ARCH 2017.1 Proceedings

Split-Atom Convolution for Probabilistic Aggregation of Catastrophe Losses

Rafał Wójcik‡, Charlie Wusuo Liu, Jayanta Guin
Financial and Uncertainty Modeling Group
AIR Worldwide
Boston, MA 02116
‡Corresponding author: rwojcik@air-worldwide.com

***Abstract*—We present a new, fast algorithm for computing a probability density function of the sum of two independent discrete random variables in the context of catastrophe loss aggregation. The algorithm is a second-order approximation to brute force convolution. Computing speed comes from the fact of separating the atoms from main parts of two convolved distributions and from arithmetization (re-gridding) of the main parts. A new 4-point re-gridding procedure is presented as an alternative to classic linear and local moment matching methods. Since processing large portfolios after catastrophic events requires convolving loss distributions over millions of locations, we discuss the impact of the order of convolutions on the second-order moments of the resulting aggregate loss distribution. Further, we analyze scalability of our method with support size and investigate an extension based on Fast Fourier Transform. We illustrate the utility of our algorithm by performing ground-up loss aggregation using 100K distributions for hurricane and earthquake peril.**

***Index Terms*—Split-Atom Convolution, Loss Distributions, Arithmetization, Loss Aggregation, Catastrophe Modeling**

I. INTRODUCTION

The purpose of catastrophe modeling (known as CAT modelling in the industry) is to anticipate the likelihood and severity of catastrophe events from earthquakes and hurricanes to terrorism and crop failure, so companies (and governments) can appropriately prepare for their financial impact. Loss estimates produced from CAT models can be deterministic for a specific event or probabilistic from an ensemble of hypothetical events [1]. The latter approach uses Monte -Carlo (MC) techniques and physical models to simulate large ensemble of events. [2]. To pass from the ensemble to financial risk, the risk analysis sums up the event losses over the locations or properties in a particular portfolio. The losses are usually characterized by probability distributions. Here, a mixture of Transformed Beta and two atoms at min/max loss is used (see Fig.1). The atoms are not only an inherent feature inferred from analyzing claims data, but also, a result of applying financial terms – deductibles and limits composed into a variety of insurance and re-insurance tiers/structures. Processing large portfolios requires loss agrregation over millions of locatons. Fast computing techniques are needed to make the analysis feasible. Ideally, algorithms that (i) do not depend on large sample size and the corresponding storage/computer memory issues as MC methods in, e.g., [3], (ii) work well with irregular supports (of the order of 300 points) as opposed to Fast Fourier Transforms in [4] , and (iii) do not assume any parametric form of distribution of sum of losses in [5], are preferred. Here, we propose to estimate the compound distribution of CAT event losses by Split-Atom convolution. This method is the second–order moment approximation to the classical *brute force* (BF) approach in [6].

The paper is organized as follows. First, we introduce the Split-Atom convolution and new arithmetization (re–gridding) algorithm. Then, we discuss a concept of the order of convolutions which impacts the accuracy of preserving the second– and higher–order moments of aggregate distribution of losses. Finally, we show the results of an example of loss analysis by convolving 100K distributions for hurricane and earthquake peril.

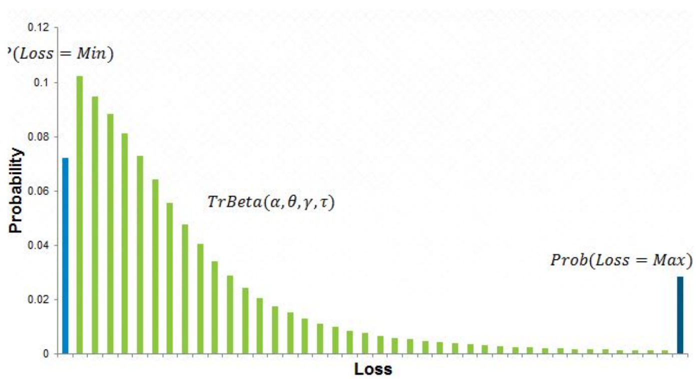

Fig. 1: A representation of loss distribution as a mixture of two atomic measures at min/max losses (blue) and main part (green) modeled as Transformed Beta.

II. LOSS AGGREGATION VIA SPLIT-ATOM CONVOLUTIONS

The density function $p_S(s)$ of the sum S of two independent discrete random variables X and Y characterizing CAT losses, with the densities p_X and p_Y respectively, can be calculated as:

$$p_S(s) = p_X \oplus p_Y = \sum_x p_X(x) p_Y(s-x) \qquad (1)$$

The most precise way of estimating (1) is BF convolution. The procedure computes all cross-products of probabilities and all cross-sums of losses. The complexity of this operation is $O(N_x N_y)$ where N_x and N_y are the number of points

Algorithm 1: Split-Atom Convolution: 9-products

Input : Two discrete probability density functions p_X and p_Y with irregular supports:
$x = \{x_1, x_2, x_2 + h_x, x_2 + 2h_x, \ldots, x_2 + (N_x - 3)h_x, x_{N_x}\}$; $y = \{y_1, y_2, y_2 + h_y, y_2 + 2h_y, \ldots, y_2 + (N_y - 3)h_y, y_{N_y}\}$ and probabilities: $p_X(x) = \sum_{i=1}^{N_x} \delta(x - x_i)p_X(x_i)$; $p_Y(y) = \sum_{j=1}^{N_y} \delta(y - y_i)p_Y(y_j)$

$x^{(1)} = \{x_1\}$; $p_X^{(1)}(x) = \delta(x - x^{(1)})p_X(x_1)$ // Split the left atom of p_X
$x^{(2)} = \{x_2, x_2 + h_x, x_2 + 2h_x, \ldots, x_2 + (N_x - 3)h_x, x_{N_x}\}$; $p_X^{(2)}(x) = \sum_{i=1}^{N_x-2} \delta(x - x_i^{(2)})p_X^{(2)}(x_i^{(2)})$
$y^{(1)} = \{y_1\}$; $p_Y^{(1)}(y) = \delta(y - y^{(1)})p_Y(y_1)$ // Split the left atom of p_Y
$y^{(2)} = \{y_2, y_2 + h_y, y_2 + 2h_y, \ldots, y_2 + (N_y - 3)h_y, y_{N_y}\}$; $p_Y^{(2)}(y) = \sum_{j=1}^{N_y-2} \delta(y - y_j^{(2)})p_Y^{(2)}(y_j^{(2)})$
Set $N_{s^\star}$ // maximum number of points for discretizing convolution grid
$h_{s^\star} = \frac{x_{N_x} + y_{N_y} - (x_1 + y_1)}{N_{s^\star} - 1}$ // theoretical step size of main part of convolution grid s
$h_s = \max(h_x, h_y, h_{s^\star})$ // final step size of main part of convolution grid s
if $h_s \geq (x_2 + y_2 - x_1 - y_1)$ **then** // set irregular convolution grid
  $s = \{x_1 + y_1, x_2 + y_2, x_2 + y_2 + h_s, \ldots, x_2 + y_2 + (N_s - 3)h_s, x_N + y_M\}$
else
  $s = \{x_1 + y_1, x_2 + y_2 - h_s, x_2 + y_2, \ldots, x_2 + y_2 + (N_s - 4)h_s, x_N + y_M\}$
end
$x^{(2,1)'} = \{x_2, x_2 + h_s, x_2 + 2h_s, \ldots, x_2 + (N_x - 3)h_s\}$ // redefine main part of $x^{(2)}$ with h_s
$p_X^{(2,1)'}(x) = \sum_{i=1}^{N'_x} \delta(x - x_i^{(2,1)'})p_X^{(2,1)'}(x_i^{(2,1)'})$ // re-grid $p_X^{(2)}$
$y^{(2,1)'} = \{y_2, y_2 + h_s, y_2 + 2h_s, \ldots, y_2 + (N_y - 3)h_s\}$ // redefine main part of $y^{(2)}$ with h_s
$p_Y^{(2,1)'}(y) = \sum_{j=1}^{N'_y} \delta(y - y_j^{(2,1)'})p_Y^{(2,1)'}(y_j^{(2,1)'})$ // re-grid $p_Y^{(2)}$
$x^{(2,2)'} = \{x_{N_x}\}$; $p_X^{(2,2)'}(x) = \delta(x - x_i^{(2,2)'}) \cdot p_X^{(2,2)'}(x_i^{(2,2)'})$ // split the right atom of p_X
$y^{(2,2)'} = \{y_{N_y}\}$; $p_Y^{(2,2)'}(y) = \delta(y - y_j^{(2,2)'}) \cdot p_Y^{(2,2)'}(y_j^{(2,2)'})$ // split the right atom of p_Y
$\mathcal{B}^{(1)} = p_X^{(2,1)'} \oplus p_Y^{(2,1)'}$ // BF convolution
$\mathcal{B}^{(2)} = p_X^{(1)} \oplus p_Y^{(1)}$ // ---''---
$\mathcal{B}^{(3)} = p_X^{(1)} \oplus p_Y^{(2,1)'}$ // ---''---
$\mathcal{B}^{(4)} = p_X^{(1)} \oplus p_Y^{(2,2)'}$ // ---''---
$\mathcal{B}^{(5)} = p_Y^{(1)} \oplus p_X^{(2,1)'}$ // ---''---
$\mathcal{B}^{(6)} = p_Y^{(1)} \oplus p_X^{(2,2)'}$ // ---''---
$\mathcal{B}^{(7)} = p_Y^{(2,1)'} \oplus p_X^{(2,2)'}$ // ---''---
$\mathcal{B}^{(8)} = p_Y^{(2,2)} \oplus p_X^{(2,1)'}$ // ---''---
$\mathcal{B}^{(9)} = p_Y^{(2,2)'} \oplus p_X^{(2,2)'}$ // ---''---
Regrid $\mathcal{B}^{(1-9)}$ onto convolution grid s

Output: Discrete probability density function p_S of independent sum $S = X + Y$ with support defined as:
$s = \{s_1, s_2, s_2 + h_s, s_2 + 2h_s, \ldots, s_2 + (N_s - 3)h_s, s_{N_s}\}$ and the associated probabilities as:
$p_S(s) = \sum_{k=1}^{N_s} \delta(s - s_k)p_S(s_k)$, where $s_{N_s} - [s_2 + (N_s - 3)h_s] \leq h_s$, $s_2 - s_1 \leq h_s$, $h_s \geq \max(h_x, h_y)$.

discretizing p_X and p_Y supports, respectively. An extra cost $N_xN_ylog(N_xN_y)$ is due to redundancy removal [6]. Clearly, it is not a good candidate for convolving CAT loss distributions due to speed and storage reasons. Our solution to (1) stems from an idea of separating two atoms at min/max losses and arithmetizing (re-gridding) the main part of loss distribution (green, Fig. 1). The former guarantees preservation of min/max losses at different loss perspectives (e.g. insured, insurer, re-insurer, FAC underwriter etc.) and the latter replaces $O(N_xN_y)$ with $O(N'_xN'_y)$ where $N'_x < N_x$ and/or $N'_y < N_y$. Splitting the atoms and compacting the main part of distributions results in a variety of convolution algorithms depending on how the original and convolution grids are defined and what computing speed and memory requirements are. Algorithm 1 (9-product Split–Atom), is an example. The original grids are non-uniform: spacing between the atoms and main part is arbitrary. Convolution grid is set the same way to assure that min/max losses are exactly preserved. Depending on the application, other possibilities are conceivable e.g. 9-products preserving not only min/max losses but also their associated probabilities or splitting only the right atom (4-product approach) for computing speed.

A. New re-gridding method

It follows from Section II-A that the computing speed of Split-Atom convolution is induced by re-gridding. The simplest way of re-gridding is linear binning [7], [8]. This method, however, only preserves the mean of the re-gridded distribution. If higher order moments are of concern, *local moment matching* [9] should be used. This algorithm sometimes produces negative probability mass. A linear optimization fix based on constrained simplex approach is discussed in [10]. We propose faster and more accurate alternative: 4-point re-gridding showed in Fig.2. Instead of solving, 2-point based linear system in [10] the algorithm redistributes the original probability mass not only to the nearest neighboring points on new, coarse support, but also to

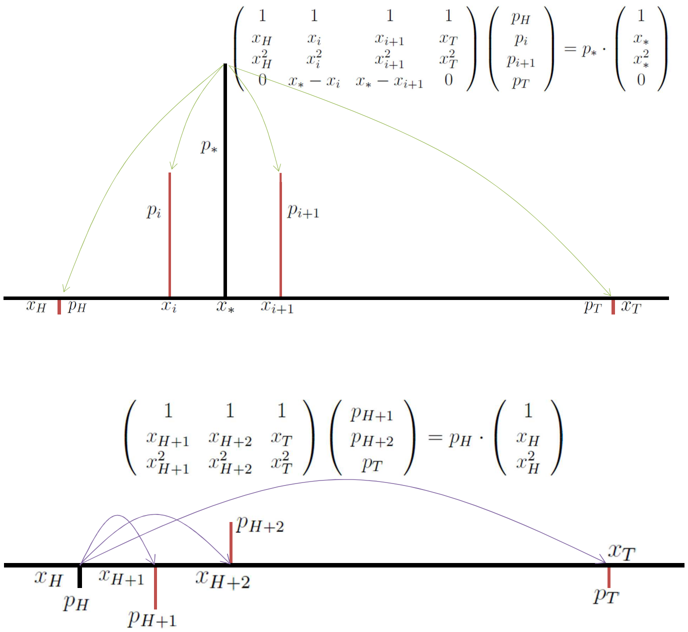

Fig. 2: 4-point re-gridding of $p^\star$ at $x^\star$ (upper panel) and mitigation of possible negative probability mass of head p_H at x_H (lower panel).

the head and tail points by recursively solving a linear system in the upper panel of Fig.2. This guarantees preservation of the second-order moments exactly (unlike [10]) at the price of generating negative probabilities at the head and tail points on the new grid. To mitigate this behavior, yet another linear system is solved in the lower panel of Fig.2. If a distribution undergoing the re-gridding is defined on sparse (˜10 points) support, this method might fail. In that case we use simple linear re-gridding [7], [8].

B. Order of convolutions

The re-gridding algorithm above, when combined with convolution, might introduce error in the second–order moments of the distribution of the sum of losses. To minimize this error, convolutions should be ordered in such a way that two distributions to be convolved are defined on more or less the same supports (equal max losses) with the same grid step ($h_x \approx h_y$ in Algorithm 1). There are several possibilities we considered: (i) sequential, where distributions are convolved according to the order of locations in a particular portfolio, (ii) sorted, where distributions are sorted in ascending order using max losses, (iii) closest pair, where depth-first search method is used to define a convolution tree and (iv) variation on the theme of balanced multi–way number partitioning referred to as the balanced largest–first differencing method (BLDM) [11]. The latter is a powerful strategy that seeks to split a collection of numbers (in this case max losses of the set of distributions to be convolved) into subsets with (roughly) the same cardinality and subset sum.

C. Use of the Fast Fourier Transform (FFT)

A further speed up of Split-Atom convolution can be achieved by applying convolution theorem for the FFT to convolve the main parts of p_X and p_Y (line 19 in Algorithm 1) instead of using BF method. We can write:

$$p_X^{(2,1)'} \oplus p_Y^{(2,1)'}(s) \overset{FFT}{\rightarrow} P_X^{(2,1)'}(\omega) P_Y^{(2,1)'}(\omega) \overset{IFFT}{\rightarrow} p_S^{(2,1)'}(s) \quad (2)$$

where ω represents frequency and $IFFT$ stands for the Inverse Fast Fourier Transform. For optimal performance, implementation of (2) requires that probabilities $p_X^{(2,1)'}$ and $p_Y^{(2,1)'}$ are defined on the same support discretized by a power of 2 number of points. If this is not the case, we simply extend the supports and zero-pad the probabilities.

III. Results

To demonstrate the loss analysis with Split-Atom convolution we performed the following experiment. We first re-sampled (without replacement) 100K loss distributions for hurricane and earthquake peril. The lookup tables with distributions are typically comprised of 16K distributions characterizing damage ratios (loss/replacement value) for a particular peril. The distributions are stratified by the mean damage ratio and defined on 64-point grid. Replacement values were sampled from $U(0, \$5 \times 10^6)$. The maximum number of points discretizing convolution grid was 256.

Linear Regrid

	Convolution \ Order of Convolution	sortedSequential	sequential	closestPair	BLDM
SD (theoretical=190)	Split-Atom: 4 products	9239	9238	784	921
	Split-Atom: 9 products	9165	9164	782	923
	Brute Force	6729	6724	576	590
Time(seconds)	Split-Atom: 4 products	0.73	0.69	1.32	1.07
	Split-Atom: 9 products	0.86	0.83	1.44	1.21
	Brute Force	7.72	7.66	16.83	18.03

4-point Regrid

	Convolution \ Order of Convolution	sortedSequential	sequential	closestPair	BLDM
SD (theoretical=190)	Split-Atom: 4 products	314	296	190	190
	Split-Atom: 9 products	303	285	190	190
	Brute Force	190	190	190	190
Time(seconds)	Split-Atom: 4 products	0.78	0.72	1.72	1.44
	Split-Atom: 9 products	1.25	1.19	2.13	1.69
	Brute Force	22.97	22.67	55.77	64.28

Table 1: The results of convolvig 100K loss distributions discretized on 64-point grid using three convolution methods: Split–Atom 4–products, 9–products (Algorithm 1) and Brute Force with linear (upper panel) and 4-point regridding (lower panel) with different orders of convolutions. We compare compute times and standard deviations (SD). Theoretical means were preserved exactly in all cases. Both losses and their probabilities were represented as doubles. Intel i7-4770 CPU @ 3.40GHz architecture with 16GB RAM was used.

Then, we convolved the re-sampled distributions using Split–Atom 4–products, 9–products (Algorithm 1) and BF approach with linear and 4-point re-gridding. The various solutions to the order of convolutions were implemented as described in Section II-B. The results are shown in Table 1. It is clear that 4-point re-gridding outperforms linear approach in terms of mitigating the huge overestimation of theoretical SD. This is particularly emphasized in the case of closest pair

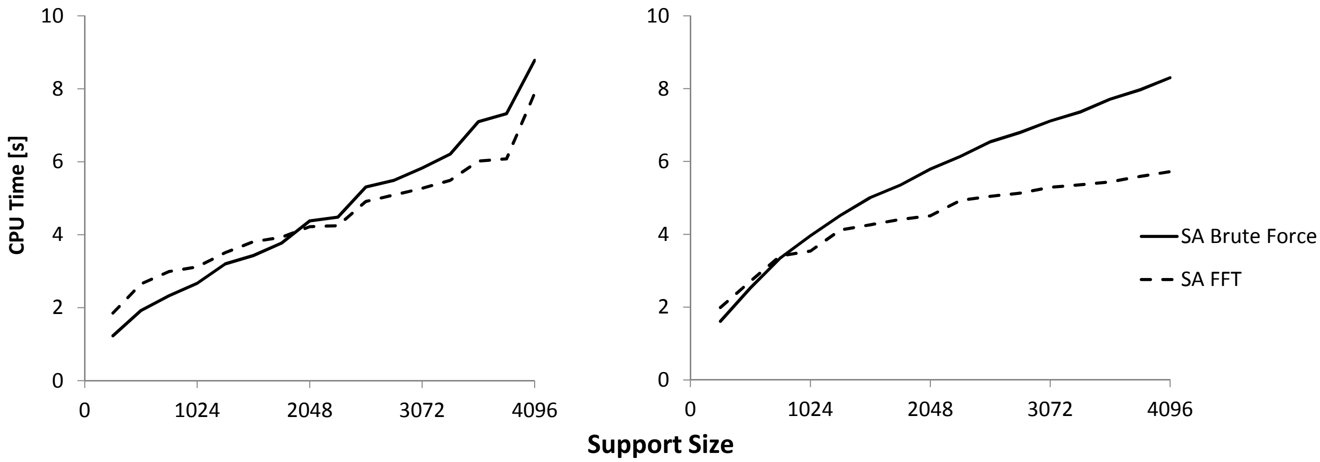

Fig. 3: Runtimes of 4-product Split-Atom (SA) convolution implemented using BF (solid line) and FFT (dashed line) as a function of the size of convolution grid. The results are based on convolving 100K earthquake and hurricane loss distributions with 4-point re-gridding and BLDM (left panel) vs. closest pair (right panel) order of convolutions. Intel i7-4770 CPU @ 3.40GHz architecture with 16GB RAM was used.

and BLDM order of convolutions where the error in SD has been eliminated. For our new re-gridding method, BLDM is also computationally fastest strategy preserving the theoretical second-order moments exactly when combined with 4-product Split–Atom convolution. Figure 3 shows computing time of Split-Atom convolution vs maximum number of points on convolution grid (support size) for both BLDM (left) and closest pair (right) order. The results confirm that Split-Atom convolution using BF implementation is the fastest strategy for CAT loss aggregation. This is due to the fact that, in practice, the maximum number of points on convolution grid never exceeds 256 for computing speed reasons. When application (e.g. processing smaller portfolios) allows larger support size, the use of closest pair order with FFT implementation is recommended. The application of FFT gives a gain in speed by a factor of about 1.5 for support size 4096 points. This indicates that the majority of speed up comes from re-gridding of main parts of two convolved distributions.

IV. Conclusions

In this paper we presented a new, fast method for convolving probability density functions characterizing losses from CAT events. This method is a second-order approximation to BF convolution. Using an example of loss aggregation for 300K hurricane and earthquake perils we have shown that Split-Atom convolution with 4-point re-gridding, executed in BLDM and closest pair order, outperforms classical BF approach. The latter is $\sim$ 30 times slower which makes it infeasible for CAT modeling applications.

References

[1] K. Clark, "Catastrophe risk," in *IAA Risk Book – Governance, Management and Regulation of Insurance Operations*. International Actuarial Association / Association Actuarielle Internationale, 2015.

[2] S. Latchman, "Quantifying the risk of natural catastrophes," 2010. [Online]. Available: http://understandinguncertainty.org/node/622

[3] P. Arbenz, C. Hummel, and G. Mainik, "Copula based hierarchical risk aggregation through sample reordering," *Insurance: Mathematics and Economics*, vol. 51, no. 1, pp. 122–133, 2012.

[4] J. Robertson, "The computation of aggregate loss distributions," in *Proceedings of the Casualty Actuarial Society*, vol. 79, 1992, pp. 57–133.

[5] H. H. Panjer, "Recursive evaluation of a family of compound distributions," *ASTIN Bulletin: The Journal of the International Actuarial Association*, vol. 12, no. 01, pp. 22–26, 1981.

[6] D. Evans and L. Leemis, "Algorithms for computing the distributions of sums of discrete random variables," *Mathematical and Computer Modelling*, vol. 40, no. 13, pp. 1429–1452, 2004.

[7] M. C. Jones and H. W. Lotwick, "On the errors involved in computing the empirical characteristic function," *Journal of Statistical Computation and Simulation*, vol. 17, no. 2, pp. 143–149, 1983.

[8] P. Hall and M. Wand, "On the accuracy of binned kernel density estimators," *Journal of Multivariate Analysis*, vol. 56, no. 2, pp. 165–184, 1996.

[9] H. Panjer and G. Willmot, *Insurance Risk Models*. Society of Actuaries, 1992.

[10] J. Vilar, "Arithmetization of distributions and linear goal programming," *Insurance: Mathematics and Economics*, vol. 27, pp. 113–122, 2000.

[11] J. Zhang, K. Mouratidis, and H. Pang, "Heuristic algorithms for balanced multi-way number partitioning," in *International Joint Conference on Artificial Intelligence (IJCAI)*. Research Collection School Of Information Systems, 2011, pp. 693–698.