COMPUTATION OF COMPOUND DISTRIBUTIONS II: DISCRETIZATION ERRORS AND RICHARDSON EXTRAPOLATION

BY

RUDOLF GRÜBEL AND RENATE HERMESMEIER

Universität Hannover

ABSTRACT

The standard methods for the calculation of total claim size distributions and ruin probabilities, Panjer recursion and algorithms based on transforms, both apply to lattice-type distributions only and therefore require an initial discretization step if continuous distribution functions are of interest. We discuss the associated discretization error and show that it can often be reduced substantially by an extrapolation technique.

KEYWORDS AND PHRASES

Total claim size distribution, ruin probabilities, discretization, acceleration of convergence.

1. INTRODUCTION

As in the first part of the paper we are interested in the numerical evaluation of a compound distribution $v = \sum_{k=0}^{\infty} p_k \mu^{*k}$, where we regard μ and $(p_k)_k \in \mathbb{N}_0$ as given and "$*$" denotes convolution. This situation arises in the standard risk theory model, where v is the distribution of the total claim size, p_k is the probability that there are k claims within the period of interest, and μ is the distribution of the individual claims: see e.g. Asmussen (1987), Beard et al. (1984), Grandell (1991) and Hipp and Michel (1990). We have $p_k = \exp(-\lambda)\lambda^k/k!$ for some $\lambda > 0$ if the claim arrival process is a Poisson process. Some additional arguments show that the ruin probabilities in the classical risk model, regarded as a function of the initial risk reserve, also fit into this framework, with $p_k = (1-\rho)^k \rho$ for some $\rho \in (0,1)$. If μ is one-sided and of lattice type, i.e. concentrated on the non-negative integer multiples of some $h > 0$, then both the compound Poisson and the compound geometric case can be handled numerically by Panjer recursion. Transform methods, in conjunction with the fast Fourier transform algorithm (FFT), apply in the

lattice case for two-sided distributions and quite general compounding sequences $(p_k)_{k\in\mathbb{N}_0}$; for a more detailed discussion and references see Part I of the present paper, Grübel and Hermesmeier (1999). In Part I it was shown that the aliasing errors that arise in connection with transform methods and are absent in Panjer recursion, can be reduced by a suitable change of the measure μ ("exponential tilting"). If μ not of lattice type, in particular if μ has a density, the use of Panjer recursion or transform methods requires an initial discretization step, which leads to a discretization error. In the present second part of the paper we show that *Richardson extrapolation*, also known as *extrapolation to the limit* or *acceleration of convergence*, can be used in many situations to reduce such discretization errors considerably.

Extrapolation methods are in the standard numerical toolbox and explained in detail in most numerical analysis textbooks, see e.g. Section 3.5 in Stoer (1994). Walz (1996) is a recent monograph on this subject, which also contains some interesting historical references; see also Section 8 in Körner (1996). The basic idea is simply this: if, for example, some real number y_0 cannot be evaluated directly but some approximation y_h is amenable for all $h > 0$, then the approximations obtained for various h-values can be combined into a better approximation if the rate of convergence of y_h to y_0 with $h \to 0$ is known. If, for example,

$$y_h = y_0 + ch^\alpha + o(h^\beta) \tag{1.1}$$

as $h \to 0$ with some constants α, β satisfying $0 < \alpha \le \beta$, then

$$\tilde{y}_h := \frac{1}{2^\alpha - 1}(2^\alpha y_{h/2} - y_h) = y_0 + o(h^\beta), \tag{1.2}$$

i.e. combining the approximations obtained for h and $h/2$ results in an approximation with an improved rate of convergence – the convergence has been accelerated. A crucial aspect of this simple idea is the fact that in passing from (1.1) to (1.2) we only need to know the value of α, but not of c. In applications, and that includes the one we are interested in, α is often known from some general qualitative considerations whereas c is difficult to obtain or even involves the quantity y_0. Of course, an actual acceleration only occurs if $c \neq 0$. Further, one can obviously extend this by combining more than two approximations if additional expansion terms in (1.1) are known.

Extrapolation to the limit underlies Romberg integration and is also a standard method in connection with the numerical treatment of differential equations where h can for example arise as the width of a grid when replacing a differential equation by a difference equation. Bohman (1977) used an extrapolation in the context of numerical inversion of characteristic functions. Here we consider the approximation v_h obtained for the compound distribution v with lattice-type methods if μ is replaced by some μ_h concentrated on the integer multiples of h. The use of extrapolation methods in this context seems to be new, but a "weak" result where integrals $\int \phi dv$ rather than the distribution v itself are considered, was obtained in Embrechts, Grübel and Pitts (1993). The

proof given there for a result of type (1.1) is entirely in the ϵ-δ-style. The next section contains a heuristic approach to that result that is based on the idea of regarding a stochastic model as a non-linear operator, an idea that was also useful in the context of analyzing aliasing errors in the first part of the present paper. The third section contains our main result, where (1.1) is extended to distributions. Whereas the weak result deals with real numbers and therefore easily fits into the above description via (1.1) and (1.2) we now have a more complicated situation, where measures rather than real numbers arise. This requires a suitable notion of distance. The proof of our main result is somewhat technical and therefore put away into an appendix, but the approach explained in Section 2 can serve as a guideline. Section 4 deals with boundary effects and numerical examples are given in Section 5.

2. A HEURISTIC APPROACH

The investigation of total claim size distributions and ruin probabilities in the risk model as quantities depending on the individual claim size distribution, the claim arrival intensity and the premium income rate can be seen as a special instance of a very general framework, where a stochastic model is regarded as a functional that relates an output quantity μ_{out} of interest to some known input quantity μ_{in}. In mathematical notation the model can be represented by a non-linear operator Ψ so that $\mu_{\text{out}} = \Psi(\mu_{\text{in}})$. When we use Panjer recursion or transform methods in the risk theory setting we essentially calculate $\Psi(\mu_h)$ and take this to be an approximation for μ_{out}, where μ_h is a discretized version of μ_{in}. In this section we argue that this general "functional" view can be used to explain a result in Embrechts et al. (1993) that deals with the approximation of integrals $\int \phi d\mu_{\text{out}}$ for a given (smooth) function $\phi : \mathbb{R} \to \mathbb{R}$.

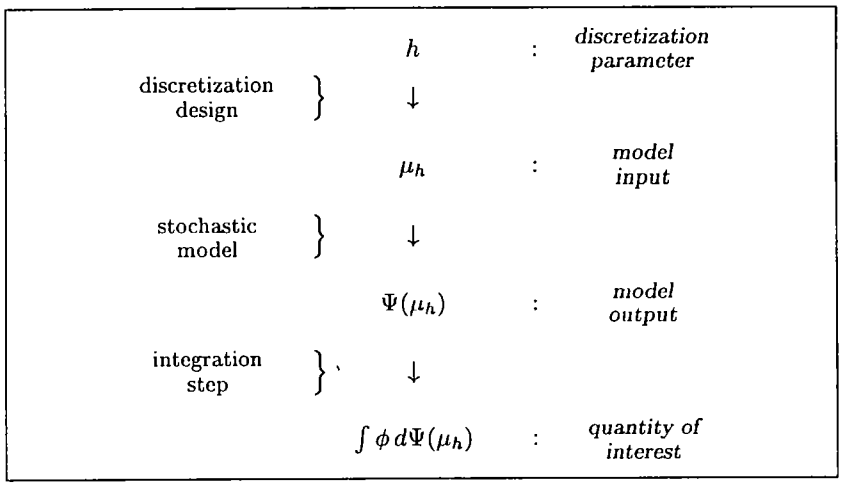

FIGURE 1: The basic decomposition.

The basis for an extrapolation to the limit is a relation of the type (1.1), which in the present context takes the form

$$I(h) := \int \phi d\Psi(\mu_h) = \int \phi d\Psi(\mu) + c(\Psi; \mu, \phi) \cdot h^{\alpha} + o(h^{\beta}) \tag{2.1}$$

as $h \downarrow 0$ with some constants $\beta \geq \alpha > 0$ and $c(\Psi; \mu, , \phi) \neq 0$. To explain the qualitative reasoning that can lead to such an expansion we decompose the function $h \to I(h)$ as in Figure 1. From this diagram it should be obvious that a relation of the type (2.1) can be obtained from the local analysis of the specific construction underlying the transition from μ to μ_h, together with a similar local analysis of the stochastic model. Differentiability (in a suitable sense) of the individual mappings, in conjunction with the chain rule, provides a mathematical frame that can be used to justify the extrapolation to the limit. If Ψ is smooth enough then the constant α will depend on the discretization design only whereas $c(\Psi; \mu, \phi)$ will involve the derivatives and will therefore not be easily accessible. We remark at this point that extrapolation can be carried further by combining more than one $I(h)$-value, and that a similar reasoning applies, now involving higher order derivatives.

As we are dealing with weak results only it is natural to regard the measures $\mu, \mu_h, \Psi(\mu)$ and $\Psi(\mu_h)$ as linear mappings $\phi \to \int \phi d\mu$ on some linear space $\mathbb{F}$ of functions ϕ, i.e. as elements of the dual $\mathbb{F}'$ and $\mathbb{F}$. Let δ_x be the one-point mass in x. We discretize μ by lumping together the masses of intervals of length h. If we centre these masses we arrive at the following discretization design:

$$\mu_h := \sum_{k \in \mathbb{Z}} \mu(I_{hk}) \delta_{hk} \text{ with } I_{hk} := \left(\left(k - \frac{1}{2} \right) h, \left(k + \frac{1}{2} \right) h \right], k \in \mathbb{Z}. \tag{2.2}$$

We now assume that μ has a density f and that f and ϕ are sufficiently smooth for the following approximation to be valid, which is based on a Taylor expansion of the function $x \to (\phi(x) - \phi(kh))f(x)$ at the point kh,

$$\int_{I_{hk}} (\phi(x) - \phi(kh))f(x)dx = \frac{1}{12}\phi'(kh)f'(kh)h^3 + \frac{1}{24}\phi''(kh)f(kh)h^3 + o(h^3).$$

Here it is crucial that the linear term in the expansion vanishes due to the symmetry of the discretization intervals. Summing over k we obtain Riemann sums for two integrals, and assuming suitable behaviour of the functions at $\pm\infty$ an integration by parts yields

$$\begin{aligned} \int \phi d\mu - \int \phi d\mu_h &= \sum_{k \in \mathbb{Z}} \int_{I_{hk}} (\phi(x) - \phi(hk))f(x)dx \\ &= \frac{h^2}{12} \int \phi'(x)f'(x)dx + \frac{h^2}{24} \int \phi''(x)f(x)dx + o(h^2) \\ &= -\frac{h^2}{24} \int \phi''(x)\mu(dx) + o(h^2). \end{aligned} \tag{2.3}$$

This is a statement on the local behaviour of the discretization mechanism, the first step in the above diagram.

For the analysis of the second step we first consider a single convolution power μ^{*n} (which could of course be regarded as a degenerate compound distribution, with $N \equiv n$). This means that our stochastic model is now the non-linear operator

$$\Phi : \mathbb{F}' \to \mathbb{F}', \quad \Phi(\mu)(\phi) := \int \phi d\mu^{*n} \quad \text{for all} \quad \phi \in \mathbb{F}.$$

The derivative of Φ at μ is a linear approximation to Φ that is locally correct at μ. For a "curve" $(\mu_\epsilon)_{\epsilon \geq 0}$ in the range of definition of Φ that is of the form $\mu_\epsilon = \mu + \epsilon\rho$, ρ a finite signed measure of total mass zero, we see that

$$\frac{1}{\epsilon}(\Phi(\mu_\epsilon) - \Phi(\mu))(\phi) \to n \int \phi d(\mu^{*(n-1)} * \rho) \text{ as } \epsilon \to 0,$$

i.e. we have $\Phi'_\mu(\rho) = n\mu^{*(n-1)} * \rho$. Note that this is linear in ρ. Returning to our original model $\Psi(\mu) = \sum_{n=0}^{\infty} p_n \mu^{*n}$ we therefore obtain

$$\Psi'_\mu(\rho) = \tilde{\nu} * \rho \quad \text{with} \quad \tilde{\nu} := \sum_{n=1}^{\infty} n p_n \mu^{*(n-1)}. \tag{2.4}$$

In order to obtain a statement of the form (2.1) from the analysis of the individual steps in the above diagram (the last step is linear and does not require any further investigation) we rephrase (2.3) as a differentiability property. Let $D : \mathbb{F} \to \mathbb{F}$ denote the "ordinary" differentiation of real functions. If we define $T : [0, \infty) \to \mathbb{F}'$ by $T(h) := \mu_{\sqrt{h}}$ for $h > 0$, $T(0) := \mu$, we see that (2.3) can be reinterpreted as

$$T'_{0+}(1) = \frac{1}{24}\mu \circ D^2, \tag{2.5}$$

where the linear map $T'_{0+} : \mathbb{R}_+ \to \mathbb{F}'$ is the right derivative of T at 0, defined by

$$T'_{0+}(\alpha)(\phi) = \lim_{h \downarrow 0} \frac{1}{h}(T(h\alpha) - T(0))(\phi).$$

Because of $T'_{0+}(\alpha) = \alpha T'_{0+}(1)$ this derivative is completely specified by (2.5). The chain rule that we now need can be written in the form

$$(\Psi \circ T)'_{0+}(1)(\phi) = \Psi'_{T(0)} \circ T'_{0+}(1)(\phi),$$

so that, replacing h by h^2, we finally arrive at

$$\frac{1}{h^2}\left(\int \phi d\Psi(\mu_h) - \int \phi d\Psi(\mu)\right) \to \frac{1}{24}\int \phi'' d(\tilde{\nu} * \mu)$$

as $h \downarrow 0$, which is (2.1) with $\alpha = \beta = 2$ and

$$c(\Psi; \mu, \phi) = \frac{1}{24} \int \phi'' d\left(\sum_{n=1}^{\infty} n p_n \mu^{*n} \right). \tag{2.6}$$

This agrees with the result obtained by Embrechts et al. (1993) and provides the required theoretical justification for an extrapolation algorithm.

3. A STRONG RESULT

We continue to use the notation from the previous section; in particular, the discretized version μ_h of μ is given by (2.2).

If we are interested in $\nu = \Psi(\mu)$ itself rather than in some specific integral $\int \phi d\nu$ we need to compare the two measures ν and $\nu_h = \Psi(\mu_h)$. We want to go beyond convergence of ν_h to ν as $h \downarrow 0$; in fact, applicability of an extrapolation technique requires at least a "next term" result, and it is a priori not clear what an expansion of the form (1.1) means if we deal with measures instead of real numbers. If ν has a smooth density f_ν we might hope that

$$f_\nu(kh) = \frac{1}{h} \nu_h(\{kh\}) + g(kh)h^\alpha + O(h^\beta) \tag{3.1}$$

holds uniformly in $k \in \mathbb{Z}$ for $h \downarrow 0$ with some function g depending on Ψ and μ. Again, we could then combine the approximations of type (3.1) obtained for different h-values into a new approximation with increased asymptotic accuracy if $\beta > \alpha$.

We will now carry out this programme rigorously under some smoothness and growth conditions on the density of μ. These conditions preclude boundary effects (see the next section) and our main result shows that Richardson extrapolation will accelerate the basic h^2-convergence to the rate h^4 as $h \downarrow 0$. In particular, under these assumptions we have $\alpha = 2$ and $\beta = 4$ in (3.1).

Let $C(4, \gamma)$ be the space of all continuous functions $f: \mathbb{R} \to \mathbb{R}$ with continuous derivatives $f^{(i)}$ up to order $i = 4$ which satisfy

$$f^{(i)}(x) = O(|x|^{-\gamma}) \quad \text{as} \quad x \to \pm\infty \quad \text{for} \quad i = 0, \ldots, 4. \tag{3.2}$$

Here we take $f^{(0)}$ to be f itself. The use of these spaces is motivated by our method of proof, which depends on Taylor series expansions to fourth order. Let $\nu := \sum_{n=1}^{\infty} p_n \mu^{*n}$ and $\nu_h := \sum_{n=1}^{\infty} p_n \mu_h^{*n}$ for all $h > 0$. Hence, in contrast to the previous section, our convolution series now begin with $n = 1$ rather than $n = 0$. As $\mu^{*0} = \mu_h^{*0} = \delta_0$ this does not affect $\nu - \nu_h$. Also, as μ and the sequence $(p_n)_{n \in \mathbb{N}_0}$ are regarded as known, we know explicitly the atom of the measure $\nu = \sum_{n=0}^{\infty} p_n \mu^{*n}$ at 0: if μ has a density then $\mu(\{0\})^n = 0$ for all $n \in \mathbb{N}$, so that $\nu(\{0\}) = p_0$.

Theorem

Assume the μ has a density $f_\mu \in C(4,\gamma)$ for some $\gamma > 1$ and that $(p_n)_{n\in\mathbb{N}}$ satisfies

$$\sup_{n\in\mathbb{N}}(1+\epsilon)^n p_n < \infty \text{ for some } \epsilon > 0. \tag{3.3}$$

Let f_ν be the continuous density of ν and let $f_{\nu,h} : h\mathbb{N}_0 \to \mathbb{R}$ be defined by $f_{\nu,h}(kh) := \nu_h(\{kh\})/h$. Then there exists a continuous function $g : \mathbb{R} \to \mathbb{R}$ such that

$$\limsup_{h\downarrow 0} \frac{1}{h^2} \sup_{k\in\mathbb{Z}} (1+|k|h)^\gamma \left| \frac{1}{h^2}\left(f_{\nu,h}(kh) - f_\nu(kh)\right) - g(kh)\right| < \infty. \tag{3.4}$$

Note that (3.3) holds for compound Poisson and compound geometric distributions. The proof of the theorem is given in the appendix which also contains an explicit formula for the limit function g. Informally, we could guess at the behaviour of $f_{\nu,h}$ be rewriting the value $c(\Psi;\mu,\phi)$ in (2.6) via partial integration such that only ϕ itself appears in the integrals rather than ϕ'', and then inserting Dirac's delta function for ϕ. This would lead to

$$g = \frac{1}{24}\left(\sum_{n=1}^{\infty} n p_n f_\mu^{*n}\right)'' = \frac{1}{24}\left(p_1 f_\mu'' + \sum_{n=2}^{\infty} n p_n f_\mu^{*(n-1)} * f_\mu''\right),$$

and this is indeed correct.

For any given $x > 0$ (3.4) implies

$$\frac{n}{x} \nu_{x/n}(\{x\}) = f_\nu(x) + \frac{x^2}{n^2} g(x) + O(n^{-4})$$

as $n \to \infty$, $n \in \mathbb{N}$. This is a result of the form (3.1) that can be used for an extrapolation to the limit. Indeed, as announced above, the rate itself does not depend on the model input provided that the general qualitative assumptions in the theorem are satisfied, but the next order term, i.e. the function g, does depend on f_μ. If extrapolation methods required the explicit form of the next term they would be useless in the present context as g is no easier to obtain than f_ν.

The theorem shows that the approximation can be done uniformly in x. In practice one would typically be interested in the values of the density of the total claim size distribution on a lattice $h\mathbb{Z}$. One could then carry out the Panjer recursion, or the transform based algorithm if μ is not concentrated on a half line, for h and $h/2$ and combine the results to obtain improved values on $h\mathbb{Z}$. Numerical examples are given in Section 5.

4. Boundary effects

In an insurance mathematics context the main (but not only) application of compound distributions appears in connection with the total claim size over a given period of time. This total claim size is the sum of a random number of individual claims of random size, and it is natural to assume that these individual claims are nonnegative. To many "one-sided" distributions of interest, such as the exponential distribution, the theorem in the previous section cannot be applied directly, as the corresponding density does not satisfy the smoothness

conditions at the left endpoint 0 of its support. In such cases we can often argue as in connection with the atoms at 0 (see the discussion following (3.2)): as we regard the compounding sequence $(p_n)_{n\in\mathbb{N}_0}$ and the individual claim size distribution μ as known, we can easily obtain the value of the compound density at 0 explicitly. If e.g. μ is the exponential distribution with parameter λ, then the convolution powers μ^{*n} with $n > 1$ do not contribute to the value $f_\nu(0)$ of the compound density at 0, which yields $f_\nu(0) = p_1 f_\mu(0)$. Here and in the following we always take the version of the density that is continuous from the right at 0.

Regardless of such a practical attitude it is mathematically interesting to understand the behaviour of extrapolation algorithms in "non-smooth" cases. The discussion in this section concentrates on the essential points but does not carry out the details; some of these are given in Hermesmeier (1997). We assume that μ is concentrated on the nonnegative real numbers and that it has a density f_μ that apart from being continuous from the right at 0 also satisfies the smoothness conditions implicitly present in the approximations made below. We ignore the atom of the compound distribution ν at 0, $f_\nu = \sum_{n=1}^{\infty} p_n f_\mu^{*n}$ is the compound density. The discretization is the same as in the previous subsections, but the step from ν_h to $f_{\nu,h}$ has to take into account that f_ν vanishes for negative arguments: we put $f_{\nu,h}(0) := \nu_h(\{0\})/(h/2)$.

We first consider weak results as in Section 2. We continue to use the discretization scheme given in (2.2). Now the linear term in the Taylor expansion for the integrals aver the I_{hk}-intervals does not disappear at the boundary $k = 0$, and we obtain

$$\int_{I_{h0}} (\phi(x) - \phi(0)) f(x)dx = \frac{1}{8}\phi'(0)f(0)h^2 + o(h^2),$$

which leads to

$$\begin{aligned}\int \phi d\mu - \int \phi d\mu_h &= \frac{h^2}{8}\phi'(0)f(0) + \frac{h^2}{12}\int_0^\infty \phi'(x)f'(x)dx \\ &\quad + \frac{h^2}{24}\int_0^\infty \phi''(x)f(x)dx + o(h^2) \\ &= -\frac{h^2}{24}\left(\int_0^\infty \phi''(x)\mu(dx) - \phi'(0)f(0)\right) + o(h^2).\end{aligned}$$

The second and third step of the decomposition in Figure 1 remain unchanged, (2.5) now becomes

$$T'_{0+}(1) = \frac{1}{24}\left(\mu \circ D^2 - f(0)\delta_0 \circ D\right),$$

so that finally (2.1) is seen to hold with $\alpha = \beta = 2$ and

$$c(\Psi; \mu, \phi) = \frac{1}{24}\left(\int \phi'' d\left(\sum_{n=1}^{\infty} np_n\mu^{*n}\right) - f(0)\int \phi' d\left(\sum_{n=1}^{\infty} np_n\mu^{*(n-1)}\right)\right).$$

Hence, for the weak result, we are essentially in the same situation as in Section 2. It is only the constant that changes, and this constant is not needed for the algorithm.

In connection with strong results it is of course the behaviour of the approximation at the boundary that is of special interest. We assume that the density f_μ of μ can be written in the form

$$f_\mu(x) = f_\mu(0) + f'_\mu(0) + xr(x)$$

with some function r satisfying $\lim_{x \downarrow 0} r(x) = 0$. Then

$$\mu_h(\{0\}) = \frac{h}{2} f_\mu(0) + \frac{h^2}{8} f'_\mu(0) + h^2 \tilde{r}(h)$$

with some function $\tilde{r}$ satisfying $\lim_{h \downarrow 0} \tilde{r}(h) = 0$. All convolution powers of μ_h are concentrated on the right half-line, hence

$$\begin{aligned} v_h(\{0\}) &= \sum_{n=1}^{\infty} n p_n \mu_h^{*n}(\{0\}) \\ &= \sum_{n=1}^{\infty} p_n (\mu_h(\{0\}))^n \\ &= p_1 \mu_h(\{0\}) + p_2 \mu_h(\{0\})^2 + o(h^2) \\ &= \frac{1}{2} p_1 f_\mu(0) h + \left(\frac{1}{8} p_1 f'_\mu(0) + \frac{1}{4} p_2 f_\mu(0)^2 \right) h^2 + o(h^2) \end{aligned}$$

(a condition such as (3.3) is needed in order to show that the terms from $n = 3$ onwards may be neglected in the above infinite sum). As $f_v(0) = p_1 f_\mu(0)$ and $f_{v,h}(0) = v_h(\{0\})/(h/2)$ this yields

$$\frac{1}{h} \left(f_{v,h}(0) - f_v(0) \right) = \frac{1}{4} p_1 f'_\mu(0) + \frac{1}{2} p_2 f_\mu(0)^2 + o(1).$$

This is again a result of type (2.1), but now we have $\alpha = 1$, whereas we had $\alpha = 2$ in all the cases considered previously. As a consequence the correct extrapolation formula would now combine $f_{v,h}(0)$ and $f_{v,h/2}(0)$ into $2f_{v,h/2}(0) - f_{v,h}(0)$ rather than $(4f_{v,h/2}(0) - f_{v,h}(0))/3$ as the case $\alpha = 2$.

Why are the conclusions different in the weak and the strong situation? Due to the final integration step the former case is basically of a global nature, which means that a single point with a differing rate does not matter.

5. Numerical examples

In this section we consider two examples, one where the conditions of the theorem in Section 3 are satisfied, the other with a claim size density that has a jump at 0. For both these examples some gamma distribution $\Gamma(\eta, \lambda)$ with shape

parameter $\eta > 0$ and scale parameter $\lambda > 0$ will be used; the corresponding density is

$$f(x|\eta, \lambda) = \frac{\lambda^\eta}{\Gamma(\eta)} x^{\eta-1} e^{-\lambda x}, x > 0.$$

Exponential distributions arise if the shape parameter has the value 1. Gamma distributions are convenient in the present context as their convolution powers can be given explicitly: if $\mu = \Gamma(\eta, \lambda)$ then $\mu^{*n} = \Gamma(n\eta, \lambda)$ for all $n \in \mathbb{N}$. We consider the compound Poisson case where we take the mean of the Poisson distribution to be 10. We then have an explicit series representation for the compound density,

$$f_v(x) = e^{-10-\lambda x} \sum_{n=1}^{\infty} \frac{10^n \lambda^{n\eta} x^{n\eta-1}}{n!\Gamma(n\eta)} \quad \text{for all } x > 0.$$

From this formula it is easy to obtain $f_v(x)$ to any desired degree of accuracy by simply summing the first few n-terms (50 are by far enough for our purposes – if N has a Poisson distribution with mean 10 then $P(N > 50) \approx 0.362 \cdot 10^{-19}$, and this is an upper bound for the L^1-distance between the true compound density and its approximation obtained by summing to $n = 50$ only). We can therefore compare the quality of the approximations obtained for different values of the discretization parameter and assess the merits of Richardson extrapolation.

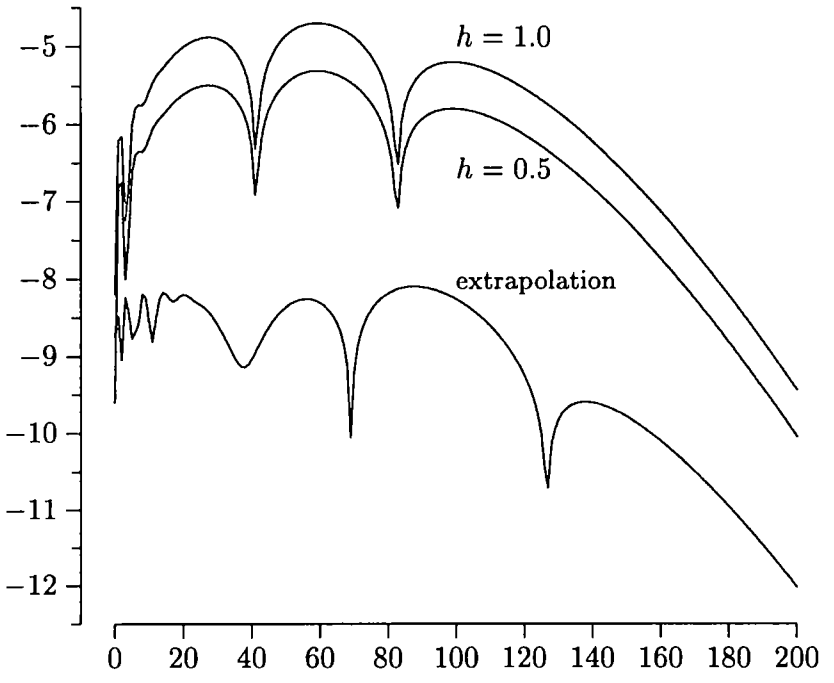

FIGURE 2: Logarithmic absolute errors for $\mu = \Gamma(6, 1)$.

Our first example has $\eta = 6$ and $\lambda = 1$, with these values the conditions of the theorem in Section 3 are satisfied. Figure 2 contains three curves which all display some error

$$\log_{10}|f_{v,\text{approx}}(x) - f_v(x)| \quad \text{for } x = 0, 1, 2, \ldots, 199, 200$$

on a logarithmic scale for this example. From top to bottom the first curve is the result with discretization parameter $h = 1.0$, the second with $h = 0.5$ and the third gives the errors for the extrapolation approximation that arises if these two approximations are combined as explained in the previous sections. Apart from some initial confusion at the left boundary, an aspect that will be discussed below, the substantial gain in accuracy achieved by the extrapolation is obvious.

In the second example we take the claim size distribution to be exponential with mean 6, Figure 3 displays the corresponding error curves. Interestingly, we again have a noticeable gain through extrapolation, even though the conditions of our main theorem are not satisfied in this case.

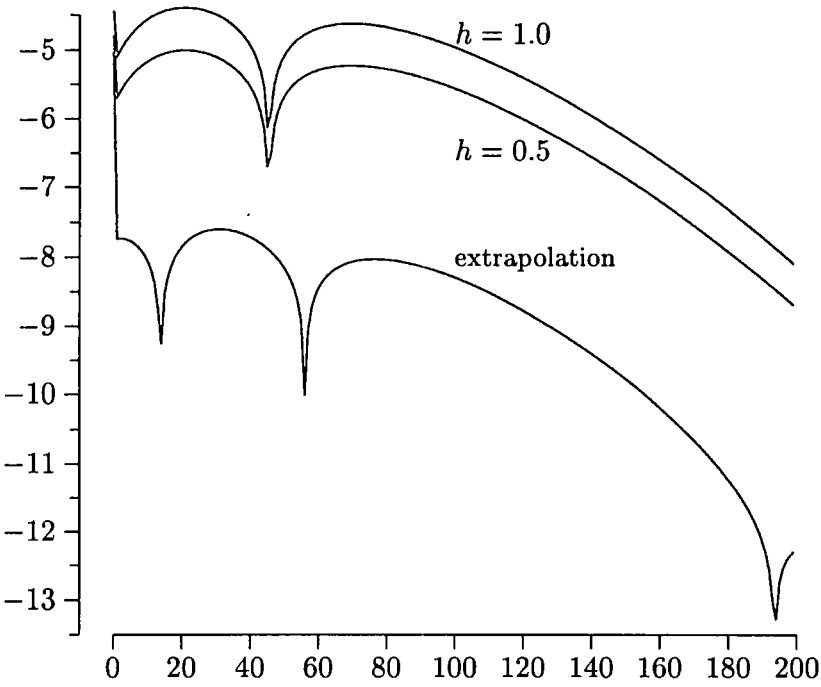

FIGURE 3: Logarithmic absolute errors for $\mu = \Gamma(1, 1/6)$.

For a better understanding of the behaviour near the boundary $x = 0$ we display magnified parts of Figures 2 and 3 in Figure 4. On first sight the behaviour observed in these two examples seems at variance with the mathematical results: the first example satisfies the conditions of the theorem

and the second does not, but we have an improvement by extrapolation at $x = 0$ in the latter case whereas the extrapolation increases the error at $x = 0$ in the first example ("$*$" is above "$\circ$" at $x = 0$ in part (a) of the figure). But behold, it is the mathematical result that helps to explain the phenomenon: the limit function g that appears in the theorem (see the discussion following the theorem) has the value 0 at $s = 0$ for $\mu = \Gamma(6, 1)$ which means that the constant c in (1.1) vanishes. In fact, this amounts to the same as using the wrong α-exponent in the extrapolation, and this is responsible for the disappointing performance of the extrapolation in this special example at this particular x-value.

From the discussion in Section 4 we know that for $\mu = \Gamma(1, 1/6)$ the correct extrapolation would use $\alpha = 1$ at $x = 0$. As the extrapolation shown in the figures is based on the assumption that $\alpha = 2$ it is surprising that we obtain an improvement nevertheless ("$*$" is below "$\circ$" at $x = 0$ in part (b) of the figure). Again, it is worth looking at the details: the values obtained for $f_v(0)$ with $h = 1.0$, $h = 0.5$ and the $\alpha = 2$ based extrapolation are 0.1111892, 0.0915200 and 0.0849636 respectively, whereas the true value is 0.0756665. An extrapolation with $\alpha = 1$ however leads to 0.0718508, which is closer to the true value (all these numbers have been multiplied by 1000). The error value obtained with the linear extrapolation is given as "$\oplus$" in Figure 4 (b). One should note, however, that the actual order of magnitude of the error at the boundary is much smaller in the first than in the second example.

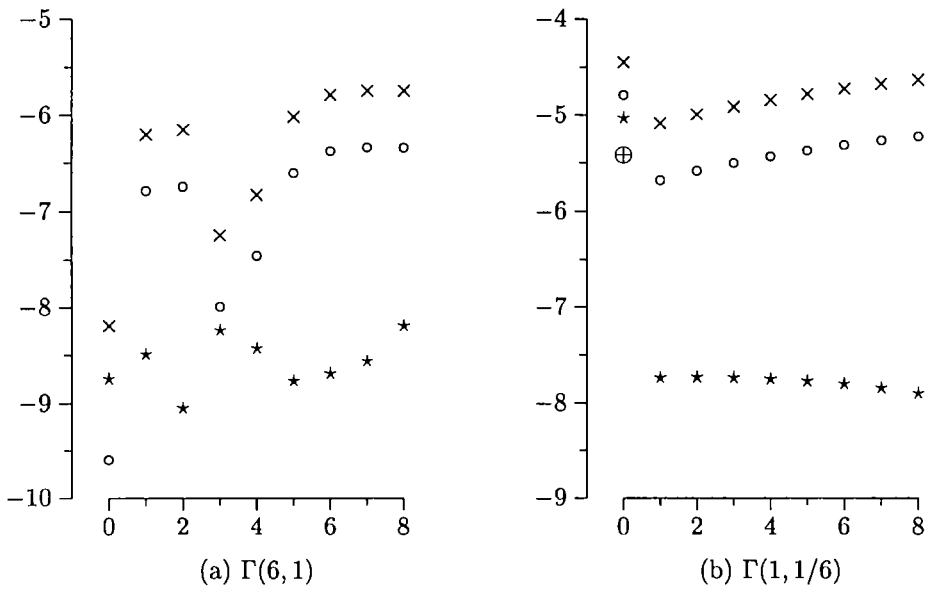

FIGURE 4: Behaviour near the boundary
$\times$: $h = 1.0$, $\circ$: $h = 0.5$, $*$: extrapolation, $\oplus$: see text.

Our general mathematical findings can be summarized as follows: smoothness of a stochastic model, regarded as a mapping in a suitable infinite-dimensional setting, implies that the order of magnitude of the discretization error on the input side will reappear at the output side, and this in turn means that an extrapolation to the limit will work, since we control the input error by choosing the discretization. This can be used as guidance in the choice of a discretization design. For example, the dicretization

$$\mu'_h := \sum_{k=0}^{\infty} \mu(I'_{hk})\delta_{hk} \quad \text{with} \quad I'_{h0} := [0, h], I'_{hk} := (k \cdot h, (k+1) \cdot h] \quad \text{for} \quad k \in \mathbb{N}$$

for one-sided distributions, which is mentioned in Hipp and Michel (1990), leads to $\alpha = 1$ rather than $\alpha = 2$ in (3.1) and is therefore inferior to the one we used in the previous sections. Also, one could take this further and argue as follows: if interest is in the compound density and if the density f_μ of the individual claim size distribution is easily accessible then a direct discretization of f_μ might well lead to a better approximation – after all, there would then be *no* discretization error on the input side. Indeed, the methods and auxiliary results introduced in the appendix for the proof of our theorem can be used to show that, under some smoothness conditions, an approximation obtained for the compound density f_v by applying Panjer recursion or transform based methods directly to the sequence of values $f_\mu(kh), k = 0, 1, 2, \ldots$ rather than to $\mu(I_{h,k})$ will give an error of magnitude $O(h^4)$. In Figure 5 we display the errors obtained in the first of the above numerical examples with the direct method and the same discretization parameters as used previously, together with the extrapolation error curve (which is the same as the bottommost curve in Figure 2). It is interesting to see that the latter is below the error obtained with the direct method on a substantial range of x-values, but not on the whole axis. An advantage of the discretization via $\mu(I_{h,k})$ is that we now know the behaviour of the error, which makes this method "extrapolatable"; a linear or quadratic extrapolation leads to a deterioration if we naively apply it to the results obtained with the direct discretization of f_μ.

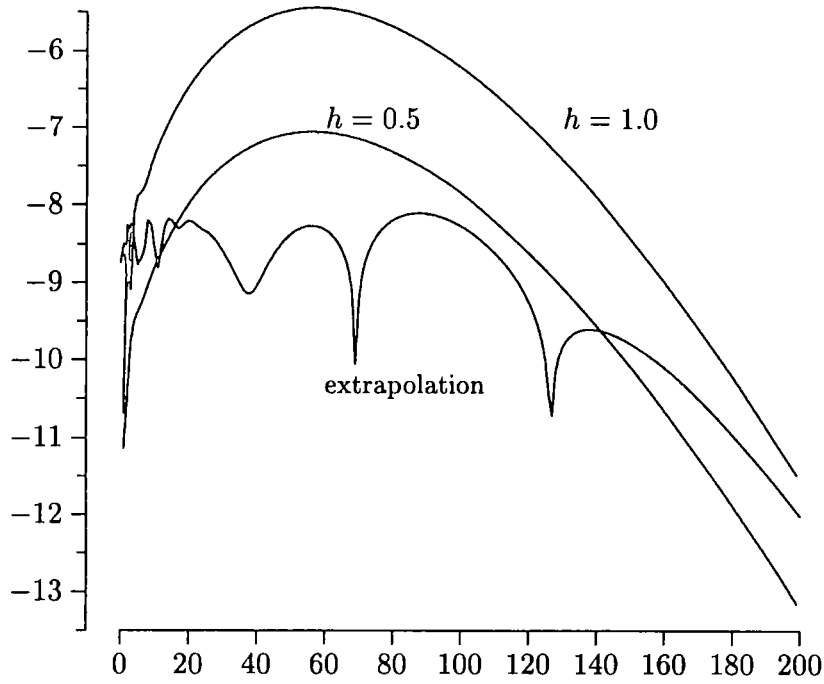

FIGURE 5: Logarithmic absolute errors for $\mu = \Gamma(6, 1)$, direct discretization.

APPENDIX

Proof of the theorem

The proof uses elementary arguments and techniques only, but is somewhat lengthy. To increase readability we divide the appendix into subsections, dealing with functions on the real line in the first, with functions on lattices in the second and relations between these via discretization operators in the third part. The final subsection then combines the material from the first three subsections into the proof of the theorem, which we first reformulate in the new notational framework. Throughout, γ is a fixed real number greater than 1. The discretization parameter h is generally assumed to vary over the interval $(0, 1]$, but the choice of 1 as the upper bound is not important. A sequence $(a_n)_{n\in\mathbb{N}}$ of real numbers is said to grow subexponentially fast if $\lim_{n\to\infty}(1+\epsilon)^{-n}a_n = 0$ for all $\epsilon > 0$, it decreases at an exponential rate if $\lim_{n\to\infty}(1+\epsilon)^{n}a_n = 0$ for some $\epsilon > 0$.

A.1 For measurable functions $f: \mathbb{R} \to \mathbb{R}$ the supremum norm with weight function $x \to (1 + |x|)^\gamma$ is given by

$$\|f\|_{0,\gamma} := \sup_{x \in \mathbb{R}} (1 + |x|)^\gamma |f(x)|.$$

Let $B_0(\gamma)$ be the linear space of such functions f with $\|f\|_{0,\gamma} < \infty$. Since $\gamma > 1$ all elements of $B_0(\gamma)$ are integrable, i.e.

$$\|f\|_{0,1} := \int |f(x)| dx < \infty \quad \text{for all } f \in B_0(\gamma).$$

(Note that $\|\cdot\|_{0,1}$ is an integral norm and thus differs in nature from $\|\cdot\|_{0,\gamma}$. We always have $\gamma > 1$, so no confusion should arise from the positional similarity of the indices, which is meant to keep the notation compact.) This implies that the convolution $f * g$ of any $f, g \in B_0(\gamma)$, defined by

$$f * g(x) := \int f(y) g(x - y) dy \quad \text{for all } x \in \mathbb{R},$$

exists. Since $\gamma > 1$ we also have

$$\kappa_0(\gamma) := \sup_{x \in \mathbb{R}} (1 + |x|)^\gamma \int (1 + |y|)^{-\gamma} (1 + |x - y|)^{-\gamma} dy < \infty,$$

which can be used to obtain inequalities relating convolution and $\|\cdot\|_{0,\gamma}$-norm. For the integral norm such an inequality is a well known consequence of Fubini's theorem, i.e.

$$\|f * g\|_{0,1} \leq \|f\|_{0,1} \|g\|_{0,1}.$$

In the corresponding statement for supremum norms an additional constant appears on the right hand side and good bounds for this constant are of some importance for us. The first part of the following lemma implies that $B_0(\gamma)$ is closed with respect to convolution, the second part shows that the $\|\cdot\|_{0,\gamma}$-norms of the convolution powers of a probability density increase subexponentially fast.

Lemma A.1 (i) $\|f * g\|_{0,\gamma} \leq \kappa_0(\gamma) \|f\|_{0,\gamma} \|g\|_{0,\gamma}$.
(ii) *For all* $n \in \mathbb{N}$, $\|f^{*n}\|_{0,\gamma} \leq n^\gamma \|f\|_{0,\gamma} \|f\|_{0,1}^{n-1}$.

Proof: Part (i) is immediate from the definitions. For (ii) we first use convexity of $x \to (1 + |x|)^\gamma$ to obtain for all $x_1, x_2, \cdots, x_n$,

$$\begin{aligned} n^{-\gamma}(1 + |x_1 + \ldots + x_n|)^\gamma &\leq \left(1 + \left|\frac{1}{n}(x_1 + x_2 + \ldots + x_n)\right|\right)^\gamma \\ &\leq \frac{1}{n}((1 + |x_1|)^\gamma + (1 + |x_2|)^\gamma + \ldots + (1 + |x_n|)^\gamma). \end{aligned}$$

This yields with $x_1 := x - x_2 - \ldots - x_n$

$$(1+|x|)^\gamma |f^{*n}(x)| \le \int \ldots \int (1+|x|)^\gamma |f(x_1)||f(x_2)| \ldots |f(x_n)| dx_2 \ldots dx_n$$

$$\le n^{\gamma-1} \Bigg(\int \ldots \int (1+|x_1|)^\gamma |f(x_1)||f(x_2)| \ldots |f(x_n)| dx_2 \ldots dx_n$$

$$+ \int \ldots \int (1+|x_2|)^\gamma |f(x_1)||f(x_2)| \ldots |f(x_n)| dx_2 \ldots dx_n$$

$$+ \ldots + \int \ldots \int (1+|x_n|)^\gamma |f(x_1)||f(x_2)| \ldots |f(x_n)| dx_2 \ldots dx_n \Bigg).$$

Using $(1+|x|)^\gamma |f(x_1)| \le \|f\|_{0,\gamma}$ we obtain

$$\int \ldots \int (1+|x_1|)^\gamma |f(x_1)||f(x_2)| \ldots |f(x_n)| dx_2 \ldots dx_n$$

$$\le \|f\|_{0,\gamma} \int |f(x_2)| dx_2 \ldots |f(x_n)| dx_n$$

$$\le \|f\|_{0,\gamma} \|f\|_{0,1}^{n-1}.$$

For $k = 2, \ldots, n$ the same inequality with x_k instead of x_1 yields

$$\int \ldots \int (1+|x_k|)^\gamma |f(x_1)||f(x_2)| \ldots |f(x_n)| dx_2 \ldots dx_n$$

$$\le \|f\|_{0,\gamma} \int \ldots \int |f(x - x_2 - \ldots - x_n)| \prod_{\substack{j=2 \\ j \ne k}}^{n} |f(x_j)| dx_2 \ldots dx_n.$$

As x_k appears only in the first factor of the integrand we can use that

$$\int |f(x - x_2 - \ldots - x_n)| dx_k = \|f\|_{0,1} \quad \text{for all } x_2, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n$$

to rewrite the multiple integral as a product of single integrals, which yields the upper bound $\|f\|_{0,\gamma} \|f\|_{0,1}^{n-1}$ again. Put together these estimates result in

$$(1+|x|)^\gamma |f^{*n}(x)| \le n^{\gamma-1} n \|f\|_{0,\gamma} \|f\|_{0,1}^{n-1}.$$

This final bound does not depend on x, hence (ii) follows. □

A.2 For functions $a_h : h\mathbb{Z} \to \mathbb{R}$ we proceed similarly to the previous subsection: we introduce the norm

$$\|a_h\|_{h,\gamma} := \sup_{k \in \mathbb{Z}} (1+|k|h)^\gamma |a_h(kh)|$$

and let $B_h(\gamma)$ be the space of such functions a_h with $\|a_h\|_{h,\gamma} < \infty$. The elements of this space, regarded as sequences, are absolutely summable, i.e.

$$\|a_h\|_{h,1} := h \sum_{k\in\mathbb{Z}} |a_h(kh)| < \infty \quad \text{for all } a_h \in B_h(\gamma).$$

In particular, we can define the convolution product $a_h * b_h$ of $a_h, b_h \in B_h(\gamma)$ by

$$a_h * b_h(kh) := h \sum_{j\in\mathbb{Z}} a_h(jh) b_h((k-j)h) \quad \text{for all } k \in \mathbb{Z}.$$

We use the same symbol to denote convolution on $B_h(\gamma)$ and on $B_0(\gamma)$ since the respective space should be clear from the context; the symbols $f, g, \cdots$ generally denote functions on $\mathbb{R}$, $a_h, b_h, \cdots$ functions on $h\mathbb{Z}$.

Since $\gamma > 1$ we have

$$\kappa_1(\gamma) := \sup_{0<h\le 1} \sup_{k\in\mathbb{Z}} h(1+|k|h)^{\gamma} \sum_{j\in\mathbb{Z}} (1+|j|h)^{-\gamma}(1+|k-j|h)^{-\gamma} < \infty,$$

which can again be used to obtain norm inequalities. We omit the proof of the following lemma as it is very similar to the proof of Lemma A.1.

Lemma A.2 (i) *For* $0 < h \le 1$, $\|a_h * b_h\|_{h,\gamma} \le \kappa_1(\gamma)\|a_h\|_{h,\gamma}\|b_h\|_{h,\gamma}$.
(ii) *For all* $n \in \mathbb{N}$, $\|a_h^{*n}\|_{h,\gamma} \le n^{\gamma}\|a_h\|_{h,\gamma}\|a_h\|_{h,1}^{n-1}$.

A.3 We now relate the function spaces introduced in the previous two subsections to each other. Two canonical families $(\Pi_h)_{0<h\le 1}$ and $(T_h)_{0<h\le 1}$ of *discretization operators* are given by

$$\Pi_h : B_0(\gamma) \to B_h(\gamma), \quad \Pi_h(f)(kh) := f(kh),$$

and

$$T_h : B_0(\gamma) \to B_h(\gamma), \quad T_h(f)(kh) := \frac{1}{h}\int_{I_{hk}} f(x)dx,$$

with

$$I_{hk} := \left(\left(k-\frac{1}{2}\right)h, \left(k+\frac{1}{2}\right)h\right] \quad \text{for all } k \in \mathbb{Z}.$$

We note in passing that $\|\Pi_h(f)\|_{h,\gamma} \le \|f\|_{0,\gamma}$, which implies that Π_h is continuous. An argument from the proof of Lemma A.3 (i) below can be used to show that $\|T_h(f)\|_{h,\gamma} \le 2^{\gamma}\|f\|_{0,\gamma}$, i.e. T_h is also a continuous linear operator.

In connection with these discretizations often a whole family $(r_h)_{0<h\le 1}$ of functions $r_h : h\mathbb{Z} \to \mathbb{R}$ arises and we are then interested in a local boundedness condition of the form

$$\text{(B)} \qquad \limsup_{h\to 0} \|r_h\|_{h,\gamma} < \infty,$$

i.e. $\|r_h\|_{h,\gamma} = O(1)$ as $h \to 0$. If we have such a family $\left(r_h^{(n)}\right)_{0<h\le 1}$ for each $n \in \mathbb{N}$, a uniform subexponential bound of the form

$$\text{(S)} \qquad \limsup_{h\to 0} \sup_{n\in\mathbb{N}} (1+\epsilon)^{-n} \left\| r_h^{(n)} \right\|_{h,\gamma} < \infty \quad \text{for all } \epsilon > 0$$

will be needed. For (B) and (S) some rules are easily established. Obviously, both conditions continue to bold for finite linear combinations of families. Moreover, if (S) holds for $\left(r_h^{(n)}\right)_{0<h\le 1}, n \in \mathbb{N}$, and if $(q_n)_{n\in\mathbb{N}}$ decreases at an exponential rate, then $(r_h)_{0<h\le 1}$ defined by $r_h := \sum_{n=1}^{\infty} q_n r_h^{(n)}$ satisfies (B). Also, if $\left(r_{1,h}^{(n)}\right)_{0<h\le 1}$, $\left(r_{2,h}^{(n)}\right)_{0<h\le 1}, n \in \mathbb{N}$, are two families satisfying (S), and if we define $\left(r_h^{(n)}\right)_{0<h\le 1}$, $n \in \mathbb{N}$, by

$$r_h^{(n)} := \sum_{k=0}^{n} r_{1,h}^{(k)} * r_{2,h}^{(n-k)},$$

then (S) also holds for $\left(r_h^{(n)}\right)_{0<h\le 1}, n \in \mathbb{N}$. These rules will be used below without further comment, with the exception of the last one which we will refer to as the convolution stability associated with condition (S). Generally, $a_h^{*0} * b_h := b_h$; below we similarly interpret $f^{*0} * g$ as g for $f, g, \in B_0(\gamma)$.

As an example consider $r_h^{(n)} := \Pi_h(f)^{*n}$ for a continuous probability density $f \in B_0(\gamma)$. Using Lebesgue's theorem on dominated convergence it is easy to see that

$$\lim_{h\to 0} \|\Pi_h(f)\|_{h,1} = \|f\|_{0,1} = 1,$$

so for any $\epsilon > 0$ we can find an $h_0 > 0$ such that $\|\Pi_h(f)\|_{h,1} \le 1 + \epsilon/2$ for all $h \le h_0$. Then, by Lemma A.2 (ii),

$$\left\| r_h^{(n)} \right\|_{h,\gamma} \le n^{\gamma} \|\Pi_h(f)\|_{h,\gamma} (1 + \varepsilon/2)^{n-1} \quad \text{for all } h \le h_0, n \in \mathbb{N}.$$

hence (S) holds for the family $\left(r_h^{(n)}\right)_{0<h\le 1}, n \in \mathbb{N}$.

We next investigate the interplay between discretization and convolution. Let $C(k,\gamma)$ be the space of k times continuously differentiable functions $f: \mathbb{R} \to \mathbb{R}$ with $f^{(i)} \in B_0(\gamma)$ for $i = 0, \cdots, k$. This is compatible with the definition given in Section 3. We note that differentiation can be carried into one of the factors of a convolution product, i.e. $(f * g)' = f * g'$ for $f \in B_0(\gamma), g \in C(1,\gamma)$.

Lemma A.3 For $f, g \in C(k,\gamma)$ let $c_{k,\gamma}(f,g)$ be defined by

$$c_{k,\gamma}(f,g) := \sum_{j=0}^{k} \binom{k}{j} \|f^{(j)}\|_{0,\gamma} \|g^{(k-i)}\|_{0,\gamma}.$$

(i) *If* $f, g \in C(2,\gamma)$ *then, for* $0 < h \le 1$,

$$\|\Pi_h(f) * \Pi_h(g) - \Pi_h(f * g)\|_{h,\gamma} \le h^2 2^{2\gamma} \kappa_1(\gamma) c_{2,\gamma}(f,g).$$

(ii) *If* $f, g \in C(4,\gamma)$ *then, for* $0 < h \le 1$,

$$\|\Pi_h(f) * \Pi_h(g) - \Pi_h(f * g)\|_{h,\gamma} \le h^4 2^{2\gamma} \kappa_1(\gamma)\big(c_{2,\gamma}(f'',g) + c_{2,\gamma}(f',g') + c_{2,\gamma}(f,g'') + c_{4,\gamma}(f,g)\big).$$

Proof: (i) We expand $\phi_k(x) := f(x)g(kh - x)$ for $x \in I_{hj}$ about $x = jh$ to obtain

$$\phi_k(x) = f(jh)g((k-j)h) + (x - jh)\phi'(jh) + \frac{1}{2}(x - jh)^2 \phi''\big(\xi_{kj}(x)\big)$$

with some $\xi_{kj}(x)$ *in* $x \in I_{hj}$. Integrating over $x \in I_{hj}$ we see that the linear term vanishes and

$$\left| hf(jh)g((k-j)h) - \int_{I_{hj}} f(x)g(kh - x)dx \right| \le \sup_{x \in I_{hj}} \left|\phi_k''\big(\xi_{kj}(x)\big)\right| \int_{I_{hj}} \frac{1}{2}(x - jh)^2 dx.$$

The integral evaluates to $h^3/24$. Clearly,

$$\phi_k''(\xi_{kj}) = f''(\xi)g(kh - \xi) - 2f'(\xi)g'(kh - \xi) + f(\xi)g''(kh - \xi).$$

For $x, y \in I_{hj}$ we have $(1 + |x|)/(1 + |y|) \le 2$, so that

$$\left|f''\big(\xi_{kj}(x)\big)\right| \le \big(1 + \xi_{kj}(x)\big)^{-\gamma} \|f''\|_{0,\gamma} \le 2^{\gamma}(1 + |j|h)^{-\gamma} \|f''\|_{0,\gamma}.$$

Similar inequalities hold for the other terms in ϕ_k'', so we obtain

$$\sup_{x \in I_{hj}} \left|\phi_k''(\xi_{kj}(x))\right| \le 2^{2\gamma}(1 + |j|h)^{-\gamma}(1 + |k - j|h)^{-\gamma}\Big(\|f\|_{0,\gamma}\|g''\|_{0,\gamma} + 2\|f'\|_{0,\gamma}\|g'\|_{0,\gamma} + \|f''\|_{0,\gamma}\|g\|_{0,\gamma}\Big),$$

which finally gives

$$\begin{aligned}
&\|\Pi_h(f) * \Pi_h(g) - \Pi_h(f * g)\|_{h,\gamma} \\
&\le \left(\sup_{k \in \mathbb{Z}} (1 + |k|h)^{\gamma}\right) \sum_{j \in \mathbb{Z}} \left| hf(jh)g((k-j)h) - \int_{I_{hj}} f(x)g(kh - x)dx \right| \\
&\le \frac{2^{2\gamma}}{24} h^2 \kappa_1(\gamma)\Big(\|f\|_{0,\gamma}\|g''\|_{0,\gamma} + 2\|f'\|_{0,\gamma}\|g'\|_{0,\gamma} + \|f''\|_{0,\gamma}\|g\|_{0,\gamma}\Big).
\end{aligned}$$

(ii) Expanding ϕ_k up to the fourth order we obtain as in the proof of the first part

$$\left| hf(jh)g((k-j)h) - \int_{I_{hj}} f(x)g(kh-x)dx - \frac{1}{24}h^3\phi_k''(jh) \right| \leq 2^{2\gamma}\frac{h^5}{1920}(1+|j|h)^{-\gamma}(1+|k-j|h)^{-\gamma}c_{4,\gamma}(f,g).$$

We have

$$h\sum_{j\in\mathbb{Z}}\phi_k''(jh) = (\Pi_h(f'')*\Pi_h(g) - 2\Pi_h(f')*\Pi_h(g') + \Pi_h(f)*\Pi_h(g''))(kh)$$
$$= \left(\Pi_h(f''*g - 2f'*g' + f*g'') + h^2 r_h\right)(kh)$$

with

$$\|r_h\|_{h,\gamma} \leq 2^{2\gamma}\kappa_1(\gamma)\left(c_{2,\gamma}(f'',g) + c_{2,\gamma}(f',g') + c_{2,\gamma}(f,g'')\right)$$

by part (i). Since $f''*g - 2f'*g' + f*g'' = 0$ we obtain the assertion on using the triangle inequality. □

Lemma A.4 *If $f \in C(4,\gamma)$ is a probability density, then*

$$\Pi_h(f)^{*n} = \Pi_h(f^{*n}) + h^4 r_h^{(n)},$$

where $\left(r_h^{(n)}\right)_{0<h\leq 1}, n \in \mathbb{N}$, *satisfies* (S).

Proof: We have

$$\Pi_h(f)^{*n} - \Pi_h(f^{*n}) = \sum_{j=1}^{n-1}\Pi_h(f)^{*(j-1)} * R_{h,n-j}$$

with

$$R_{h,j} := \Pi_h(f)*\Pi_h(f^{*j}) - \Pi_h(f*f^{*j}).$$

We know that $(\Pi_h(f)^{*n})_{0<h\leq 1}, n \in \mathbb{N}$, satisfies (S). By Lemma A.3 (ii)

$$\|R_{h,n}\|_{h,\gamma} \leq h^4 2^{2\gamma}\kappa_1(\gamma)c(f,n,\gamma),$$

with $c(f,n,\gamma)$ a linear combination of terms of the form

$$\|f^{(l)}\|_{0,\gamma}\left\|\left(f^{*(n-1)}\right)^{(k)}\right\|_{0,\gamma}, 0 \leq l,k \leq 4.$$

These grow subexponentially fast, hence $(h^{-4}R_{h,n})_{0<h\leq 1}, n \in \mathbb{N}$, also satisfies (S) and the assertion of the lemma follows since (S) is stable with respect to convolution. □

Lemma A.5 *If $f \in C(4,\gamma)$ is a probability density, then*

$$T_h(f)^{*n} = \Pi_h(f^{*n}) + \frac{nh^2}{24}\Pi_h(f)^{*(n-1)} * \Pi_h(f'') + h^4 r_h^{(n)},$$

where $\left(r_h^{(n)}\right)_{0<h\le 1}, n \in \mathbb{N}$, *satisfies* (S).

Proof: Expanding f about $x = kh$ we obtain

$$f(x) = f(kh) + (x-kh)f'(kh) + \frac{1}{2}(x-kh)^2 f''(kh) + \frac{1}{6}(x-kh)^3 f^{(3)}(kh) + \frac{1}{24}(x-kh)^4 f^{(4)}(\xi_k(x))$$

with some $\xi_k(x) \in I_{hk}$ if $x \in I_{hk}$. Integration over $x \in I_{hk}$ results in

$$\left(T_h(f) - \Pi_h(f) - \frac{h^2}{24}\Pi_h(f'')\right)(kh) = h^4 r_h(kh)$$

with

$$r_h(kh) := \frac{1}{24h^5}\int_{I_{hk}} (x-kh)^4 f^{(4)}(\xi_k(x))dx.$$

Continuing in the style of the proof of Lemma A.3 (i) and using $\int_{I_{hk}} (x-kh)^4 dx = h^5/80$ we obtain

$$\|r_h\|_{h,\gamma} \le \frac{2^\gamma}{1920}\|f^{(4)}\|_{0,\gamma}.$$

As in the proof of Lemma A.4 we now use a suitable decomposition of the difference,

$$T_h(f)^{*n} - \Pi_h(f)^{*n} = (T_h(f) - \Pi_h(f)) * \sum_{j=1}^{n-1} T_h(f)^{*(j-1)} * \Pi_h(f)^{*(n-j)}.$$

It is known that (S) holds for $(\Pi_h(f)^{*n})_{0<h\le 1}, n \in \mathbb{N}$. Similarly, (S) also holds for $(T_h(f)^{*n})_{0<h\le 1}$, $n \in \mathbb{N}$, as we even have $\|T_h(f)^{*n}\|_{h,1} = 1$ for all $h > 0$, $n \in \mathbb{N}$. Convolution stability gives (S) for the sum on the right hand side. Multiplying this sum by some $(r_h)_{0<h\le 1}$ satisfying (B) does not destroy (S). Hence, in summary, we have

$$T_h(f)^{*n} - \Pi_h(f)^{*n} - \frac{nh^2}{24}\Pi_h(f)^{*(n-1)} * \Pi_h(f'') = h^4 r_h^{(n)},$$

where $\left(r_h^{(n)}\right)_{0<h\le 1}, n \in \mathbb{N}$, satisfies (S). By Lemma A.4 we may replace $\Pi_h(f)^{*n}$ by $\Pi_h(f^{*n})$, and the statement of the lemma follows. □

A.4 For probability densities $f \in B_0(\gamma)$ and probability functions $a_h \in B_h(\gamma)$ (i.e. a_h has non-negative entries and $\|a_h\|_{h,1} = 1$) we define $\Psi(f), \Psi(a_h)$ by

$$\Psi(f) := \sum_{n=1}^{\infty} p_n f^{*n}, \qquad \Psi(a_h) := \sum_{n=1}^{\infty} p_n a_h^{*n}.$$

Again it should be clear from the context whether it is the continuous or the discrete version of the functional that Ψ refers to. We generally assume that the probabilities $(p_n)_{n \in \mathbb{N}_0}$ decrease at an exponential rate. Lemma A.1 (ii) and Lemma A.2 (ii) then imply that $\Psi(f)$ and $\Psi(a_h)$ are again elements of $B_0(\gamma)$ and $B_h(\gamma)$ respectively.

In terms of the operator and norm notation introduced above (3.4) can be written as

$$\left\| \frac{1}{h^2} \left(\Psi(T_h(f)) - \Pi_h(\Psi(f)) - \Pi_h(g) \right) \right\|_{h,\gamma} = O(h^2),$$

hence the theorem in Section 3 is a consequence of the following result.

Theorem A.6 *Let $f \in C(4, \gamma)$ be a probability density. Then*

$$g := \frac{1}{24} \sum_{n=1}^{\infty} n p_n f^{*(n-1)} * f'' \in B_0(\gamma)$$

and

$$\Psi(T_h(f)) = \Pi_h(\Psi(f)) + h^2 \Pi_h(g) + h^4 r_h,$$

where $(r_h)_{0<h\le 1}$ satisfies (B).

Proof: To see that g is an element of $B_0(\gamma)$ we use parts (i) and (ii) of Lemma A.1. With Lemma A.5 and the continuity of Π_h we obtain

$$\begin{aligned} \Psi(T_h(f)) - \Pi_h(\Psi(f)) &= \sum_{n=1}^{\infty} p_n (T_h(f)^{*n} - \Pi_h(f^{*n})) \\ &= \sum_{n=1}^{\infty} p_n \left(\frac{nh^2}{24} \Pi_h(f)^{*(n-1)} * \Pi_h(f'') + h^4 r_h^{(n)} \right) \\ &= \sum_{n=1}^{\infty} p_n \frac{nh^2}{24} \Pi_h(f)^{*(n-1)} * \Pi_h(f'') + h^4 r_{1,h} \end{aligned}$$

where (B) holds for $(r_{1,h})_{0<h\le 1}$ by the remark following the definition of (S). Similarly Lemma A.3 yields

$$\sum_{n=1}^{\infty} n p_n \Pi_h(f)^{*(n-1)} * \Pi_h(f'') = \sum_{n=1}^{\infty} n p_n \Pi_h\left(f^{*(n-1)} * f'' \right) + h^2 r_{2,h},$$

with $(r_{2,h})_{0<h\leq 1}$ satisfying (B). Put together, these two equations imply the assertion of the theorem. □

REFERENCES

ASMUSSEN, S. (1987) *Applied Probability and Queues*. Wiley, Chichester.

BOHMAN, H. (1977) Numerical Fourier inversion. *TIMS Studies in the Management Sciences* 7, 3-11.

BEARD, R.E., PENTIKÄINEN, T. and PERSONEN, E. (1984) *Risk Theory: The Stochastic Basis of Insurance*. (3rd ed.) Chapman and Hall, London.

EMBRECHTS, P., GRÜBEL, R. and PITTS, S.M. (1993) Some applications of the fast Fourier transform algorithm in insurance mathematics. *Statistica Neerlandica* **47**, 59-75.

GRANDELL, J. (1991) *Aspects of Risk Theory*. Springer, New York.

GRÜBEL, R. and HERMESMEIER, R. (1999) Computation of compound distributions I: aliasing errors and exponential tilting. *ASTIN Bulletin*, **29**, No. 2, 197-214.

HERMESMEIER, R. (1997) *Transformationsmethoden und Diskretisierungseffekte bei stochastischen Modellen in der Versicherungsmathematik*. Doctoral thesis, University of Hannover.

HIPP, C. and MICHEL, R. (1990) *Risikotheorie: Stochastische Modelle und Statistische Methoden*. Verlag Versicherungswirtschaft e.V., Karlsruhe.

KÖRNER, T.W. (1996) *The Pleasures of Counting*. Cambridge University Press, Cambridge.

STOER, J. (1994) *Numerische Mathematik I*, 7. Aufl. Springer, Berlin.

WALZ, G. (1996) *Asymptotics and Extrapolation Methods*. Mathematical Research **88**. Akademi Verlag, Berlin.

RUDOLF GRÜBEL and RENATE HERMESMEIER
Institut für Mathematische Stochastik
Universität Hannover
Postfach 60 09
D-30060 Hannover
E-mail: rgrubel@stochastik.uni-hannover.de