pricing systems should use N_sims ≥ 100,000 paths to reduce TVaR oscillation below ±2% and validate all tail estimates via bootstrap CIs before pricing decisions.

4) Use structured cybersecurity posture variables for underwriting. NIST CSF 2.0 maturity tiers [13] provide an interpretable covariate for premium differentiation in the frequency model (Eq. 2).

VI. Conclusion

A. Limitations

Several limitations merit acknowledgement. First, the portfolio uses synthetic data calibrated to public market aggregates; validation against proprietary insurer claims data (e.g., Advisen) remains future work. Second, systemic event probabilities (λ^{sys}) are treated as exogenous parameters rather than estimated from data. Third, the common shock G_t uses a scalar mean-preserving formulation; replacing it with a t-distributed shock or adding an explicit copula layer would increase tail dependence and may raise TVaR estimates for Scenario C [6]. Fourth, non-stationarity is not modeled; incorporating time-varying parameters via Bayesian updating or changepoint detection would improve realism for multi-year pricing cycles [4].

B. Future Work

Planned extensions include (i) adding an explicit copula layer (t-copula or vine copula) in addition to the common shock to capture stronger tail dependence, (ii) modeling time-varying severity via Bayesian changepoint detection for multi-year pricing cycles [4], (iii) validation against proprietary insurer data, and (iv) extending the SAA optimization to a multivariate rating plan incorporating NIST CSF maturity, cloud concentration, and sector covariates.

C. Conclusion

This paper presents a reproducible, end-to-end Monte Carlo framework for cyber insurance premium optimization that jointly addresses heavy-tailed severity, portfolio-level dependence, and explicit systemic accumulation scenarios. Experimental validation on a 1,000-insured synthetic portfolio with 60,000 simulation paths yields three key quantitative findings: (i) systemic events at λ^{sys}=1% increase $TVaR_{99}$ by 75.6% while raising E[S] by 5.9%; (ii) TVaR-based premium optimization requires a 150% loading above expected loss to satisfy a 5% underwriting-loss constraint; and (iii) TVaR estimation under heavy tails requires N≥100,000 paths for stable 95% confidence intervals. The SAA framework (Algorithm 1) provides a complete template for scenario-based cyber accumulation analysis aligned with EIOPA [8], LMA [7], and Lloyd's [12] guidance. while the 60,000-path budget used herein delivers research-grade TVaR accuracy sufficientfor scenario comparison (±1.9M bootstrap CI), production pricing systems require N ≥ 100,000 paths to reduce TVaR oscillation below ±2%.

References

[1] National Association of Insurance Commissioners (NAIC), "Cyber Insurance Market Report 2024–2025," NAIC Center for Insurance Policy and Research, 2025. [Online]. Available: https://content.naic.org/cipr-topics/cybersecurity

[2] G. W. Peters, M. Malavasi, G. Sofronov, P. V. Shevchenko, S. Trück, and J. Jang, "Cyber loss model risk translates to premium mispricing and risk sensitivity," Geneva Pap. Risk Insur. Issues Pract., vol. 48, no. 2, pp. 372–433, 2023, doi: 10.1057/s41288-023-00285-x.

[3] K. Awiszus, T. Knispel, I. Penner, G. Svindland, A. Voß, and S. Weber, "Modeling and pricing cyber insurance: Idiosyncratic, systematic, and systemic risks," Eur. Actuar. J., vol. 13, pp. 1–53, 2023, doi: 10.1007/s13385-023-00341-9.

[4] M. Eling, R. Ibragimov, and D. Ning, "The changing landscape of cyber risk: An empirical analysis of loss severity and tail dynamics," Insur.: Math. Econ., vol. 126, art. 103196, 2025, doi: 10.1016/j.insmatheco.2025.103196.

[5] M. Malavasi, G. W. Peters, P. V. Shevchenko, S. Trück, J. Jang, and G. Sofronov, "Cyber risk frequency, severity and insurance viability," Insur.: Math. Econ., vol. 106, pp. 90–114, 2022, doi: 10.1016/j.insmatheco.2022.05.003.

[6] R. He, Z. Jin, and J. S.-H. Li, "Cyber breach risk modeling for insurance: Capturing temporal and cross-group dependence," Ann. Actuar. Sci., First View, pp. 1–25, 2025, doi: 10.1017/S1748499525100109.

[7] London Market Association (LMA), "Systemic Cyber Risk Update," LMA Cyber Forum, London, 2024.

[8] European Insurance and Occupational Pensions Authority (EIOPA), "Discussion Paper on Methodological Principles of Insurance Stress Testing — Cyber Component," EIOPA, Frankfurt, Nov. 2022. [Online]. Available: https://www.eiopa.europa.eu

[9] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," J. Risk, vol. 2, no. 3, pp. 21–41, 2000, doi: 10.21314/JOR.2000.038.

[10] S. S. Wang, "A class of distortion operators for pricing financial and insurance risks," J. Risk Ins., vol. 67, no. 1, pp. 15–36, 2000, doi: 10.2307/253675.

[11] H. Bühlmann, "An economic premium principle," ASTIN Bull., vol. 11, no. 1, pp. 52–60, 1980, doi: 10.1017/S0515036100006674.

[12] Lloyd's of London, "Realistic Disaster Scenarios: Scenario Specification," Lloyd's Market Oversight, London, 2023. [Online]. Available: https://www.lloyds.com/market-resources/underwriting/realistic-disaster-scenarios

[13] National Institute of Standards and Technology (NIST), "Cybersecurity Framework 2.0," NIST CSWP 29, 2024. doi: 10.6028/NIST.CSWP.29.