

## EXPECTED SHORTFALL IN CREDIT PORTFOLIOS WITH EXTREMAL DEPENDENCE

Achal Bassamboo

Kellogg School of Management  
Northwestern University  
Evanston, IL 60208, U.S.A.

Sandeep Juneja

School of Technology and Computer Science  
Tata Institute of Fundamental Research,  
Mumbai, INDIA

Assaf Zeevi

Graduate School of Business,  
Columbia University,  
New York, NY 10027, U.S.A.

### ABSTRACT

We consider the risk of a portfolio comprised of loans, bonds, and financial instruments that are subject to possible default. We are interested in efficiently estimating expected excess loss conditioned on the event that the portfolio incurs large losses over a fixed time horizon; this risk measure is often referred to as *expected shortfall*. We consider a heterogeneous mix of obligors and assume a portfolio dependence structure that supports *extremal dependence* among obligors and does not hinge solely on correlation. We first derive sharp asymptotics that illustrate the implications of extremal dependence among obligors in the risk of the portfolio. Using this as a stepping stone, we develop a multi-stage importance sampling algorithm that is shown to have bounded relative error in estimating expected shortfall.

### 1 INTRODUCTION

Market conditions over the past few years combined with regulatory arbitrage have lead to significant interest and activity in trading and transferring of credit-related risk. Since most financial institutions are exposed to multiple sources of credit risk, a portfolio approach is needed to adequately measure and manage this risk. One of the most fundamental problems in this context is that of modeling dependence among a large number of obligors (consisting, for example, of companies to which a bank has extended credit), and assessing the impact of this dependence on the likelihood of multiple defaults and large losses.

A common framework for modeling a credit portfolio is the so-called latent variable approach in which dependence among obligors is captured through latent variables; the

latter often arise from factor analysis, and hence may be used to capture macroeconomic or industry-specific effects. The risk of default is then determined by the distance between the underlying variables and a given threshold. This methodology underlies essentially all models that descend from Merton's seminal firm-value work (Merton 1974).

The *normal copula model* which assumes that the latent variables follow a multivariate normal distribution is one of the most widely used models in practice. It has been incorporated into many popular risk management systems such as J.P. Morgan's CreditMetrics (Gupta, Finger, and Bhatia 1997), Moody's KMV system (Kealhofer and Bohn 2001), and is also prominently featured in the latest Basel accords that regulate capital allocation in banks (Basel Committee on Banking Supervision 2002); see also Li(2000) and the survey paper by Crouhy, Galai, and Mark (2000).

In recent years empirical work has argued that financial variables often exhibit stronger dependence than that captured in the correlation-based normal model. The stronger linkage is often manifested in large *joint* movements. In particular, in the credit risk context it has been argued that the main source of risk in large balanced loan portfolios is the occurrence of many joint defaults – what might be termed as “extreme credit risk.” These observations strongly suggest that in many instances the normal copula may not be an adequate way to model dependencies.

An attractive alternative to the normal model is one based on the multivariate Student *t* distributions, known as the *t-copula model*. While generalizing the normal copula model, the *t*-based model remain simple, parsimonious and analytically tractable. Recent work has shown that at least in certain instances this model provides a better fit to empirical financial data in comparison with the normal copula (e.g.,

Mashal and Zeevi 2003). It is important to note that unlike the normal copula the  $t$ -based model supports *extremal dependence* between the underlying variables. Roughly speaking, this means that variables may simultaneously take on very large values with non-negligible probability; for further discussion see Embrechts, Lindskog, and McNeil (2003). A useful interpretation of extremal dependence follows from the construction of a multivariate  $t$  distribution as a ratio of a multivariate normal and the square-root of a scaled Chi-squared random variable. When the denominator takes values close to zero, coordinates of the associated vector of  $t$ -distributed random variable may register large co-movements; see further discussion in Embrechts, Lindskog, and McNeil (2003) and Glasserman, Heidelberger and Shahabuddin (2002). Hence the Chi-squared random variable plays the role of a “common multiplicative shock.”

In this paper, we are interested in estimating expected shortfall, which is defined as the expected excess loss conditioned on the event that the loss exceeds a large threshold. Unlike value-at-risk (VaR), which is insensitive to the magnitude of loss beyond a certain percentile, ES weights large losses by their magnitude. The model that we consider builds on the latent variable approach, blending in a common multiplicative shock. The distributional assumptions are quite general and include as a particular instance the  $t$ -copula model discussed above. Our objective is twofold: to derive asymptotics for the expected shortfall; and to develop a provably efficient Importance Sampling (IS) algorithm for estimating it.

The main contributions of this paper include the following:

- We derive a sharp asymptotic for expected shortfall which illustrates in a precise manner the effects of extremal dependence (see Theorem 2).
- We construct an IS algorithm to efficiently estimate the risk of a portfolio via simulation. The algorithm uses a multi-stage exponential twist, and is shown to have bounded relative error as the number of obligors increase to infinity (see Theorem 3).

Numerical results illustrate the performance of the algorithm.

The paper is organized as follows. This section ends with a brief review of related literature, and section 2 describes the model. Section 3 and 4 contain our main results: the former derives the asymptotics and the latter describes the IS algorithm and investigates its performance. Section 5 presents numerical results. All proofs are relegated to the appendix.

## 2 RELATED LITERATURE

A recent paper by Glasserman and Li (2003) focuses on IS procedures for portfolio credit risk in the normal copula

model. Glasserman, Heidelberger and Shahabuddin (2002) describe efficient methods for computing equity portfolio VaR in the presence of heavy-tailed risk factors. They also argue that the proposed change of measure will be effective for estimating conditional loss. The point of departure of the current paper is the recent work of Bassamboo, Juneja, and Zeevi (2005) that develops a framework to model extremal dependence and in that setting derives IS algorithms for efficiently estimating the probability of large losses in heterogenous credit portfolios. In this paper we use their framework and consider the problem of estimating expected shortfall.

## 3 PROBLEM FORMULATION

### 3.1 The Portfolio Structure and Loss Distribution

Consider a portfolio of loans consisting of  $n$  obligors. Our interest centers on the distribution of losses from defaults over a fixed time horizon and the associated conditional mean losses above a certain threshold. The probability of default for the  $i$ th obligor over the time horizon of interest is  $p_i \in (0, 1)$ , and is used as an input to our model. This value is often set based on the average historical default frequency of companies with similar credit ratings. The associated exposure to default of counterparty  $i$  is assumed to be given by  $e_i > 0$ , that is, the default event results in a fixed and given loss of  $e_i$  monetary units. (We note that it is easy to generalize the main results of the paper to the case where the loss size is random under mild regularity conditions.)

For the determination of the portfolio loss distribution, the specification of dependence between defaults is of paramount importance. The dependence model that we consider is closely related to the widely used CreditMetrics model; see Gupta, Finger, Bhatia (1997), Crouhy, Galai, Mark (2000) and Li (2000). In particular, we assume that there exists a vector of underlying latent variables  $(X_1, \dots, X_n)$  so that the  $i$ th default occurs if  $X_i$  exceeds some given threshold  $x_i$  (the distributional assumptions related to the latent variables will be discussed in Section 3.2). The loss incurred from defaults is then given by

$$L = e_1 \mathbb{I}\{X_1 > x_1\} + \dots + e_n \mathbb{I}\{X_n > x_n\}, \quad (1)$$

where  $\mathbb{I}\{\cdot\}$  is the indicator function. The threshold  $x_i$  is chosen according to the marginal default probabilities so that  $\mathbb{P}(X_i > x_i) = p_i$ . Our main interest is in developing sharp asymptotics and efficient simulation techniques to estimate the expected shortfall of the portfolio, given by  $\mathbb{E}[L - x | L > x]$ , for a large threshold  $x$ .

The normal copula model that is widely used in the financial industry and that forms the basis of the CreditMetrics and KMV models assumes that the vector of latent

variable follows a multivariate normal distribution. Hence the dependence between the default events is determined by the correlation structure of the latent variables, in particular,  $(\mathbb{I}\{X_1 > x_1\}, \dots, \mathbb{I}\{X_n > x_n\})$  has a normal copula as its dependence structure (Embrechts, Lindskog, McNeil 2003). The underlying correlations are often specified through a linear factor model

$$X_i = c_{i1}Z_1 + \dots + c_{id}Z_d + c_i\eta_i,$$

where: i.)  $Z_1, \dots, Z_d$  are iid standard normal rv's that measure, for example, global, country and industry effects impacting all companies; ii.)  $c_{i1}, \dots, c_{id}$  are the loading factors; iii.)  $\eta_i$  is a normal rv that captures idiosyncratic risk, and is independent of the  $Z_i$ 's; and iv.)  $c_i$  and the loading factors are chosen so that the variance of  $X_i$  is equal to one (without loss of generality). To keep the notation simple, we restrict attention to single factor models ( $d = 1$ ). Extension to multi factor models is not difficult and follows along the lines outlined in Bassamboo, Juneja, Zeevi (2005).

The multivariate normal that underlies CreditMetrics/KMV provides a limited form of dependence between obligors, which, in particular, may not assign sufficient probability to the occurrence of many simultaneous defaults in the portfolio. As indicated in the introduction, one of the primary objectives of the current paper is to extend the normal copula model to incorporate "stronger" dependence among obligors, so that the corresponding dependence structure is more in line with recently proposed models of extremal dependence (e.g., Frey and McNeil 2001 and Embrechts, Lindskog, and McNeil 2003) and empirical findings (e.g., Mashal and Zeevi 2003), both of which suggest consideration of  $t$ -copula models and the like over the normal copula.

### 3.2 Extremal Dependence

Let  $(\eta_i : 1 \leq i \leq n)$  denote iid random variables and let  $Z$  denote another rv independent of  $(\eta_i : 1 \leq i \leq n)$ . Fix  $0 < \rho < 1$  and put

$$X_i = \frac{\rho Z + \sqrt{1 - \rho^2}\eta_i}{W}, \quad i = 1, \dots, n \quad (2)$$

where  $W$  is a non-negative rv independent of  $Z$  and  $(\eta_i : 1 \leq i \leq n)$  and its probability density function (pdf)  $f_W(\cdot)$  satisfies

$$f_W(w) = \alpha w^{\nu-1} + o(w^{\nu-1}) \quad \text{as } w \downarrow 0, \quad (3)$$

for some constants  $\alpha > 0$  and  $\nu > 0$ . Here and in what follows, we write  $h(x) = o(g(x))$  if  $h(x)/g(x) \rightarrow 0$  as  $x \rightarrow 0$  or as  $x \rightarrow \infty$ , where the limit considered is obvious from the context. If  $Z$  and  $\{\eta_i\}$  are iid having a

normal distribution and  $W$  is removed from (2), then this model reduces to a single factor latent variable instance of CreditMetrics/KMV. As alluded to earlier, our aim is to model economies where the dependence amongst obligor defaults is primarily due to *common shocks*, and this is captured in (3) through the random variable  $W$ . When  $W$  takes values close to zero, all the  $X_i$ 's are likely to be large leading to many simultaneous defaults. The parameter  $\nu$  measures the likelihood of common shocks: smaller values imply a higher probability that  $W$  takes values close to zero. This class of models has been recently proposed in the context of credit risk modeling (Frey and McNeil 2001) and references therein; in the particular instance where  $(Z, \eta)$  follow a bivariate normal distribution, this is often referred to as a mean-variance normal mixture, with  $1/W$  providing the mixing distribution.

**Example 1** Let  $W$  follow a Gamma( $\beta, \gamma$ ) distribution, with  $\gamma, \beta > 0$ , whose pdf is given by

$$f_W(x) = \frac{\beta^\gamma x^{\gamma-1}}{\Gamma(\gamma)} e^{-\beta x}, \quad x \geq 0.$$

Then this distribution satisfies (3) with  $\nu = \gamma$ ,  $\alpha = \beta^\gamma / \Gamma(\gamma)$ .

**Example 2** For a positive integer  $k$ , let  $W = \sqrt{k^{-1}\text{Gamma}(1/2, k/2)}$  so that

$$f_W(x) = \frac{2k^{k/2}x^{k-1}}{2^{k/2}\Gamma(k/2)} e^{-kx^2/2}, \quad x \geq 0.$$

This pdf satisfies (3) with  $\nu = k$ ,  $\alpha = 2(k/2)^{k/2} / \Gamma(k/2)$ .

Note that for  $\gamma = k/2$  and  $\beta = 1/2$ , the distribution discussed in Example 1 is Chi-squared with  $k$  degrees-of-freedom (df). Note that when a linear combination of  $Z$  and  $\eta_i$  follows a normal distribution and  $W$  has the distribution given in Example 2, then the vector  $(X_i : 1 \leq i \leq n)$  follows a multivariate  $t$ -distribution, whose dependence structure is given by a  $t$ -copula with  $k$  degrees of freedom.

**Technical assumptions:** Let  $F_Z(\cdot)$  and  $F_\eta(\cdot)$  denote the distribution functions of  $Z$  and  $\eta_i$ , respectively. In what follows we restrict  $Z$  to be light-tailed, i.e.,  $1 - F_Z(x)$  is upper bounded by an exponentially decaying term as  $x \rightarrow \infty$ . As far as the "noise" variable  $\eta$  is concerned, we make the following technical assumption: the distribution of  $\eta$  possesses a density which is positive on the real-line. (The latter assumption is made to facilitate analysis and can be generalized at the expense of further technical details.) In what follows we refer to (3) together with the above conditions collectively as *the distributional assumptions* associated with our model.

**4 ASYMPTOTIC ANALYSIS OF LOSS DISTRIBUTION AND EXPECTED SHORTFALL**

Since it is virtually impossible to exactly compute the portfolio loss distribution or expected shortfall, we focus on an asymptotic regime which is of practical interest and supports a tractable analysis. This regime is one where the portfolio of interest is comprised of a “large number” of obligors, each individual obligor defaults with “small” probability, and the focus is on “large” portfolio losses. The mathematical meaning of these terms is spelled out in section 4.1 and subsequently in section 4.2 we describe the main results.

**4.1 Preliminaries**

Let  $f(x)$  denote an increasing function so that  $f(x) \rightarrow \infty$  as  $x \rightarrow \infty$ . Fix  $n$  (the number of obligors in the portfolio), and let  $\{a_1, \dots, a_n\}$  be strictly positive constants. Set the default thresholds for the individual obligors to be  $x_i^n = a_i f(n)$ , so that obligor  $i$  defaults if  $X_i > a_i f(n)$  and obligors may have different marginal default probabilities. The overall portfolio loss is given by

$$L_n = e_1 \mathbb{I}\{X_1 > a_1 f(n)\} + \dots + e_n \mathbb{I}\{X_n > a_n f(n)\}, \quad (4)$$

where  $e_i, i = 1, \dots, n$ , is the exposure associated with the  $i$ th obligor. We are interested in studying the probability that  $L_n$  takes on large values when  $n$  is large, and use this to compute the expected shortfall. In particular, we focus on the probability of the event  $\{L_n > nb\}$ , and the conditional expectation  $\mathbb{E}[L_n - nb | L_n > nb]$  for  $b > 0$ . Hence as the size of the portfolio,  $n$ , grows large, the individual probability of default decreases in a manner that is governed by the function  $f(n)$ , and the loss level of interest,  $nb$ , scales up with the size of the portfolio.

We assume that  $f(n)$  increases at a sub-exponential rate so that  $f(n) \exp(-\beta n)$  is a bounded sequence that converges to zero as  $n \rightarrow \infty$  for all  $\beta > 0$ . By suitably selecting the function  $f(n)$  we can model portfolios of varying credit ratings classes. For example, letting  $f(n)$  increase polynomially in  $n$  we can model a portfolio with high quality obligors, while if  $f(n)$  increases, say, at a logarithmic rate, then the loans are considered more risky.

To deal with the heterogeneity among obligors, captured by the sequences  $\{e_i, a_i\}_{i=1}^n$ , we impose the following assumption.

**Assumption 1** *The non-negative sequence  $((e_i, a_i) : i \geq 1)$  takes values in a finite set  $\mathcal{V}$ , with cardinality  $|\mathcal{V}|$ . In addition, the proportion of each element  $(e_j, a_j) \in \mathcal{V}$  in the portfolio converges to  $q_j > 0$  as  $n \rightarrow \infty$  (so that  $\sum_{j \leq |\mathcal{V}|} q_j = 1$ ).*

In practice, the loan portfolio may be partitioned into a finite number of homogeneous loans based on factors such as industry, quality of risk, and exposure sizes. Assumption 1

allows this flexibility. While our analysis easily generalizes to the case where each obligor corresponds to the pair  $(e_j, a_j)$  with probability  $q_j$ , and  $e_j$  is a light tailed random variable, we avoid overburdening the notation by simply assuming a constant exposure level  $e_j$ , and that for a given portfolio a fraction  $q_j$  of the obligors correspond to class  $j$ . (In the remainder of the paper we ignore the non-integrality of  $q_j n$  for simplicity and clarity of exposition.)

**4.2 Sharp Asymptotics for the Probability of Large Portfolio Losses**

Let  $\bar{e} = \sum_{j \leq |\mathcal{V}|} e_j q_j$ , i.e., the limiting average loss when all the obligors default. Recall that the portfolio loss,  $L_n$ , is given in (4). The following theorem derived in Bassamboo, Juneja, and Zeevi (2005) gives a sharp asymptotic for the probability of large portfolio losses that provides an important building block in evaluating the asymptote of the expected shortfall. The function  $w(z)$  used in the statement of the theorem is defined precisely in Appendix A.1. Essentially, conditioned on  $Z = z$ ,  $w(z)$  denotes the threshold value so that for  $W \in (0, \frac{w(z)}{f(n)})$  the mean loss from the portfolio is greater than  $b$ ; for  $W \in (\frac{w(z)}{f(n)}, \infty)$ , the mean portfolio loss is less than  $b$ .

**Theorem 1** *Bassamboo, Juneja, and Zeevi 2005*  
Fix  $0 < b < \bar{e}$ , and let Assumption 1 as well as the distributional assumptions on  $(Z, \eta, W)$  hold true. Then,

$$\lim_{n \rightarrow \infty} f(n)^v \mathbb{P}(L_n > nb) = \frac{\alpha}{v} \int_{-\infty}^{\infty} w(z)^v dF_Z(z). \quad (5)$$

As alluded to above, this theorem is the key to establishing an asymptotic for conditional shortfall in Theorem 2. The function  $r(w, z)$  used in the statement of Theorem 2 is defined precisely in Appendix A.1. Essentially,  $r(w, z)$  denotes the mean loss from the portfolio conditioned on  $Z = z$  and  $W = w/f(n)$ . Let  $(Y)^+ := \max(0, Y)$ .

**Theorem 2** *Fix  $0 < b < \bar{e}$ , and suppose Assumption 1 as well as the distributional assumptions on  $(Z, \eta, W)$  hold true. Then*

$$\frac{\mathbb{E}[L_n - nb | L_n > nb]}{n} \rightarrow \frac{v \int_{-\infty}^{\infty} \int_0^{w(z)} (r(w, z) - b)^+ w^{v-1} dw dF_Z(z)}{\int_{-\infty}^{\infty} w(z)^v dF_Z(z)} \quad a.s.$$

as  $n \rightarrow \infty$ .

This asymptotic may be briefly understood as follows:

$$\mathbb{E}[L_n - nb | L_n > nb] = \frac{\mathbb{E}[(L_n - nb)I(L_n > nb)]}{P(L_n > nb)}.$$

The numerator may be asymptotically approximated by noting that the set of values of  $W$  and  $Z$ , for which the

mean portfolio loss is less than  $b$  contributes negligibly to it (because, in that region, the probability of  $\{L_n > nb\}$  decays exponentially with  $n$ ). On the remaining set, the portfolio loss amount may be replaced by its conditional expectation (conditioned on value of  $W$  and  $Z$ ) and since in this region  $W$  is small, its pdf may be approximated using (3). Similarly, the denominator may be asymptotically approximated by focussing on the region where conditioned on  $W$  and  $Z$ , the mean portfolio loss is  $> b$ , and in that region, approximating the conditional probability of  $\{L_n > nb\}$  by 1.

### 4.3 Numerical Illustration

In this section, we illustrate the accuracy of the asymptotic approximation for expected shortfall as the number of obligors becomes large. Table 1 compares the accuracy of the sharp asymptotic with the simulation based results as a function of  $n$ . Model parameters are taken to be  $\nu = 4$ ,  $f(n) = \sqrt{n}$ ,  $\rho = 0.25$ , each  $a_i = 0.5$  and  $b = 0.25$ . The accuracy improves significantly for large values of  $n$ . The cases where  $n = 100$  and  $250$  correspond to portfolio sizes of practical interest. Note that the asymptote is off by about 6-10% for these values. We may also expect the accuracy of the asymptotic to reduce for heterogeneous portfolios, especially when the exposures  $e_i$  are randomly distributed.

Table 1: The Expected Shortfall and Its Sharp Asymptotic as a Function of the Number of Obligor ( $n$ )

$n$	$\hat{\beta}(n, b)$ [95% C.I.]	Asymptote
100	5.4[±1.3%]	4.8
250	13.0[±1.3%]	12.3
500	24.9[±1.5%]	24.4
1000	48.8[±1.6%]	48.8
2000	95.3[±1.7%]	97

## 5 IMPORTANCE SAMPLING SIMULATION

The asymptotic presented in Theorem 2 may lead to significant inaccuracies in estimation of expected shortfall in many practically relevant cases. Hence Monte Carlo methods become an attractive alternative to accurately estimate this quantity. Since the event on which the expected shortfall is computed is typically small, naive simulation would require a very large number of runs to provide a satisfactory estimate. Importance sampling provides an efficient means of generating low variance estimates, essentially by placing further probability mass on the rare event of interest and then suitably unbiased the resultant simulation output.

### 5.1 Preliminaries

For notational convenience, assume that  $Z$  and  $W$  have probability density functions  $f_Z(\cdot)$  and  $f_W(\cdot)$ , respectively (though in our analysis we do not require that the distribution of  $Z$  have a density function). Let  $(p_j : j \leq |\mathcal{V}|)$  denote the probabilities associated with the Bernoulli variables  $(\mathbb{I}\{X_i > a_i f(n)\} : i \leq n)$ , as a function of the generated  $Z$  and  $W$ . We suppress this dependence from the notation for ease of presentation (this dependence is explicitly displayed in the proofs given in Appendix A). For notational purposes, let  $A_n = \{L_n > nb\}$  denote the event in which portfolio losses exceed a level  $nb$  in a portfolio with  $n$  obligors. Suppose that under an importance sampling distribution we generate samples of  $Z$ ,  $W$  and the Bernoulli variables  $(\mathbb{I}\{X_i > a_i f(n)\} : i \leq n)$ , and hence  $L_n$  and  $\mathbb{I}\{A_n\}$ , using density functions  $\tilde{f}_Z(\cdot)$ ,  $\tilde{f}_W(\cdot)$  and probabilities  $(\tilde{p}_j : j \leq |\mathcal{V}|)$ , where the distribution of  $W$  may depend upon the generated value of  $Z$ , and the distribution of the Bernoulli success probabilities may depend upon the generated values of  $Z$  and  $W$  (this dependence is also suppressed in the notation here). Let  $\tilde{\mathbb{P}}$  denote the corresponding probability measure. The sample output is then unbiased by multiplying it with the likelihood ratio  $\tilde{\mathcal{L}}$  that equals

$$\frac{f_Z(Z)f_W(W)}{\tilde{f}_Z(Z)\tilde{f}_W(W)} \prod_{j \leq |\mathcal{V}|} \left(\frac{p_j}{\tilde{p}_j}\right)^{Y_j q_j n} \left(\frac{1-p_j}{1-\tilde{p}_j}\right)^{(1-Y_j)q_j n},$$

where  $Y_j q_j n$  denotes the number of defaults in class  $j$  obligors.

We now discuss characterizations of performance for importance sampling estimators. To estimate the expected shortfall  $\mathbb{E}[L_n - nb | L_n > nb]$ , denoted by  $\beta(n, b)$ , we generate  $m$  iid samples  $((L_n^i, \tilde{\mathcal{L}}^i) : i \leq m)$  of  $(L_n, \tilde{\mathcal{L}})$  and compute the following estimate

$$\hat{\beta}_m(n, b) = \frac{\sum_{i=1}^m \tilde{\mathcal{L}}^i (L_n^i - nb) \mathbb{I}\{L_n^i > nb\}}{\sum_{i=1}^m \tilde{\mathcal{L}}^i \mathbb{I}\{L_n^i > nb\}}.$$

Using the delta-method (e.g., Serfling 1981) we note that the following central limit theorem holds:

$$\sqrt{m}[\hat{\beta}_m(n, b) - \beta(n, b)] \Rightarrow \sigma(n, b)\mathcal{N}(0, 1),$$

as  $m \rightarrow \infty$  where  $\Rightarrow$  denotes convergence in distribution, and

$$\sigma^2(n, b) = \frac{\sigma_1^2(n, b)}{\mu_2^2(n, b)} + \frac{\mu_1^2(n, b)\sigma_2^2(n, b)}{\mu_2^4(n, b)} + 2\frac{\sigma_{12}(nb)\mu_1(n, b)}{\mu_2^3(n, b)}, \quad (6)$$

with

$$\begin{aligned} \mu_1(n, b) &= \tilde{\mathbb{E}}[\tilde{\mathcal{L}}(L_n - nb)\mathbb{I}\{L_n > nb\}], \\ \mu_2(n, b) &= \tilde{\mathbb{E}}[\tilde{\mathcal{L}}\mathbb{I}\{L_n > nb\}], \\ \sigma_1^2(n, b) &= \tilde{\mathbb{E}}[(\tilde{\mathcal{L}}(L_n - nb))^2\mathbb{I}\{L_n > nb\}] - \mu_1^2(n, b), \\ \sigma_2^2(n, b) &= \tilde{\mathbb{E}}[\tilde{\mathcal{L}}^2\mathbb{I}\{L_n > nb\}] - \mu_2^2(n, b), \\ \sigma_{12}(n, b) &= \tilde{\mathbb{E}}[\tilde{\mathcal{L}}(L_n - nb)\mathbb{I}\{L_n > nb\}] - \mu_1(n, b)\mu_2(n, b). \end{aligned}$$

**Definition 1** We say that the change of measure has bounded relative error if

$$\limsup_{n \rightarrow \infty} \frac{\sigma(n, b)}{\beta(n, b)} < \infty.$$

The key to the occurrence of the large loss events in the portfolio corresponds to  $W$  taking small values so that the mean loss conditioned on  $W$  and  $Z$ , exceeds a level  $b$ . In Sections 5.2 we describe an IS algorithm for estimating the expected shortfall, that judiciously assigns large probability to event  $A_n$  to reduce simulation variance. The algorithm generates a new distribution of  $W$  by *exponentially twisting* the original one (see, e.g., Juneja and Shahabuddin (2005) for an introduction to exponential twisting). This is exactly the IS measure proposed by Bassamboo, Juneja, and Zeevi (2005) and it will be shown to have bounded relative error in estimating the expected shortfall.

When conditional on  $(W, Z)$  the mean loss is less than  $b$ , it may be a good practice (though not essential for the asymptotic optimality of the algorithm) to generate the corresponding Bernoulli random variables under an *exponentially twisted* distribution so that the event  $A_n$  is no longer rare, and the mean loss under the new distribution equals  $b$ . For any random variable  $X$  with pdf  $f_X(\cdot)$ , the associated distribution that is exponentially twisted by parameter  $\theta$  has the form

$$\exp(\theta x - \Lambda_X(\theta))f_X(x),$$

where  $\Lambda_X(\cdot)$  denotes the log-moment generating function of  $X$ . For  $\theta \geq 0$ , let  $\Lambda_j(\theta)$  denote  $\log(\exp(\theta e_j)p_j + (1 - p_j))$ . It is well known, and easily checked through differentiation, that  $\Lambda_j(\cdot)$  is strictly convex when  $0 < p_j < 1$  (e.g., Dembo and Zeitouni 1993). Let

$$p_j^\theta = \Lambda_j'(\theta) = \frac{\exp(\theta e_j)p_j}{\exp(\theta e_j)p_j + (1 - p_j)} = \frac{\exp(\theta e_j - \Lambda_j(\theta))p_j}{\exp(\theta e_j - \Lambda_j(\theta))p_j + (1 - p_j)},$$

where  $e_j$  is the exposure to the  $j$ th obligor, and  $p_j$  the probability that the  $j$ th obligor defaults. Put  $1 - p_j^\theta = \exp(-\Lambda_j(\theta))(1 - p_j)$ , and note that  $p_j^\theta$  is strictly increasing in  $\theta$ . For the case where the mean loss  $\sum_{j \leq |\mathcal{V}|} e_j q_j p_j < b$ ,

consider the new default probabilities  $(p_j^{\theta^*} : j \leq |\mathcal{V}|)$ , where  $\theta^* > 0$  is the unique solution to the equation

$$\sum_{j \leq |\mathcal{V}|} e_j q_j p_j^\theta = b.$$

This choice of twisting parameter induces a probability distribution under which the mean loss is  $b$ , hence the event of incurring such loss in a sample is no longer rare. In what follows we suppress the dependence of  $\theta^*$  on  $W$  and  $Z$ , in the notation, although it is noteworthy that  $\theta^*$  increases with  $W$  and decreases with  $Z$ .

### 5.2 The Importance Sampling Algorithm

This algorithm consists of three stages. First a sample of  $Z$  is generated using the original distribution. Depending on the value of  $Z$ , a sample of  $W$  is generated using appropriate importance sampling. Depending on the value of samples of  $Z$  and  $W$ , samples of the Bernoulli variable  $\mathbb{I}\{X_i > a_i f(n)\}$  are generated for  $i \leq n$ , using naive simulation or importance sampling. For a fixed positive constant  $\xi$ , put  $\tilde{w}(z) = \max(\xi, w(z))$ .

#### Importance Sampling Algorithm

1. Generate a sample of  $Z$  according to the original distribution  $F_Z(\cdot)$ .
2. Generate a sample of  $W$  using the density  $f_W^*$  obtained by exponentially twisting  $f_W$  with parameter  $-\theta_{Z,n}$ , where

$$\theta_{Z,n} = \frac{v f(n)}{\tilde{w}(Z)}.$$

This choice of the twisting parameter is based on the analysis in Bassamboo, Juneja, and Zeevi (2005).

3. For each  $i \leq n$ , generate samples of  $\mathbb{I}\{X_i > a_i f(n)\}$  independent of each other using the distribution:  $\tilde{p}_i = p_i$  if the mean loss under the generated  $W$  and  $Z$  is greater than  $b$ ; and using  $\tilde{p}_i = p_i^{\theta^*}$  otherwise.

Let  $\tilde{\mathbb{P}}$  denote the probability measure corresponding to this algorithm and  $\tilde{\mathbb{E}}$  the expectation operator under this measure. Again, let  $Y_j q_j n$  denote the number of class  $j$  defaults in a single simulation run. The likelihood ratio  $\tilde{\mathcal{L}}$  is then given by

$$\exp[\theta_{Z,n} W + \Lambda_W(-\theta_{Z,n})] \prod_{j \leq |\mathcal{V}|} \left( \frac{p_j}{\tilde{p}_j} \right)^{Y_j q_j n} \left( \frac{1 - p_j}{1 - \tilde{p}_j} \right)^{(1 - Y_j) q_j n}. \quad (7)$$

Now we state the main result of this section.

**Theorem 3** Under Assumption 1 and the distributional assumptions on  $(Z, \eta, W)$ , the proposed IS algorithm has bounded relative error for estimating the expected short-fall  $\beta(n, b)$ .

### 6 NUMERICAL RESULTS

In this section we compare the performance of the proposed algorithm with naive simulation, and investigate sensitivity to  $\nu$ . The broad conclusion is that the algorithm provide significant improvement over the performance of naive simulation. This improvement increases as the probability of large losses becomes more rare (e.g., as  $\nu$  increases). This supports our theoretical conclusion that the relative performance, as measured by the ratio of the standard deviation of the estimate to the mean of the estimate, remains well behaved for the algorithm even as the probability of large losses becomes increasingly rare.

Motivated by the  $t$ -copula model, we set the distribution of  $W$  in our numerical experiments as in Example 2, the random variable  $Z$  is chosen to follow a standard Normal distribution (mean zero, variance 1) and each  $\eta_i$  is normally distributed with mean 0 and variance 9. (We set the value of variance to 9 instead of 1 simply to ensure that the loss probability is sufficiently large to be practically relevant). The random variables  $W, Z$  and  $(\eta_i : i \leq n)$  are mutually independent so that  $X = (X_1, \dots, X_n)$  has a multi-dimensional Student  $t$ -distribution, with the dependence structure given by a  $t$ -copula. Implementation issues concerning this algorithm are discussed in Bassamboo, Juneja, and Zeevi (2005).

#### 6.1 Performance of the Proposed Algorithm

Table 2 compares the performance of the IS Algorithm with naive simulation as  $\nu$  varies. The model parameters are chosen to be  $n = 250, f(n) = \sqrt{n}, \rho = 0.25, b = 0.25$ , each  $a_i = 0.5$  and  $e_i = 1$ . For each  $\nu$ , we generate 50,000 samples under the original measure and IS measure. We then compute the variance reduction obtained under the two algorithms. This equals the ratio of the variance of the estimator under the importance sampling measure to the variance of the estimator under the original measure. We also report the probability of large loss, i.e.,  $\mathbb{P}(L_n > nb)$ . For  $df = 12$  and  $df = 16$ , we observed  $L_n < nb$  under naive simulation for all the 50, 000 sample paths generated.

Table 2: Performance of the IS Algorithm as a Function of  $\nu$ . Variance Reduction is Measured Relative to Naive Simulation. (\*) Denotes That the Event of Interest was not Observed in Any Sample Path Using Naive Simulation.

$df$	$\hat{\beta}(n, b)[95\% \text{ C.I.}]$	Var. red.	$\mathbb{P}(L_n > nb)$
4	13.20[±1.5%]	62	$8.06 \times 10^{-3}$
8	7.84[±2.6%]	743	$2.41 \times 10^{-4}$
12	5.81[±4.1%]	(*)	$1.07 \times 10^{-5}$
16	4.67[±6.9%]	(*)	$6.18 \times 10^{-7}$

### A PROOFS OF MAIN RESULTS

#### A.1 Preliminaries

We first introduce some preliminary notation and observations that are useful in proving the main theorems. Let

$$\begin{aligned}
 p_{w,z,i} &:= \mathbb{P}\left(\eta > \frac{a_i W f(n) - \rho Z}{\sqrt{1 - \rho^2}} \mid W = \frac{w}{f(n)}, Z = z\right) \\
 &= \mathbb{P}\left(\eta > \frac{a_i w - \rho z}{\sqrt{1 - \rho^2}}\right).
 \end{aligned}$$

Let  $\mathbb{P}_{w,z}$  denote the probability measure conditioned on the event  $W = w/f(n)$  and  $Z = z$  and let  $\mathbb{E}_{w,z}$  denote the corresponding expectation operator. Let

$$\begin{aligned}
 r(w, z) &:= \sum_{j \leq |\mathcal{V}|} e_j q_j p_{w,z,j} \\
 &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n e_i p_{w,z,i},
 \end{aligned} \tag{8}$$

where the limit follows from Assumption 1. For  $w > 0, r(w, z)$  denotes the limiting average portfolio loss (as  $n \rightarrow \infty$ ), and also the mean portfolio loss when  $W = \frac{w}{f(n)}$  and  $Z = z$ . Note that  $r(w, z)$  is non-decreasing in  $z$  and non-increasing in  $w$ .

Let  $z_b$  denote the unique value of  $z$  that solves

$$\bar{e}\mathbb{P}\left(\eta \geq \frac{-\rho z}{\sqrt{1 - \rho^2}}\right) = b.$$

(Note that our assumption that  $\eta$  has a positive density function on the real line ensures that there exists a unique  $z_b$  that solves the above equation.) The term  $z_b$  assumes significance in our analysis since for  $Z < z_b$  the event of average loss exceeding  $b$  remains a rare event for all values  $W > 0$ . For  $z \geq z_b$ , let  $w(z)$  be defined as the unique solution to

$$r(w, z) = b. \tag{9}$$

Note that  $w(z)$  is strictly positive for each  $z > z_b$ . Note also that for  $w \leq w(z)$ , under  $\mathbb{P}_{w,z}$  the average loss amount  $\frac{1}{n} \sum_{i=1}^n e_i \mathbb{I}\{X_i > a_i f(n)\}$  in the limit as  $n \rightarrow \infty$  has mean which is greater than or equal to  $b$ , and hence the probability of large loss is no longer a rare event. Set  $w(z) = 0$  for  $z \leq z_b$ .

To perform asymptotic analysis, we need additional notation obtained by perturbing certain parameters. For each  $\delta$ , let  $z_{b_\delta}$  denote the unique solution to

$$\bar{e} \mathbb{P} \left( \eta \geq \frac{-\rho z}{\sqrt{1-\rho^2}} \right) = b - \delta.$$

Note that  $z_{b_0} \equiv z_b$ , and  $z_{b_\delta}$  is a decreasing function of  $\delta$ . Further, we have  $z_{b_\delta} \rightarrow z_b$  as  $\delta \rightarrow 0$ . Let  $w_\delta(z) \geq 0$  denote the unique solution to the equation  $r(w, z) = b - \delta$  for  $z \geq z_{b_\delta}$ . Note that  $w(z) = w_0(z)$ ,  $w_\delta(z)$  is a strictly increasing function of  $z$  for  $z \geq z_{b_\delta}$ , and using continuity and monotonicity of  $r(w, z)$  in  $w$ , we have

$$w_\delta(z) \rightarrow w(z) \tag{10}$$

as  $\delta \rightarrow 0$ . The following upper bound on  $w_\delta(z)$  is easily seen and is useful to our analysis:

$$w_\delta(z) \leq \frac{\rho}{\min_i a_i} (z - z_{b_\delta}) \quad \text{for all } z > z_{b_\delta}. \tag{11}$$

### A.2 Proof of Theorem 2.

Using Theorem 1, it suffices to show that

$$\begin{aligned} & \frac{f(n)^v}{n} \mathbb{E}[(L_n - nb)^+] \\ & \rightarrow \alpha \int_{z_b}^\infty \int_0^{w(z)} \left( \sum_{j=1}^{|\mathcal{V}|} e_j q_j p_{w,z,j} - b \right) \\ & \quad \times w^{v-1} dw dF_Z(z), \end{aligned} \tag{12}$$

as  $n \rightarrow \infty$ . Here,  $(Y)^+ := \max(0, Y)$ .

Fix  $\delta > 0$ . We decompose the left hand side of (12) into the following three terms

$$\begin{aligned} & \mathbb{E}[(L_n - nb)^+] = \mathbb{E}[(L_n - nb)^+ \mathbb{I}\{z \leq z_b\}] \\ & \quad + \mathbb{E} \left[ (L_n - nb) \mathbb{I} \left\{ W > \frac{w_\delta(Z)}{f(n)}, Z \geq z_b \right\} \right] \\ & \quad + \mathbb{E} \left[ (L_n - nb) \mathbb{I} \left\{ W \leq \frac{w_\delta(Z)}{f(n)}, Z > z_b \right\} \right]. \end{aligned} \tag{13}$$

We divide the remaining part of the proof into four steps. The first and the second step show that the first and second term on the right hand side of (13), respectively, are asymptotically negligible. The third and the fourth step develop upper and

lower bound on the third term on the right-hand-side of (13).

**Step 1.** We show that

$$\lim_{n \rightarrow \infty} \frac{f(n)^v}{n} \mathbb{E}[(L_n - nb)^+ \mathbb{I}\{Z \leq z_b\}] = 0. \tag{14}$$

Note that

$$\frac{(L_n - nb)}{n} < (\max_i e_i - b).$$

Thus, we have

$$\begin{aligned} & \frac{f(n)^v}{n} \mathbb{E}[(L_n - nb)^+ \mathbb{I}\{Z \leq z_b\}] \\ & = \mathbb{E}[(L_n - nb) \mathbb{I}\{L_n > nb, Z \leq z_b\}] \\ & \leq f(n)^v (\max_i e_i - b) \mathbb{P}(L_n > nb, Z \leq z_b). \end{aligned}$$

The assertion in (14) now follows from Step 1 of proof of Bassamboo, Juneja, and Zeevi (2005), Theorem 1.

**Step 2.** We show that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{f(n)^v}{n} \mathbb{E} \left[ (L_n - nb)^+ \mathbb{I} \left\{ W > \frac{w_\delta(Z)}{f(n)}, Z \geq z_b \right\} \right] \\ & = 0. \end{aligned} \tag{15}$$

As in Step 1, the left hand side is bounded above by

$$\begin{aligned} & \lim_{n \rightarrow \infty} (\max_i e_i - b) f(n)^v \\ & \quad \mathbb{P} \left( L_n > nb, W > \frac{w_\delta(Z)}{f(n)}, Z \geq z_b \right), \end{aligned}$$

which by Step 2 of the proof of Bassamboo, Juneja, and Zeevi (2005), Theorem 1, gives (15).

**Step 3.** We now develop an asymptotic upper bound on the third term on the right hand side of (13), which in turn gives an asymptotic upper bound on  $\mathbb{E}[L_n - nb]$ . To this end, we show that for  $\delta > 0$ ,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{f(n)^v}{n} \mathbb{E} \left[ (L_n - nb)^+ \mathbb{I} \left\{ W \leq \frac{w_\delta(Z)}{f(n)}, Z > z_{b_0} \right\} \right] \\ & \leq \alpha \int_{z_b}^\infty \int_0^{w(z)} \left( \sum_{j=1}^{|\mathcal{V}|} e_j q_j p_{w,z,j} - b \right) w^{v-1} dw dF_Z(z). \end{aligned}$$

To see this, note that

$$\begin{aligned} & \mathbb{E} \left[ (L_n - nb)^+ \mathbb{I} \left\{ W \leq \frac{w_\delta(Z)}{f(n)}, Z > z_{b_0} \right\} \right] \\ & = \int_{z_b}^\infty \int_0^{\frac{w_\delta(Z)}{f(n)}} \mathbb{E}_{f(n)w,z} [(L_n - nb)^+] f_W(w) dF_Z(z). \end{aligned} \tag{16}$$

For any  $0 < \kappa < 1$ , this is upper bounded by

$$\int_{z_b}^{f(n)^\kappa} \int_0^{\frac{w_\delta(z)}{f(n)}} \mathbb{E}_{f(n)w,z}[(L_n - nb)^+] f_W(w) dw dF_Z(z) \quad (17)$$

$$+ \mathbb{E}[(L_n - nb)^+] \mathbb{I}\{Z \geq f(n)^\kappa\}.$$

Note from (3) that for any  $\epsilon > 0$  there exists  $n$  sufficiently large such that  $f_W(w) \leq \alpha(1 + \epsilon)w^{v-1}$  for  $0 \leq w \leq \frac{w_\delta(z)}{f(n)}$  and  $z \in (z_b, f(n)^\kappa)$ . (This follows since  $w_\delta(z)$  increases at most at a linear rate as a function of  $z$ ). Thus, for sufficiently large  $n$ , (17) is upper bounded by

$$\alpha(1 + \epsilon) \int_{z_b}^\infty \int_0^{\frac{w_\delta(z)}{f(n)}} \mathbb{E}_{w,z}[(L_n - nb)^+] w^{v-1} dw dF_Z(z)$$

$$+ n(\max_i e_i - b) \mathbb{P}(Z \geq f(n)^\kappa).$$

The last term multiplied by  $\frac{f(n)^v}{n}$  vanishes in the limit as

$$\lim_{n \rightarrow \infty} f(n)^v \mathbb{P}(Z \geq f(n)^\kappa) = 0.$$

Next consider the first term, changing the variable and letting  $y = wf(n)$  we get

$$\frac{\alpha(1 + \epsilon)}{f(n)^v} \int_{z_b}^\infty \int_0^{w_\delta(z)} \mathbb{E}_{y,z}[(L_n - nb)] y^{v-1} dy dF_Z(z).$$

The desired upper bound follows by multiplying the above by  $f(n)^v/n$ , taking limits as  $n \rightarrow \infty$ , noting that  $\epsilon$  is arbitrary,  $L_n/n$  is bounded, and the fact that

$$\lim_{n \rightarrow \infty} \mathbb{E}_{y,z} \left[ \left( \frac{L_n}{n} - nb \right)^+ \right] = \left( \sum_{j=1}^{|\mathcal{V}|} e_j q_j p_{y,z,j} - b \right)^+.$$

Using the above three steps together with (13) establishes that

$$\lim_{n \rightarrow \infty} \frac{f(n)^v}{n} \mathbb{E}[(L_n - nb)^+]$$

$$\leq \alpha \int_{z_b}^\infty \int_0^{w_\delta(z)} \left( \sum_{j=1}^{|\mathcal{V}|} e_j q_j p_{w,z,j} - b \right)$$

$$\times w^{v-1} dw dF_Z(z).$$

Note that the left hand side is independent of  $\delta$ ;  $w_\delta(z)$  is bounded from above by a linear function in  $z$ ;  $w_\delta(z) \rightarrow w(z)$  as  $\delta \rightarrow 0$ ; and  $Z$  is light tailed. Using the dominated convergence theorem when letting  $\delta \rightarrow 0$ , we deduce that

$$\lim_{n \rightarrow \infty} \frac{f(n)^v}{n} \mathbb{E}[(L_n - nb)^+]$$

$$\leq \alpha \int_{z_b}^\infty \int_0^{w(z)} \left( \sum_{j=1}^{|\mathcal{V}|} e_j q_j p_{w,z,j} - b \right) w^{v-1} dw dF_Z(z).$$

**Step 4.** We now prove the following lower bound

$$\liminf_{n \rightarrow \infty} \frac{f(n)^v}{n} \mathbb{E}[(L_n - nb)^+]$$

$$\times \mathbb{I} \left\{ W \leq \frac{w_\delta(Z)}{f(n)}, Z > z_{b_0} \right\}$$

$$\geq \alpha \int_{z_b}^\infty \int_0^{w(z)} \left( \sum_{j=1}^{|\mathcal{V}|} e_j q_j p_{w,z,j} - b \right)$$

$$w^{v-1} dw dF_Z(z). \quad (18)$$

To see this, note that for a given  $\tilde{\delta} > 0$ , there exists  $N$  such that  $\mathbb{E}_{w,z}[L_n] \geq n \left( \sum_{j=1}^{|\mathcal{V}|} e_j p_{w,z,j} - \tilde{\delta} \right)$  for all  $n > N$ . Thus, we have that the left-hand-side of (18) is lower bounded by

$$\liminf_{n \rightarrow \infty} f(n)^v \int_{z_b}^\infty \int_0^{\frac{w(z)}{f(n)}} \left( \sum_{j=1}^{|\mathcal{V}|} p_{f(n)w,z,j} e_j - b - \tilde{\delta} \right)$$

$$\times f_W(w) dw dF_Z(z)$$

$$\geq \alpha(1 - \epsilon)$$

$$\times \liminf_{n \rightarrow \infty} f(n)^v \int_{z_b}^\infty \int_0^{\frac{w(z)}{f(n)}} \left( \sum_{j=1}^{|\mathcal{V}|} p_{f(n)w,z,j} e_j - b - \tilde{\delta} \right)$$

$$\times w^{v-1} dw dF_Z(z),$$

for any  $\epsilon > 0$ . The last inequality follows from (3). Let  $y = f(n)w$ . Thus the above expression equals

$$\alpha(1 - \epsilon) \int_{z_b}^\infty \int_0^{w(z)} \left( \sum_{j=1}^{|\mathcal{V}|} e_j q_j p_{y,z,j} - b - \tilde{\delta} \right)$$

$$\times y^{v-1} dy dF_Z(z).$$

Taking limits as  $\epsilon \rightarrow 0$  and  $\tilde{\delta} \rightarrow 0$ , we get the desired result. This completes the proof. ■

### A.3 Proof of Theorem 3.

Using Theorem 2, it suffices to prove that

$$\limsup_{n \rightarrow \infty} \frac{\sigma^2(n, b)}{n^2} < \infty.$$

To this end, we will prove that each term on the right-hand-side of (6) scaled by  $n^{2\nu}$  is finite. Consider the first term on the right-hand-side of (6). We first observe that

$$\limsup_{n \rightarrow \infty} \frac{f(n)^{2\nu}}{n^2} \tilde{\mathbb{E}}[L^2(L_n - nb)^2 \mathbb{I}\{L_n > nb\}] < \infty. \quad (19)$$

To see this, note that  $(L_n - nb)^2 \leq (\max_j e_j n)^2$ . Also,

$$\limsup_{n \rightarrow \infty} f(n)^{2\nu} \tilde{\mathbb{E}}[\tilde{\mathcal{L}}^2 \mathbb{I}\{L_n > nb\}] < \infty$$

follows from Theorem 2 of Bassamboo, Juneja, and Zeevi (2005) which states that the proposed algorithm has bounded relative error for estimating  $\mathbb{P}(L_n > nb)$ . Thus, we have

$$\limsup_{n \rightarrow \infty} \frac{\sigma_1^2(n, b)}{n^2 \mu_2^2(n, b)} < \infty,$$

since

$$\limsup_{n \rightarrow \infty} \frac{\sigma_1^2(n, b) f(n)^{2\nu}}{n^2} < \infty, \quad (20)$$

and

$$\limsup_{n \rightarrow \infty} \mu_2^2(n, b) f(n)^{2\nu} < \infty. \quad (21)$$

Here (20) follows from (19) and (21) follows from Theorem 1. Similarly,

$$\limsup_{n \rightarrow \infty} \frac{\mu_1^2(n, b) \sigma_2^2(n, b)}{n^2 \mu_2^4(n, b)} < \infty.$$

For the last term, note that

$$\limsup_{n \rightarrow \infty} \frac{f(n)^{2\nu} \tilde{\mathbb{E}}[(\tilde{\mathcal{L}})^2(L_n - nb) \mathbb{I}\{L_n > nb\}]}{n} < \infty.$$

Therefore,  $\limsup_{n \rightarrow \infty} \frac{\sigma_{12}(nb) \mu_1(n, b)}{n^2 \mu_2^3(n, b)} < \infty$ , and the proof is complete. ■

## REFERENCES

- Basel Committee on Banking Supervision. 2002. The New Basel Capital Accord, Basel.
- Bassamboo, A., S. Juneja, and A. Zeevi. 2005. Portfolio credit risk with external dependence. Working paper, Stanford University.
- Crouhy, M., D. Galai, and R. Mark. 2000. A comparative analysis of current credit risk models. *J. of Banking and Finance* 24: 59–117.
- Dembo, A. and O. Zeitouni. 1993. *Large Deviations Techniques and Applications*. Jones and Bartlett Publishers.
- Embrechts, P., F. Lindskog, and A. McNeil. 2003. Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance*, ed. S. Rachev. Elsevier, NY.
- Feller, W. 1970. *An Introduction to Probability Theory and Its Applications, Volume 1*. John Wiley & Sons.
- Frey, R. and A. McNeil. 2001. Modelling dependent defaults. Working paper, Universität and ETH Zürich.
- Glasserman, P., P. Heidelberger, and P. Shahabuddin. 2002. Portfolio value-at-risk with heavy-tailed risk factors. *Math. Finance* 12: 239–269.
- Glasserman, P. and J. Li. 2003. Importance sampling for portfolio credit risk. Working paper, Columbia University.
- Gupta, G., C. Finger, and M. Bhatia. 1997. Credit metrics technical document. Technical report, J.P. Morgan & Co., New York.
- Juneja, S. and P. Shahabuddin. 2005. Rare event simulation techniques: an introduction and recent advances. To appear in *Handbook on Simulation*, ed. S. Henderson and B. Nelson.
- Kealhofer, S. and J. Bohn. 2001. Portfolio management of credit risk. Technical report, KMV working paper, New York.
- Li, D. 2000. On default correlation: a copula function approach. *J. of Fixed Income* 9: 43–54.
- Mashal, R. and A. Zeevi. 2003. Beyond correlation: extreme co-movements between financial assets. Working paper, Columbia University.
- Merton, R. 1974. On the pricing of corporate debt: the risk structure of interest rates. *J. of Finance* 29: 449–470.
- Serfling, R. J. 1981. *Approximation Theorems of Mathematical Statistics*. Wiley-Interscience.

## AUTHOR BIOGRAPHIES

**ACHAL BASSAMBOO** is Donald P. Jacobs Scholar in managerial economics and decision sciences at Kellogg School of Management. His research interests are stochastic systems, revenue management and rare event simulation. His e-mail address is <a-bassamboo@northwestern.edu>.

**SANDEEP JUNEJA** is a faculty in the School of Technology and Computer Science at the Tata Institute of Fundamental Research. His research interests include Monte-Carlo methods for stochastic systems. His e-mail address is <juneja@tifr.res.in> and his web page is <www.tcs.tifr.res.in/~sandeepj>.

**ASSAF ZEEVI** is Nathan Gantcher Associate Professor of Business in the Graduate School of Business, Columbia University. His research focuses on stochastic models of service systems, with other recent research on problems in financial economics, simulation, statistics and applied probability. His email address is <assaf@gsb.columbia.edu>.