ASYMPTOTICS FOR VAR AND CTE OF TOTAL AGGREGATE LOSSES IN A BIVARIATE OPERATIONAL RISK CELL MODEL

YISHAN GONG
Department of Mathematical Sciences, Xi’an Jiaotong Liverpool University
Suzhou, Jiangsu 215123, China

YANG YANG∗
School of Mathematics and Statistics, Nanjing Audit University
Nanjing, Jiangsu 211815, China

(Communicated by Ka Chun Cheung)

Abstract. This paper considers a bivariate operational risk cell model, in which the loss severities are modeled by some heavy-tailed and weakly (or strongly) dependent nonnegative random variables, and the frequency processes are described by two arbitrarily dependent general counting processes. In such a model, we establish some asymptotic formulas for the Value-at-Risk and Conditional Tail Expectation of the total aggregate loss. Some simulation studies are also conducted to check the accuracy of the obtained theoretical results via the Monte Carlo method.

1. Introduction. Consider a bivariate operational risk cell model, in which the severities \( \{(X^{(1)}_k, X^{(2)}_k), k \geq 1\} \) form a sequence of independent and identically distributed (i.i.d.) nonnegative random vectors with generic random vector \((X^{(1)}, X^{(2)})\) having marginal distributions \(F_1\) and \(F_2\), respectively; the frequency processes \(\{N_1(t), t \geq 0\}\) and \(\{N_2(t), t \geq 0\}\) describe the numbers of loss events of two business lines in the time interval \([0, t]\) for \(t \geq 0\), respectively. The severity processes \(\{(X^{(1)}_k, X^{(2)}_k), k \geq 1\}\) are independent of the frequency processes \(\{(N_1(t), N_2(t)), t \geq 0\}\). However, \(N_1(t)\) and \(N_2(t)\) are arbitrarily dependent, and for each \(k \geq 1\), some certain dependence structure may exist between \(X^{(1)}_k\) and \(X^{(2)}_k\). In this model, for each business line \(i = 1, 2\), the aggregate loss \(S_i(t)\) in \([0, t]\) constitutes a process

\[ S_i(t) = \sum_{k=1}^{N_i(t)} X^{(i)}_k, \quad t \geq 0, \]

and the bank’s total aggregate loss process is defined as

\[ S(t) = S_1(t) + S_2(t), \quad (1) \]

with distribution \(G_t(x) = P(S(t) \leq x)\).

Such a bivariate operational risk cell model aims to implement the Advanced Measurement Approach (AMA), which is proposed by Basel Committee on Banking
Supervision [6] according to Basel II accord and allows for explicit correlations between different operational risk events. Precisely speaking, banks should allocate losses to more than one business line or loss event type, which leads to the core problem of the multivariate modelling with all different risk type/business line cells. In the framework of mathematics, we consider the bivariate cell model in this paper for simplicity.

A frequency/severity approach, called Loss Distribution Approach (LDA), is one of the most popular approaches satisfying the AMA standards, and is widely used in banks and insurance companies, see, e.g. [1] and [40]. [16] and [28] showed that this model best fits the distribution of empirical data. In this paper, we aim to investigate the bivariate behavior of operational risk defined as the high quantile of a loss distribution, i.e. the Value-at-Risk (VaR) of the total aggregate loss \( S(t) \) in (1), and its corresponding Conditional Tail Expectation (CTE) as the confidence level increases.

[12] stated that although the LDA can be used to calculate the standard capital requirement of operational risk in insurance companies, the bridge between losses and business drivers, and dependencies among losses are not explicit and formal due to the lack of historical loss data. Therefore, a fundamental issue in this paper is to carefully address the dependence structure between different cells. Among some early works, [9] derived a simple closed-form expression for VaR in a single operational risk cell model, in which the severities are described as some i.i.d. heavy-tailed random variables. Later, some asymptotic results for VaR in a multivariate model were further provided by [10], in which a Lévy copula dependence is used to model different cells but in each cell the severities are still assumed to be i.i.d. and the frequency processes are restricted to homogeneous Poisson processes. Recently, [36] then established some approximations for both VaR and CTE in which some weak tail dependence among the severities of each cell are allowed and the frequency processes are extended to some general counting processes. Some related discussions on dependence modelling include [13], [11], [4], [40].

Remark that each pair of generic severity losses of a bank may be weakly tail dependent under the circumstance that the international financial and monetary system are stable, while they often exhibit strong tail dependence if the financial system is experiencing the economic downturns, see e.g. [23]. In this paper, we shall model the dependence between each pair of severity losses not only by some certain weak dependence, but also by some strong dependence to some extent. Concretely speaking, the asymptotic independence and the bivariate regular variation are considered, which are weakly and strongly dependent, respectively. In addition, we allow arbitrary dependence between the two general frequency processes.

The rest of this paper is organized as follows. After preparing some preliminaries in Section 2, the main results of this paper as well as some simulation studies are presented in Section 3. Section 4 concludes the paper. All proofs are postponed in Section 5.

2. Preliminaries. Throughout the paper, all limit relationships are according to \( x \to \infty \) unless otherwise stated. For two positive functions \( g_1(\cdot) \) and \( g_2(\cdot) \), we write \( g_1(x) \preceq g_2(x) \) or \( g_2(x) \succeq g_1(x) \) if \( \limsup g_1(x)/g_2(x) \leq 1 \), write \( g_1(x) \sim g_2(x) \) if \( \lim g_1(x)/g_2(x) = 1 \), and write \( g_1(x) = o(g_2(x)) \) if \( \lim g_1(x)/g_2(x) = 0 \). For a non-decreasing function \( g: \mathbb{R} \to \mathbb{R} \), denote by \( g^- \) the general inverse, that is, for \( y \in \mathbb{R} \), \( g^-(y) = \inf\{x \in \mathbb{R} : g(x) \geq y\} \), where \( \inf \emptyset = \infty \) by convention. For any set \( A \) and
any measure $\nu$, denote by $1_A$ the indicator function of $A$, and abbreviate $\nu(A)$ to $\nu A$ as long as no confusion arises.

2.1. Two risk measures. The calculation for capital requirement of operational risk is equivalent to find the VaR of total aggregate loss at confidence level 99.9% during one year period, see [6]. In our bivariate cell model, the VaR of total aggregate operational loss can be defined as the $q$-quantile of $G_t$,

$$\text{VaR}_q(S(t)) = G_t^{\leftarrow}(q),$$

for confidence level $q \in (0,1)$. Due to the lack of the sub-additivity, however, the VaR measure has been gradually replaced by the CTE, see Basel III [7]. The latter is used as the average capital requirement beyond the VaR, which can be defined as

$$\text{CTE}_q(S(t)) = \mathbb{E}[S(t)|S(t) > \text{VaR}_q(S(t))].$$

Some discussions on the merits and demerits of these two risk measures can be found in [20], but there is no evidence for global advantage of one risk measure against the other. Therefore, in this paper, we aim to establish some asymptotic formulas for both the VaR and CTE based on operational risks.

2.2. Heavy-tailed distributions and dependence structures. Investigating asymptotic approximations of any quantity of interest including our specific VaR and CTE measures, requires knowledge about the behavior in the extreme region. On the other hand, extremal/heavy-tailed distributions are often used to model loss variables, since many loss data are characterised by right heavy-tailedness, see [30], [15], [18], [5]. We start with the concept of subexponential distributions. By definition, a distribution $V$ on $\mathbb{R}^+ = [0,\infty)$ is said to be subexponential, written as $V \in S$, if

$$V(x) = 1 - \frac{V(x)}{V(x)} > 0 \text{ for all } x \geq 0 \text{ and the limit}$$

$$\lim_{x \to \infty} \frac{V^\ast_n(x)}{V(x)} = n$$

holds for all (or, equivalently, for some) $n \geq 2$, where $V^\ast_n$ is the $n$-fold convolution of $V$. Some typical examples of subexponential distributions are the Lognormal distribution, the Weibull distribution of the form

$$V(x) = 1 - e^{-cx^\tau}, \quad x \geq 0,$$

for some $c > 0$ and $0 < \tau < 1$, and the Pareto distribution of the form

$$V(x) = 1 - \left(\frac{1}{x + \sigma}\right)^\alpha, \quad x \geq 1 - \sigma,$$

with parameters $\alpha > 0$ and $\sigma > 0$.

One of the most useful subclass of subexponential distributions is that of regularly varying tailed distributions. Recall that a positive measurable function $h$ on $\mathbb{R}^+$ is said to be regularly varying at $\infty$ with index $\alpha \in \mathbb{R}$, written as $h \in \text{RV}_\alpha$, if

$$\lim_{x \to \infty} \frac{h(xy)}{h(x)} = y^\alpha, \quad y > 1.$$
\( V \in \mathcal{R}_{-\alpha} \) (or \( V \in \mathcal{R}_{-\infty} \)). Clearly, if a distribution \( V \in \mathcal{R}_{-\alpha} \) for some \( \alpha > 0 \), then \( V \in \mathcal{S} \); the Pareto distribution (3) belongs to \( \mathcal{R}_{-\alpha} \), and the Weibull distribution (2) belongs to \( \mathcal{S} \cap \mathcal{R}_{-\infty} \). Some applications to heavy-tailed distributions can be found in [39], [38], [37], among many others.

The concept of multivariate regular variation (MRV) is a natural extension of univariate regular variation, which provides an integrated framework for modelling extreme losses with both heavy tails and asymptotic (in)dependence in finance, insurance and risk management after the pioneer work from [17]. More related works on MRV include [24], [3], [35], [34], [33] among many others. We briefly introduce a special case of MRV with two dimensions.

A random vector \((\xi_1, \xi_2)\) taking values in \([0, \infty) \times (0, \infty] \) is said to follow a distribution with a bivariate regularly varying (BRV) tail if there exist a reference distribution \( V \) and a non-degenerate (i.e. not identically 0) limit measure \( \nu \) such that

\[
\lim_{x \to \infty} \frac{\mathbb{P}(1/x (\xi_1, \xi_2) \in \cdot)}{\nu(\cdot)} = \nu(\cdot) \quad \text{on } [0, \infty) \times \{0\},
\]

where \( \nu \to \) denotes vague convergence, meaning that the relation

\[
\lim_{x \to \infty} \frac{\mathbb{P}(\xi_1 > x, \xi_2 > x)}{\nu(1, \infty \setminus \{0\})} = 0;
\]

they are said to be asymptotically independent (AI), if

\[
\lim_{x \to \infty} \frac{P(\xi_1 > x, \xi_2 > x)}{\mathbb{V}_1(x) + \mathbb{V}_2(x)} = 0; \quad (7)
\]

they are said to be asymptotically dependent (AD), if they are tail proportional (i.e. \( \mathbb{V}_2(x) \sim c \mathbb{V}_1(x) \) for some \( c > 0 \)) and the limit in (7) is positive. It is noticeable that AD exhibits some strong dependence between variables. See [27] and [29]. We remark that if \((\xi_1, \xi_2) \in \text{BRV}_{-\alpha}(V)\) and \( \nu((1, \infty), (0, 1), \infty) > 0 \), then \( \xi_1 \) and \( \xi_2 \) are AD.

We now introduce some specific asymptotic independence structures. Assumption 2.1 below is introduced by [22] and is related to the so-called negative (or positive) regression dependence proposed by [25].
Assumption 2.1. Let \( \xi_1 \) and \( \xi_2 \) be two nonnegative random variables with distributions \( V_1 \) and \( V_2 \), respectively. There exist two positive constants \( x_0 \) and \( M \) such that

\[
P(\xi_i > x | \xi_j = y) \leq M V_i(x), \quad i \neq j = 1, 2,
\]

hold for all \( \min\{x, y\} \geq x_0 \), where the conditional probability in (8) is understood as 0 when \( y \) is not a possible value of \( \xi_j \), i.e. \( P(\xi_j \in \Delta) = 0 \) for some open set \( \Delta \) containing \( y \).

We remark that if \( \xi_1 \) and \( \xi_2 \) satisfy Assumption 2.1, then they are AI. Indeed, for sufficiently large \( x \),

\[
P(\xi_1 > x, \xi_2 > x) = \int_{x}^{\infty} P(\xi_1 > x | \xi_2 = y) V_2(dy)
\]

\[
\leq M V_1(x) V_2(x)
\]

\[
= o(1) \left( V_1(x) + V_2(x) \right),
\]

which implies (7). In addition, Assumption 2.1 is closely related to the following dependence proposed by [2].

Assumption 2.2. Let \( \xi_1 \) and \( \xi_2 \) be two nonnegative random variables with distributions \( V_1 \) and \( V_2 \), respectively. There exists a measurable function \( h(\cdot) : [0, \infty) \rightarrow (0, \infty) \) such that

\[
P(\xi_i > x | \xi_j = y) \sim h(y) V_i(x), \quad i \neq j = 1, 2,
\]

hold uniformly for all \( y \in [0, \infty) \).

Remark 1. Clearly, if \( h(\cdot) \) is bounded, then Assumption 2.2 implies Assumption 2.1. [26] verified Assumption 2.2 through copulas. It can be checked that the generic random pair \((X(1), X(2))\) coupled with some commonly-used copula, such as theAli-Mikhail-Haq copula, the Farlie-Gumbel-Morgenstern copula and the Frank copula, satisfies Assumption 2.2 with the corresponding measure function \( h(\cdot) \) bounded.

3. Main results and numerical studies.

3.1. Main results. We remark that the following propositions and theorems can be extended to the case of multivariate generic severity vector \((X^{(1)}, \ldots, X^{(d)})\), \( d \geq 2 \). Although our goal is to study the asymptotics for both VaR and CTE of total aggregate loss when confidence level increases, it is necessary to investigate the total aggregate loss’s tail probability. We make up several propositions in three different cases. The first one considers the case that each pair of the severities \((X^{(1)}, X^{(2)})\) is AI through satisfying Assumption 2.1 and with subexponential marginal distributions.

Proposition 1. Consider the bivariate cell model, in which the generic random vector \((X^{(1)}, X^{(2)})\) satisfies Assumption 2.1, and \( F_i(x) \sim c_i F(x) \) for some positive constants \( c_i, i = 1, 2 \), and some reference distribution \( F \in S \). If there exists some \( \varepsilon > 0 \) such that \( E \left[ (1 + \varepsilon)^{N_i(t)} \right] < \infty, i = 1, 2 \), then, regardless of arbitrary dependence between \( N_1(t) \) and \( N_2(t) \), it holds that

\[
P(S(t) > x) \sim (c_1 E[N_1(t)] + c_2 E[N_2(t)]) \cdot F(x).
\]
Example 3.1. Suppose that \((X^{(1)}, X^{(2)})\) follows the common marginal Weibull distribution \(F\) of the form (2), the dependence between \(X^{(1)}\) and \(X^{(2)}\) is modelled via the Frank copula of the form (16) below, and the two frequency processes \(N_1(t)\) and \(N_2(t)\) are described by two arbitrarily dependent homogeneous Poisson processes with intensities \(\lambda_1 > 0\) and \(\lambda_2 > 0\), respectively. Then, the asymptotic estimate in (9) can be reduced to

\[
P(S(t) > x) \sim (\lambda_1 + \lambda_2)t \cdot F(x). \tag{10}
\]

Compared with Proposition 1, the second proposition below strengthens the distribution of the severities to be regularly varying tailed, but allows each pair of the severity losses to be more general AI, and the two frequency processes to have some power moments.

Proposition 2. Consider the bivariate cell model, in which \(X^{(1)}\) and \(X^{(2)}\) are AI with \(F_i(x) \sim c_i F(x)\) for some positive constants \(c_i, i = 1, 2\), and some reference distribution \(F \in \mathcal{R}_{-\alpha}\), \(\alpha > 0\). If there exists some \(p > \alpha\) such that \(E[(N_i(t))^p] < \infty, i = 1, 2\), then, regardless of arbitrary dependence between \(N_1(t)\) and \(N_2(t)\), relation (9) holds.

Example 3.2. All assumptions are the same as those in Example 3.1, except replacing the common marginal distribution \(F\) by the Pareto distribution (3). Then, the asymptotic result in (9) can also be simplified to (10).

The third proposition considers an opposite case that each pair of severity losses are modelled by the bivariate regular variation, which is AD rather than AI.

Proposition 3. Consider the bivariate cell model, in which \((X^{(1)}, X^{(2)}) \in \text{BRV}_{-\alpha}(F)\) for some reference distribution \(F\) and some \(\alpha > 0\). If \(\nu([1, \infty]) > 0\) and there exists some \(p > \alpha\) such that \(E[(N_i(t))^p] < \infty, i = 1, 2\), then, regardless of arbitrary dependence between \(N_1(t)\) and \(N_2(t)\), it holds that

\[
P(S(t) > x) \sim \gamma F(x), \tag{11}
\]

where the coefficient

\[
\gamma = E \left[ \left( \nu(A) - \nu((0,1], \infty) \right) N_1(t) + \nu((0,1], \infty) N_2(t) \right] \mathbb{1}_{(N_1(t)<N_2(t))}
\]

\[
\quad + \left( \nu((1,0], \infty) N_1(t) + \left( \nu(A) - \nu((1,0], \infty) \right) N_2(t) \right) \mathbb{1}_{(N_1(t)\geq N_2(t))}, \tag{12}
\]

and \(A = \{(u,v) \in [0, \infty] : u + v > 1\}\).

Remark 2. If \(N_1(t) = N_2(t) = N(t)\), then the coefficient in (12) can be reduced to \(\gamma = \nu(A)E[N(t)]\).

Example 3.3. Suppose that \((X^{(1)}, X^{(2)})\) follows the common Pareto distribution \(F\) of the form (3), \(X^{(1)}\) and \(X^{(2)}\) are dependent according to the Gumbel copula of the form (17) below, and the frequency process \(N_1(t) = N_2(t) = N(t)\) is a homogeneous Poisson process with intensity \(\lambda > 0\). Then the asymptotic formula in (11) can be simplified to \(P(S(t) > x) \sim \nu(A) \cdot \lambda t \cdot F(x)\), where the coefficient \(\nu(A)\) can be further calculated, see Section 3.2.

Based on the above three propositions, we can establish some approximations for the VaR or CTE of total aggregate loss.
Theorem 3.1. Under the conditions of Proposition 1, if further $F_1 = F_2 = F \in S \cap R_{-\infty}$, then it holds that as $q \uparrow 1$,
\[ \text{VaR}_q(S(t)) \sim F^{\gamma^{-}}(q). \quad (13) \]

Theorem 3.2. Under the conditions of Proposition 2 with $F_1 = F_2 = F$, it holds that as $q \uparrow 1$,
\[ \text{VaR}_q(S(t)) \sim (c_1 E[N_1(t)] + c_2 E[N_2(t)])^{\frac{1}{\alpha}} F^{\gamma^{-}}(q), \quad (14) \] and if further $\alpha > 1$, then as $q \uparrow 1$,
\[ \text{CTE}_q(S(t)) \sim \frac{\alpha}{\alpha - 1} (c_1 E[N_1(t)] + c_2 E[N_2(t)])^{\frac{1}{\alpha}} F^{\gamma^{-}}(q). \quad (15) \]

Theorem 3.3. Under the conditions of Proposition 3, it holds that as $q \uparrow 1$,
\[ \text{VaR}_q(S(t)) \sim \gamma^{\frac{1}{\alpha}} F^{\gamma^{-}}(q), \] and if further $\alpha > 1$, then as $q \uparrow 1$,
\[ \text{CTE}_q(S(t)) \sim \frac{\alpha}{\alpha - 1} \gamma^{\frac{1}{\alpha}} F^{\gamma^{-}}(q), \]

where $\gamma$ is defined in (12).

3.2. Numerical studies. In this subsection, we perform some Monte Carlo simulation studies by using software R to check the accuracy of the theoretical Theorems 3.1–3.3. Firstly, we introduce some examples of AI and AD via copulas, which will be used to model the dependence between each pair of the severities or the two frequency processes. Assume that $H$, the joint distribution of $(\xi_1, \xi_2)$, has continuous marginal distributions $V_1$ and $V_2$, respectively. Then, by Sklar’s theorem, there exists a unique copula $C(u_1, u_2) : [0, 1]^2 \mapsto [0, 1]$ such that
\[ H(x_1, x_2) = C(V_1(x_1), V_2(x_2)). \]

Example 3.4. The Frank copula is of the form
\[ C(u_1, u_2) = -\frac{1}{\theta} \log \left(1 + \frac{e^{-\theta u_1} - 1}{e^{-\theta} - 1} \right), \quad \theta > 0. \quad (16) \]

It can be verified that if identically distributed $\xi_1$ and $\xi_2$ are dependent through the Frank copula, then they are AI, but the following Gumbel copula is a typical AD case.

Example 3.5. The Gumbel copula is of the form
\[ C(u_1, u_2) = \exp \left\{ -((\log u_1)^{\rho} + (\log u_2)^{\rho})^{\frac{1}{\rho}} \right\}, \quad \rho \geq 1. \quad (17) \]

Throughout this section, model specifications for the numerical studies are listed below:
- The severities $\{X_k^{(1)}, X_k^{(2)}\}, k \geq 1$ follow the common Weibull distribution (2) with parameters $c > 0$ and $0 < \tau < 1$, or the common Pareto distribution (3) with parameters $\alpha > 0$ and $\sigma > 0$. Clearly, $F_1 = F_2 = F \in S \cap R_{-\infty}$ for the former and $F_1 = F_2 = F \in R_{-\infty}$ for the latter.
- The frequency processes $N_1(t)$ and $N_2(t)$ are described by two homogeneous Poisson processes with intensities $\lambda_1 > 0$ and $\lambda_2 > 0$, respectively.
- The AI dependence between two variables is modelled via the Frank copula (16), whereas the AD dependence is according to the Gumbel copula (17).
All the numerical studies are conducted with the sample size \( m = 10^6 \) and the time period \( t = 1 \).

We first consider Theorem 3.1. For the simulated estimate \( \hat{\text{VaR}}_q(S(t)) \), we first generate \( m \) samples \( (N_{1(j)}(t), N_{2(j)}(t)), j = 1, \ldots, m \), where for each \( j \), \( N_{1(j)}(t) \) and \( N_{2(j)}(t) \) are both homogeneous Poisson processes with intensities \( \lambda_1 > 0 \) and \( \lambda_2 > 0 \), respectively, and they are dependent via the Frank (or Gumbel) copula (16) (or (17)) with parameter \( \theta_N \) (or \( \rho_N \)). For each \( j = 1, \ldots, m \), take \( N_\text{max}^{(j)}(t) = \max\{N_{1(j)}(t), N_{2(j)}(t)\} \), then generate \( N_\text{max}^{(j)}(t) \) pairs of the severities \( (X_{k(j)}^{(1)}, X_{k(j)}^{(2)}), k = 1, \ldots, N_\text{max}^{(j)}(t) \), such that the common marginal distribution is Weibull distributed of the form (2) with \( c > 0 \) and \( 0 < \tau < 1 \), and each pair is dependent via the Frank copula (16) with parameter \( \theta_X \). Thus, the total aggregate loss can be calculated according to (1),

\[
S_{(j)}(t) = \sum_{k=1}^{N_{1(j)}(t)} X_{k(j)}^{(1)} + \sum_{k=1}^{N_{2(j)}(t)} X_{k(j)}^{(2)}, \quad j = 1, \ldots, m.
\]

In this way, the VaR \( q \) is estimated by

\[
\hat{\text{VaR}}_q(S(t)) = \inf \left\{ x : \frac{1}{m} \sum_{j=1}^{m} I_{S_{(j)}(t) \leq x} \geq q \right\},
\]

for confidence level \( q \). The asymptotic value on the right hand side of (13) is computed by \( (\lambda_1 - 1 \log(1 - q))^{-1/\tau} \). By Remark 1, the Frank copula between each pair of severities satisfies Assumption 2.1, hence all conditions in Theorem 3.1 are satisfied.

The various parameters are set to \( \lambda_1 = 6 \), \( \lambda_2 = 7 \), \( \tau = 0.1 \), \( c = 0.5 \), \( \theta_X = 1 \), \( \theta_N = 1.2 \), \( \rho_N = 1.2 \). Then the simulation results for \( \hat{\text{VaR}}_q(S(t)) \) in Theorem 3.1 are shown in Figure 1. It can be seen that with the increase of confidence level \( q \), the simulated estimates increase quickly and get closer to the asymptotic estimates. As shown in Figure 1(a), the curves of the two simulated estimates under the Frank and
Gumbel copulas are almost overlapped with the asymptotic estimates. Furthermore, in Figure 1(b), their ratios fluctuate around 1 as \( q \) increases to 1. This indicates that the simulated VaR of total aggregate loss is insensitive with respect to different dependence structures between the two frequency processes, which coincides with our theoretical result. Therefore, our obtained asymptotic equation in Theorem 3.1 is reasonable.

We next consider Theorem 3.2. The procedure and all assumptions for generating the simulated estimate \( \hat{\text{Var}}_q(S(t)) \) in (18) are the same as those in Theorem 3.1, except replacing the severity distribution by the Pareto distribution (3) with \( \alpha > 0 \) and \( \sigma > 0 \). The asymptotic value on the right hand side of (14) is computed by

\[
E[N_1(t)] + E[N_2(t)] \left( 1 - q \right)^{-\frac{1}{\alpha}} - \sigma.
\]

As for \( \text{CTE}_q(S(t)) \), it can be estimated by

\[
\hat{\text{CTE}}_q(S(t)) = \frac{\sum_{j=1}^m S(j)1(S(j)>\hat{\text{VaR}}_q(S(t)))}{\sum_{j=1}^m 1(S(j)>(\hat{\text{VaR}}_q(S(t)))}.
\]

Figure 2. Comparison of the simulated and asymptotic estimates for \( \text{VaR}_q(S(t)) \), with Frank dependent and Pareto distributed severities and AI or AD dependent frequency processes in Theorem 3.2

The various parameters are set to \( \lambda_1 = 6, \lambda_2 = 7, \alpha = 0.5 \) (or \( \alpha = 1.78 \)), \( \sigma = 1, \theta_X = 1, \theta_N = 1.2, \rho_N = 1.2 \). Then the simulation results for \( \text{VaR}_q(S(t)) \) and \( \text{CTE}_q(S(t)) \) in Theorem 3.2 are shown in Figures 2 and 3, respectively. Despite some subtle difference, the curves in different copulas almost coincide, and the ratios of simulated and asymptotic values fluctuate around 1 as confidence level increases to 1, which indicates that under our settings, the \( \text{VaR}_q(S(t)) \) and \( \text{CTE}_q(S(t)) \) are affected little by the dependence between the two frequency processes. Comparing Figure 2 with Figure 1, it is noticeable that under the same confidence level \( q \), the \( \text{VaR}_q(S(t)) \) in Theorem 3.1 is much greater than that in Theorem 3.2. This is because the tail of the Pareto distribution is heavier than that of the Weibull distribution. In other words, the total aggregate loss \( S(t) \) is relatively smaller for Weibull distributed severities, and hence the corresponding \( q \)-quantile of \( S(t) \) with Weibull distributed severities is larger than that with Pareto distributed severities, which is consistent with the general rules of the operational risk management.

Finally, we turn to Theorem 3.3, for which the settings for \( N_1(t) \) and \( N_2(t) \) are the same as those in Theorem 3.1 with parameters \( \theta_N \) and \( \rho_N \) for the Frank and
Gumbel copulas, respectively; the generic severity random vector \((X^{(1)}, X^{(2)})\) is modelled via the Gumbel copula of the form (17) with parameter \(\rho_X \geq 1\), and has a common distribution \(F \in \mathbb{R}_{-\alpha}\) for some \(\alpha > 0\). Then, by Lemma 5.2 of [35], \((X^{(1)}, X^{(2)}) \in \text{BRV}_{-\alpha}(F)\) for some reference distribution \(F\), and it can be calculated that for any Borel set \(A \subset [0, \infty]\),

\[
\nu(A) = \alpha^2(\rho_X - 1) \times \int_A (u^{-\alpha\rho_X} + v^{-\alpha\rho_X})^\frac{1}{\rho_X - 1} u^{-\alpha\rho_X - 1} v^{-\alpha\rho_X - 1} dudv.
\]

This leads to \(\nu((1, 0), \infty] = \nu((0, 1), \infty] = 1\). We further specify \(F\) to be Pareto distributed of the form (3) with \(\alpha > 0\) and \(\sigma > 0\). Then all conditions in Theorem 3.3 are satisfied.
The various parameters are set to $\lambda_1 = 6$, $\lambda_2 = 7$, $\alpha = 0.7$ (or $\alpha = 1.5$), $\sigma = 1$, $\rho_X = 1.2$, $\theta_N = 1.5$ (or $\theta_N = 1.2$), $\rho_N = 2$. It can be calculated through software Mathematica that $\nu(A) \approx 1.7127$ if $\alpha = 0.7$ when estimating $\text{VaR}_q(S(t))$, and $\nu(A) \approx 2.68771$ if $\alpha = 1.5 > 1$ when estimating $\text{CTE}_q(S(t))$. The comparison of the simulated and asymptotic estimates for the VaR and CTE of total aggregate loss as well as their corresponding ratios are shown in Figures 4 and 5, respectively.

At the end of this subsection, we aim to investigate the influence of the AI and AD dependences between each pair of the severities to the behavior of the VaR of total aggregate loss. For simplicity, we assume that the two frequency processes follow a common Poisson process $N(t)$ with intensity $\lambda = 6$; the AI and AD dependences between each pair of the severities are specialized to the Frank and Gumbel copulas with parameters $\theta_X = \rho_X = 1.8$, respectively; and the severities follow a common Pareto distribution with parameters $\alpha = 0.5$ and $\sigma = 1$. By Remark 2, the coefficient $\gamma = \nu(A) E[N(t)] \approx 9.4213$ in (12).

**Figure 5.** Comparison of the simulated and asymptotic estimates for $\text{CTE}_q(S(t))$, with Gumbel dependent and Pareto distributed severities and AI or AD dependent frequency processes in Theorem 3.3

**Figure 6.** Comparison of the simulated and asymptotic estimates for $\text{VaR}_q(S(t))$, with common frequency process, and Gumbel or Frank dependent Pareto distributed severities in Theorems 3.2 and 3.3
The simulation results are shown in Figure 6. From Figure 6(a), it can be seen that both the simulated and asymptotic estimates for $\text{VaR}_q(S(t))$ under Gumbel dependent severities (i.e. the AD case) are smaller than those under Frank dependent severities (i.e. the AI case). This is reasonable, because the potential total aggregate loss for financial institutions is more likely to increase when a stable financial system turns to the economic downturns, and, hence, in a given time period and with a fixed confidence level $q$, the $q$-quantile of total aggregate loss in a crisis (i.e. the AD case) is smaller than that in a stable economic circumstance (i.e. the AI case). As a result, the $\text{VaR}_q(S(t))$ is sensitive to the AI and AD dependences between each pair of severities.

4. Concluding remarks. In this paper, we consider some dependent bivariate operational risk cell models, and derive some asymptotic formulas for the two risk measures Value-at-Risk and Conditional Tail Expectation of total aggregate loss as confidence level increases. A major contribution of the study is that each pair of the severities is allowed to be modelled not only by some weak dependence, but also by some strong dependence, and the frequency processes can be arbitrarily dependent. Our results can help to develop a very useful early-warning measure for operational risk, which are suitable for both a stable and an unstable economic environment. Another significant contribution of the paper is the demonstration that the severities are modelled through both extremely and moderately heavy-tailed distributions, which exhibit the high-severity characteristics of empirical data from banking and insurance industries, and include some commonly-used Pareto distribution, heavy-tailed Weibull distribution, Lognormal distribution, and among many others.

5. Proofs.

5.1. Some lemmas. Before proving our main results, we firstly cite a series of lemmas. The first lemma gives two important probabilistic inequalities for the tail probability of sum of i.i.d. random variables, which are the so-called Kesten’s bounds, and can be found in Lemma 1.3.5 of [19] and the proof of Lemma 4.5 of [32].

**Lemma 5.1.** Let $\{\xi_i, i \geq 1\}$ be a sequence of i.i.d. real-valued random variables with common distribution $V$. (1) If $V \in S$, then for any $\varepsilon > 0$, there exists some $C > 0$, such that for all $n \geq 1$ and $x \geq 0$, 

$$P\left(\sum_{i=1}^{n} \xi_i > x\right) \leq C(1 + \varepsilon)^n V(x).$$

(2) If $V \in \mathbb{R}_{-\alpha}$ for some $\alpha > 0$, then for any $p > \alpha$, there exists some $C > 0$ such that for all $n \geq 1$ and $x \geq 0$, 

$$P\left(\sum_{i=1}^{n} \xi_i > x\right) \leq Cn^p V(x).$$

The next two lemmas further investigate the asymptotic behavior for the tail probability of the sum of dependent random variables, which are derived from [22] and [14], respectively.
Lemma 5.2. Let \( \xi_1 \) and \( \xi_2 \) be two nonnegative random variables with distributions \( V_1 \) and \( V_2 \), respectively. Under Assumption 2.1, if \( V_1 \in \mathcal{S} \), \( \mathbb{V}_1(x) \sim \mathbb{V}_2(x) \) and \( V_1 * V_2 \in \mathcal{S} \), then
\[
P(\xi_1 + \xi_2 > x) \sim \mathbb{V}_1(x) + \mathbb{V}_2(x).
\]

Lemma 5.3. Let \( \xi_1, \ldots, \xi_n \) be \( n \) pairwise AI nonnegative random variables with distributions \( V_i \in \mathcal{R}_1 \alpha \) for all \( 1 \leq i \leq n \) and some \( \alpha > 0 \). Then, it holds that
\[
P \left( \sum_{i=1}^{n} \xi_i > x \right) \sim \sum_{i=1}^{n} V_i(x).
\]

The last lemma is due to [10].

Lemma 5.4. Let \( V_1 \) and \( V_2 \) be two distributions satisfying \( \mathbb{V}_1(x) \sim \mathbb{V}_2(x) \). If \( V_1 \in \mathcal{R}_1 \alpha \) for some \( 0 < \alpha \leq \infty \), then
\[
\left( \frac{1}{\mathbb{V}_1} \right)^{<}(x) \sim \left( \frac{1}{\mathbb{V}_2} \right)^{<}(x).
\]

5.2. Proofs of main results. In this subsection, we aim to prove our main results of this paper.

Proof of Proposition 1. We split the tail probability \( P(S(t) > x) \) according to whether \( N_1(t) \) or \( N_2(t) \) is larger:
\[
P(S(t) > x)
= \sum_{m=0}^{\infty} \sum_{n=m+1}^{\infty} P \left( \sum_{k=1}^{m} (X_k^{(1)} + X_k^{(2)}) + \sum_{l=m+1}^{n} X_l^{(2)} > x \right) P(N_1(t) = m, N_2(t) = n)
+ \sum_{m=0}^{\infty} \sum_{n=0}^{m} P \left( \sum_{k=1}^{m} (X_k^{(1)} + X_k^{(2)}) + \sum_{l=n+1}^{m} X_l^{(1)} > x \right) P(N_1(t) = m, N_2(t) = n)
= I_1 + I_2.
\]

(19)

We mainly deal with \( I_1 \). Clearly, \( F_1 * F_2 \in \mathcal{S} \) due to Corollary 3.17 of [21]. Then, by Lemma 5.2 we have
\[
P(X^{(1)} + X^{(2)} > x) \sim (c_1 + c_2) \mathcal{F}(x),
\]
which implies that the distribution of \( X^{(1)} + X^{(2)} \) is subexponential as well. For any fixed \( n > m \geq 0 \), since \( X_1^{(1)} + X_1^{(2)}, \ldots, X_m^{(1)} + X_m^{(2)}, X_{m+1}^{(2)}, \ldots, X_n^{(2)} \) are independent, by (20), \( F \in \mathcal{S} \) and Corollary 3.16 of [21], we have
\[
P \left( \sum_{k=1}^{m} (X_k^{(1)} + X_k^{(2)}) + \sum_{l=m+1}^{n} X_l^{(2)} > x \right)
\sim m P(X^{(1)} + X^{(2)} > x) + (n - m) c_2 \mathcal{F}(x)
\sim (c_1 m + c_2 n) \mathcal{F}(x).
\]

(21)

Again by (20), there exists some large \( x_1 > 0 \) such that for all \( x \geq x_1 \),
\[
P(X^{(1)} + X^{(2)} > x) \leq 2(c_1 + c_2) \mathcal{F}(x).
\]

Define a new positive random variable \( Y \) with tail distribution
\[
\mathbb{G}(x) = \begin{cases} 
2(c_1 + c_2) \mathcal{F}(x) & x \geq x_1 \\
1 & x < x_1.
\end{cases}
\]
Clearly, $G \in \mathcal{S}$ and $X^{(2)} \leq X^{(1)} + X^{(2)} \leq_{s.t.} Y$, where the symbol $\leq_{s.t.}$ denotes “stochastically not larger than”. Construct a sequence of i.i.d. positive random variables $\{Y_k, k \geq 1\}$ with common distribution $G$. By Lemma 5.1(1), for any $\varepsilon > 0$, there exists some $m_0 > 0$, by (23), (6), and according to Lemma 5.3, we can obtain our second result.

**Proof of Proposition 3.** The proof is similar to that of Proposition 1. We still use the decomposition (19). By $(X^{(1)}, X^{(2)}) \in \text{BRV}_{-\alpha}(F)$, we have

$$P(X^{(1)} + X^{(2)} > x) \sim \nu(A)F(x),$$

implying that the distribution of $X^{(1)} + X^{(2)}$ belongs to $\mathcal{R}_{-\alpha}$. For any fixed $n > m \geq 0$, by (23), (6), and according to Lemma 5.3,

$$P \left( \sum_{k=1}^{m} (X^{(1)}_k + X^{(2)}_k) + \sum_{l=m+1}^{n} X^{(2)}_l > x \right) \sim \left( m\nu(A) + (n - m)\nu((0, 1), \infty) \right)F(x).$$

Again by (23), there exists some large $x_1$ such that for all $x \geq x_1$,

$$P(X^{(1)} + X^{(2)} > x) \leq 2\nu(A)F(x).$$

Define a new positive random variable $Y$ with the tail distribution

$$G(x) = \begin{cases} 2\nu(A)F(x), & x \geq x_1 \\ 1, & x < x_1 \end{cases}.$$
Then, \( X^{(2)} \leq X^{(1)} + X^{(2)} \leq_{s.t.} Y \). Construct a sequence of i.i.d. positive random variables \( \{Y_k, k \geq 1\} \) with common distribution \( G \in R_\alpha \). By Lemma 5.1(2), similarly to (22), it holds that for all \( x \geq 0 \),

\[
\sum_{m=0}^{\infty} \sum_{n=m+1}^{\infty} P(\sum_{k=1}^{n} Y_k > x) \frac{P(N_1(t) = m, N_2(t) = n)}{G(x)} \leq C \sum_{m=0}^{\infty} \sum_{n=m+1}^{\infty} \sum_{p} P(N_1(t) = m, N_2(t) = n) \leq CE[(N_2(t))^p] < \infty.
\]

Combining (24), \( G(x) \sim 2\nu(A)F(x) \) and applying the dominated convergence theorem yields that

\[
\lim_{x \to \infty} \frac{I_1}{F(x)} = E\left[\left((\nu(A) - \nu((0,1),\infty))N_1(t) + \nu((0,1),\infty)N_2(t)\right)1_{(N_1(t) < N_2(t))}\right].
\]

Similarly,

\[
\lim_{x \to \infty} \frac{I_2}{F(x)} = E\left[\left(\nu((1,0),\infty)N_1(t) + (\nu(A) - \nu((1,0),\infty))N_2(t)\right)1_{(N_1(t) \geq N_2(t))}\right].
\]

Thus, the above two estimates lead to (11).

**Proof of Theorem 3.1.** By \( F \in R_{-\infty} \) and Lemma 5.4, Proposition 1 implies that as \( q \uparrow 1 \),

\[
\text{VaR}_q(S(t)) = \left(\frac{1}{G_t}\right)^\leftarrow \left(\frac{1}{1-q}\right) \sim \left(\frac{1}{F}\right)^\leftarrow \left(\frac{E[c_1N_1(t) + c_2N_2(t)]}{1-q}\right) \sim \left(\frac{1}{F}\right)^\leftarrow \left(\frac{1}{1-q}\right) = F^\leftarrow(q),
\]

where in the third step we used the fact \( \left(\frac{1}{F}\right)^\leftarrow \in RV_0 \) according to Theorem 2.4.7(ii) of [8].

**Proof of Theorem 3.2.** Along the line of the proof of Theorem 3.1, relation (14) follows immediately from Proposition 2 and Lemma 5.4.

As for (15), by \( \alpha > 1 \) and Karamata’s theorem, as \( q \uparrow 1 \),

\[
\text{CTE}_q(S(t)) = \text{VaR}_q(S(t)) + \int_{\text{VaR}_q(S(t))}^{\infty} \frac{P(S(t) > x)dx}{P(S(t) > \text{VaR}_q(S(t)))} \sim \frac{\alpha}{\alpha - 1} \text{VaR}_q(S(t)),
\]

as claimed.

**Proof of Theorem 3.3.** The proof can be done in the same way as that of Theorem 3.2 with slight modification.
Acknowledgments. We are very grateful to the referee for his/her constructive suggestions. This paper was supported by the Humanities and Social Sciences Foundation of the Ministry of Education of China (No. 20YJA910006), Natural Science Foundation of Jiangsu Province of China (No. BK20201396), Natural Science Foundation of the Jiangsu Higher Education Institutions (No. 19KJA180003), Postgraduate Education Reform Project of Jiangsu Province (No. JGLX19_091), the Project of Construction for Superior Subjects of Mathematics/Statistics of Jiangsu Higher Education Institutions.

REFERENCES

[1] C. Angela, R. Bisignani, G. Masala and M. Micocci, Advanced operational risk modelling in banks and insurance companies, Investment Management and Financial Innovations, 6 (2009), 78–83.

[2] A. V. Asimit and A. L. Badescu, Extremes on the discounted aggregate claims in a time dependent risk model, Scandinavian Actuarial Journal, (2010), 93–104.

[3] A. V. Asimit, E. Furman, Q. Tang and R. Vernic, Asymptotics for risk capital allocations based on conditional tail expectation, Insurance: Mathematics and Economics, 49 (2011), 310–324.

[4] A. L. Badescu, G. Lan, X. S. Lin and D. Tang, Modeling correlated frequencies with application in operational risk management, Journal of Operational Risk, 10 (2015), 1–43.

[5] S. A. Bakar, N. A. Hamzah, M. Maghsoudi and S. Nadarajah, Modeling loss data using composite models, Insurance: Mathematics and Economics, 61 (2015), 146–154.

[6] Basel Committee on Banking Supervision, International Convergence of Capital Measurement and Capital Standards, Report of Basel Committee on Banking Supervision, 2004.

[7] Basel Committee on Banking Supervision, Consultative Document, Fundamental Review of the Trading Book: A revised Market Risk framework, Report of Basel Committee on Banking Supervision, 2013.

[8] N. H. Bingham, C. M. Goldie and J. L. Teugels, Regular Variation, Cambridge university press, Cambridge, 1989.

[9] K. Böcker and C. Klüppelberg, Operational VaR: A closed-form approximation, Risk, 12 (2005), 90–93.

[10] K. Böcker and C. Klüppelberg, Multivariate models for operational risk, Quantitative Finance, 10 (2010), 855–869.

[11] E. Brechmann, C. Czado and S. Paterlini, Modeling dependence of operational loss frequencies, Journal of Operational Risk, 8 (2013), 105–126.

[12] N. Cantle, D. Clark, J. Kent and H. Verheugen, A brief overview of current approaches to operational risk under Solvency II, Milliman White Paper, (2012).

[13] V. Chavez-Demoulin, P. Embrechts and J. Neslehová, Quantitative models for operational risk: Extremes, dependence and aggregation, Journal of Banking & Finance, 30 (2006), 2635–2658.

[14] Y. Chen and K. C. Yuen, Sums of pairwise quasi-asymptotically independent random variables with consistent variation, Stochastic Models, 25 (2009), 76–89.

[15] K. Coorey and M. M. Ananda, Modelling actuarial data with a composite log-normal-Pareto model, Scandinavian Actuarial Journal, (2005), 321–334.

[16] P. de Fontnouvelle, D. Jesus-Rueff, J. S. Jordan and E. S. Rosengren, Using Loss Data to Quantify Operational Risk, Working Paper, 2003.

[17] L. de Haan and S. I. Resnick, On the observation closest to the origin, Stochastic Processes and their Applications, 11 (1981), 301–308.

[18] M. Eling Fitting insurance claims to skewed distributions: Are the skew-normal and skew-student good models?, Insurance: Mathematics and Economics, 51 (2012), 239–248.

[19] P. Embrechts, C. Klüppelberg and T. Mikosch Modelling Extremal Events for Insurance and Finance, Springer, Berlin, 1997.

[20] S. Emmer, M. Kratz and D. Tasche What is the best risk measure in practice? A comparison of standard measures, Journal of Risk, 18 (2015), 31–60.

[21] S. Foss, D. Korshunov and S. Zachary, An Introduction to Heavy-tailed and Subexponential Distributions, Springer, New York, 2011.
[22] J. Geluk and Q. Tang, Asymptotic tail probabilities of sums of dependent subexponential random variables, Journal of Theoretical Probability, 22 (2009), 871–882.
[23] P. Hartmann, S. Straetmans and C. G. de Vries, Heavy tails and currency crises, Journal of Empirical Finance, 17 (2010), 241–254.
[24] H. Joe and H. Li, Tail risk of multivariate regular variation, Methodology and Computing in Applied Probability, 13 (2011), 671–693.
[25] E. L. Lehmann, Some concepts of dependence, The Annals of Mathematical Statistics, 37 (1966), 1137–1153.
[26] J. Li, Q. Tang and R. Wu, Subexponential tails of discounted aggregate claims in a time-dependent renewal risk model, Advances in Applied Probability, 42 (2010), 1126–1146.
[27] A. J. McNeil, R. Frey and P. Embrechts, Quantitative Risk Management: Concepts, Techniques and Tools, Princeton University Press, Princeton, 2015.
[28] M. Moscadelli, The Modelling of Operational Risk: Experience with the Analysis of the Data Collected by the Basel Committee, Technical report by Bank of Italy, 2004.
[29] S. I. Resnick, Extreme Values, Regular Variation and Point Processes, Springer, New York, 1987.
[30] S. I. Resnick, Discussion of the Danish data on large fire insurance losses, ASTIN Bulletin, 27 (1997), 139–151.
[31] S. I. Resnick, Heavy-tail Phenomena: Probabilistic and Statistical Modeling, Springer, New York, 2007.
[32] Q. Tang, The finite-time ruin probability of the compound Poisson model with constant interest force, Journal of Applied Probability, 42 (2005), 608–619.
[33] Q. Tang, Z. Tang and Y. Yang, Sharp asymptotics for large portfolio losses under extreme risks, European Journal of Operational Research, 276 (2019), 710–722.
[34] Q. Tang and Y. Yang Interplay of insurance and financial risks in a stochastic environment, Scandinavian Actuarial Journal, (2019), 432–451.
[35] Q. Tang and Z. Yuan, Asymptotic analysis of the loss given default in the presence of multivariate regular variation, North American Actuarial Journal, 17 (2013), 253–271.
[36] Y. Yang, Y. Gong and J. Liu, Measuring tail operational risk in univariate and multivariate models under extreme losses, Working Paper, (2019).
[37] Y. Yang, T. Jiang, K. Wang and K. C. Yuen, Interplay of financial and insurance risks in dependent discrete-time risk models, Statistics & Probability Letters, 162 (2020), 108752, 11 pp.
[38] Y. Yang, K. Wang, J. Liu and Z. Zhang, Asymptotics for a bidual-dimensional risk model with two geometric Lévy price processes, Journal of Industrial & Management Optimization, 15 (2019), 481–505.
[39] Y. Yang, K. C. Yuen and J. Liu, Asymptotics for ruin probabilities in Lévy-driven risk models with heavy-tailed claims, Journal of Industrial & Management Optimization, 14 (2018), 231–247.
[40] X. Zhu, Y. Wang and J. Li, Operational risk measurement: A loss distribution approach with segmented dependence, Journal of Operational Risk, 14 (2019), 1–20.