Dynamic Initial Margin Estimation Based on Quantiles of Johnson Distributions

Thomas A. McWalter*[,a,b], Jörg Kienitz[a,c,d], Nikolai Nowaczyk[e], Ralph Rudd[a] and Sarp K. Acar[d]

[a]The African Institute of Financial Markets and Risk Management (AIFMRM), University of Cape Town
[b]Department of Statistics, University of Johannesburg
[c]Fachbereich Mathematik und Naturwissenschaften, Bergische Universität Wuppertal
[d]Quaternion - Quantitative Services Acadia
[e]Independent Researcher

September 24, 2018
(Latest revision July 1, 2022)

Abstract

The estimation of dynamic initial margin (DIM) is a challenging problem. We describe an accurate new approach using Johnson-type distributions, which are fitted to conditional moments, estimated using a least-squares Monte Carlo simulation (the JLSMC algorithm). We compare JLSMC DIM estimates with those computed using an accurate nested Monte Carlo simulation as a benchmark, and with another method that assumes portfolio changes are Gaussian. The comparisons reveal that the JLSMC algorithm is accurate and efficient, producing results that are comparable with nested Monte Carlo while using an order of magnitude less computational effort. We provide illustrative examples using the Hull-White and Heston models for different derivatives and portfolios. A further advantage of our new approach is that it relies only on the readily available data that is needed for any exposure or XVA calculation.

Keywords: dynamic initial margin (DIM); margin value adjustment (MVA); quantiles; Johnson distributions; least squares Monte Carlo.

The authors thank the organisers and participants of The Factory, held at the University of Cape Town (July 2018), for critical feedback; in particular: Sam Cohen, Marc Henrard, Martin Larsson, Andrea Macrina, Erik Schlögl, David Taylor and James Taylor. We also thank Bernhard Hientzsch for identifying a few errors in an earlier draft of the paper.

*Corresponding author

1 Introduction

Following the economic and financial crisis of 2007/8 there has been a coordinated global regulatory response of the Basel Committee on Banking Supervision (BCBS) and the International Swaps and Derivatives Association (ISDA) that has increasingly lead to mandatory margining of bilateral over-the-counter (OTC) derivative trades. The volume in this market is very large[1] and the lifetime of the contracts can be very long.[2] Consequently, significant resources have been allocated to the computation of collateral estimates for such trades. In the present paper we describe a novel method that provides efficient and accurate estimates for the dynamic initial margin associated with such trades.

The counterparty credit risk in a derivative trade is a result of outstanding cash flows that are not yet settled. Consequently, financial regulation promotes, and increasingly requires, counterparties to collateralize their derivative trades. This collateralization is implemented in two ways to cover two distinct sources of risk. The first source is the *current exposure*, that is the immediate loss suffered by the surviving counterparty. This is mitigated by the counterparties exchanging *variation margin (VM)*, which is roughly equal to the present value of the trades. A lesson learned from the financial crisis is that even a fully VM-collateralized trade has some non-zero credit risk, as the surviving counterparty needs some time to close out positions after default. This close-out time is called the *margin period of risk (MPOR)* and is typically 10 business days. As the trades can move quite adversely during the MPOR, additional collateral, called *initial margin (IM)* is required to mitigate this *potential exposure*. Initial margin is posted at inception, but, despite its name, it is adjusted (potentially daily) over the lifetime of the trades, in the same manner as VM.

The BCBS requires that the IM collateral for bilateral OTC derivatives should be at least at the level of a 99% quantile of the changes in value of the trades over the MPOR, see (Basel Committee on Banking Supervision 2015, Sect. 3.1)[3]. This requirement poses the practical challenge of accurately computing not only the spot IM, but also of forecasting future IM levels. The latter problem is called *dynamic initial margin (DIM)*. Even for a bank running an internal model approach, computing such a quantile naively using a general nested Monte Carlo simulation is too computationally expensive for practical purposes.[4] The problem of developing an easy-to-implement, yet accurate, DIM model is the subject of the present paper.

There are a number of characteristics that any new algorithm should possess. First, the approach should use easily obtainable data such as quantities used for computing any other exposure or XVA calculation. Second, it should be efficient, requiring significantly less computational resources than the gold standard of nested Monte Carlo. Third, it should provide more accurate results than other existing approximations. We shall show that our new method meets these criteria when compared to other approaches proposed in the literature.

[1] The total outstanding notionals are about USD 500 trillion, see Bank of International Settlements (2017) for detailed statistics.

[2] Maturities of several decades are not uncommon in the interest rate derivatives market.

[3] Technically this reference only applies to non-cleared bilateral OTC derivatives. The IM for derivatives cleared through a central counterparty (CCP), or for exchange traded derivatives (ETDs), is set by the CCP or exchange respectively. While the details of these methodologies for computing IM are proprietary, the guideline of setting the IM at a conservative quantile still seems to apply.

[4] The fact that margin methods used for uncleared vs. cleared derivatives are different, and that methods may differ between clearing houses, complicates matters further.

1.1 State of the Art

ISDA has proposed a simple method for computing spot IM, called ISDA SIMM, see International Swaps and Derivatives Association (2016). This method relies on computing the quantile via a sensitivity based VaR, using parameters set by ISDA, which are recalibrated yearly. This method has the advantage that the two trading counterparties only have to compute the sensitivities of the trades, alleviating the problem of margin disputes. Moreover, the method can be used by smaller banks who might not have an internal model. Computing the associated DIM, i.e., forecasting future IM levels under ISDA SIMM, is therefore reduced to the problem of computing forward sensitivities of derivatives. This problem is challenging in itself and has a lot of research devoted to it, in particular in the context of fast sensitivity computation using adjoint algorithmic differentiation (AAD). Although many banks have struggled to implement AAD in practice, the requirement to compute DIM has sparked new and promising research in this area, see for example Antonov (2017); Antonov et al. (2017a,b); Fries (2017b,a,c); and Caspers and Lichters (2018).

A modelling approach that does not require forward sensitivities, but only the cube of *net present values (NPVs)* of trades, has been introduced in Anfuso et al. (2017). This *Gaussian Least-Squares Monte Carlo (GLSMC)* approach uses regression techniques and the assumption that the changes in value of the trades over the MPOR are approximately Gaussian in distribution. An evaluation of the approach using an open source risk engine[5], carried out in Caspers et al. (2017), found that while for trades with linear payoffs, like interest rate swaps, this model performs well, it performs poorly with complex portfolios of more exotic trades.

The practical business impact of IM collateralizing derivative trades is two-fold: On the one hand, as the collateral for IM can usually not be rehypothecated[6], both counterparties incur funding costs. These funding costs are immediate, mandatory and devalue the trades by a significant *margin value adjustment (MVA)*. On the other hand, as the collateralization reduces the credit risk in the trades, both parties can in theory save capital. However, the savings in capital can only be realized by a bank if it manages to develop and implement a DIM model, prove that that the model is conservative and obtain regulatory approval to use it. As pointed out in Nazneen (2017), the industry is still struggling to develop an easy to implement yet accurate DIM model to realize the savings, yet is starting to suffer from the funding costs, which makes this an increasingly pressing problem. In addition, estimating the possible savings for individual banks depends on subtle modelling questions, for instance how to treat unpaid cash flows during the MPOR, see Andersen et al. (2016, 2017). The debate as to what extent collateralization reduces systemic risk is the subject of ongoing research, see O'Halloran et al. (2017, 2019).

1.2 Method and Approach

Based on the guidelines set by the Basel Committee on Banking Supervision (2015), we present a formal mathematical definition of the DIM problem in Section 2. Three approaches for computing DIM estimates are described in Section 3, being Nested Monte Carlo, Gaussian Least-Squares Monte Carlo and our new proposed method, the Johnson Least-Squares Monte Carlo (JLSMC) Algorithm. The first approach is guaranteed to produce an unbiased estimate

[5] See `https://www.opensourcerisk.org`.

[6] This means that a bank cannot re-use the IM collateral received from one counterparty to post it to another.

for DIM, but is prohibitively expensive from a computational perspective to be practical. The second is the above-mentioned approach by Anfuso et al. (2017).

The main contribution of our paper appears in Section 3.3, which provides the specification of an easy-to-implement, efficient and accurate new method. The JLSMC algorithm is based on fitting Johnson (1949) distributions to the first four conditional moments of the portfolio changes over the MPOR, as determined using a least-squares Monte Carlo simulation. With the Johnson curves fitted to these conditional moments, we can fit a parametric function to the quantiles of the Johnson curves thus providing a functional form for the conditional quantile. DIM is estimated using this parametric function. As with the approach of Anfuso et al. (2017), our proposed approach only requires the estimated future values of the portfolio under consideration for each simulated path. These values are already computed when banks estimate their exposure under various scenarios. Thus, all the values that are necessary for our approach can be assumed available.

We evaluate the accuracy of our algorithm in Section 4. Section 5 concludes the paper with a discussion of the approach and its limitations and further avenues of research.

1.3 Summary of Results

We test JLSMC for various models and products, and compare it to GLSMC, with nested Monte Carlo and, where possible, analytic solutions as benchmarks.

In Section 4.1 we analyse a European put option on a stock under geometric Brownian motion. A detailed comparison of the GLSMC and JLSMC approaches shows that JLSMC produces significantly more accurate DIM estimates, when compared to a nested Monte Carlo estimated and the analytical solution, see Figure 2. This example shows that, even for vanilla derivatives (a put option), the inaccuracy of the GLSMC algorithm is driven primarily by violation of its assumption of normality—the conditional density of returns is far from normal and hence Gaussian quantiles are bound to be inaccurate.

In Section 4.2 we study various interest rate products (caplet, floorlet, swaptions and a portfolio thereof) under a Hull-White model, see Figures 4 and 5. In Section 4.3 the Heston model is used for underlying stock dynamics and we study various equity products (puts and calls and portfolios thereof), see Figures 6 and 7. For both these models and related products, we observe that the JLSMC estimates of DIM differ significantly from GLSMC estimates, the former agreeing with unbiased nested Monte Carlo estimates and analytic values, where available.

2 Problem Formulation

Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a maturity time, $T > 0$, and a filtration of $\mathscr{F}$ given by $(\mathscr{F}_t)_{0 \le t \le T}$. Let X_t be an $\mathscr{F}_t$-adapted multivariate process, which represents the state variables in our financial system, for instance stock prices, stochastic variance, interest rates etc.[7] Finally, let V_t, which we think of as the value of a portfolio of assets and claims, be an $\mathscr{F}_t$-adapted process that is a function of the state variables. The algorithm that we shall

[7] Some of these quantities are observable, some of them may be traded and some of them are neither traded nor observable. As an aside, risk practitioners often think in terms of *risk factors*, RF, which are market observable. Formally, RF_t is a multivariate process that is adapted to $\mathscr{F}_t$, consisting of a function of a subset of the state variables. Generally, knowledge of RF_t does not uniquely specify state, i.e., the filtration generated by RF_t is a sub-filtration of $\mathscr{F}_t$.

later introduce is agnostic concerning the relationship between the portfolio value and the state variables.

We define the IM and DIM of V_t, using a *margin period of risk* (MPOR) given by $\delta < T-t$, as

$$\mathsf{IM}_{t,\delta} := Q_\alpha(\Delta V_t | \mathscr{F}_t) \tag{1}$$

and

$$\mathsf{DIM}_{t,s,\delta} := \mathbb{E}^{\mathbb{P}}[\mathsf{IM}_{s,\delta} | \mathscr{F}_t], \tag{2}$$

respectively, where $\Delta V_t := V_{t+\delta} - V_t$ and Q_α is the conditional quantile at the α-level of confidence (for instance $\alpha = 0.99$), see Appendix A.1. Here, for simplicity, we ignore the problem of handling cash flows over the MPOR.[8] The financial meaning of $\mathsf{DIM}_{t,s,\delta}$ is that it is the expected level at time t of the initial margin applicable at time $s > t$.

An example of the above setup, which we use later, is that X, the vector of state variables, is specified by a system of SDEs of the form

$$dX_t = \mu_t(X_t)\, dt + \sigma_t(X_t)\, dW_t, \qquad X_0 = x_0, \tag{3}$$

where W_t is a multivariate Brownian motion and $\mu_t(\cdot)$ and $\sigma_t(\cdot)$ are appropriately defined vector and matrix valued functions. The filtration $\mathscr{F}_t$ is then the augmentation of the filtration generated by W_t. We model $\mathscr{F}_t$-adapted *asset prices*, $S_t = (S_t^0, \ldots, S_t^k)$, as functions of the state variables, choose $B_t := S_t^0$ as the numéraire and assume that there exists at least one risk-neutral measure, $\mathbb{Q} \sim \mathbb{P}$, for the numéraire denominated assets. We let $H_t = (H_t^1, \ldots, H_t^l)$ be contingent claims on S_t, which can then be risk-neutrally priced using

$$H_t = B_t \mathbb{E}^{\mathbb{Q}} \left[\tfrac{H_T}{B_T} \mid \mathscr{F}_t \right]. \tag{4}$$

Finally, $V_t := \sum_i a_i S_t^i + \sum_j b_j H_t^j$ is the value of a portfolio of assets and derivatives, for some constants $a_j, b_j \in \mathbb{R}$.

The $\mathsf{IM}_{t,\delta}$ feeds into two relevant business quantities: First, the *expected positive exposure* (EPE), which in turn feeds into the *risk weighted assets* (RWA)—a key quantity in determining regulatory capital reserves, and thus the cost of capital. Assuming that VM_t is a process modelling the variation margin, the exposure may be computed as

$$\mathsf{E}_t = (V_t - \mathsf{VM}_{t-d} - \mathsf{IM}_{t-d,\delta})^+,$$

where d is the period of time that has elapsed since the last margin posting and the realisation of the exposure. Here we are making the "classical" assumption that both counterparties cease to post margin at the same time (see Andersen et al. (2017)). Since d may vary depending on the details of the default, we shall assume, for the purposes of modelling EPE, that $d = \delta$. This leads to the following expression

$$\mathsf{EPE}_{0,t} := \mathbb{E}^{\mathbb{P}}[(V_t - \mathsf{VM}_{t-\delta} - \mathsf{IM}_{t-\delta,\delta})^+ \mid \mathscr{F}_0].$$

[8] Otherwise one would have to replace ΔV_t by $\Delta V_t + \mathrm{CF}_t$ where CF_t are the cash flows during the MPOR.

Second, the *margin value adjustment* (MVA) may be computed as

$$\mathsf{MVA}_t := B_t \int_t^T \mathbb{E}^{\mathbb{P}} \left[f_s \frac{\mathsf{IM}_{s,\delta}}{B_s} \mid \mathscr{F}_t \right] ds, \tag{5}$$

where f_s is the funding spread at time s.

Note that the above two equations (and some that follow below) use expectations under $\mathbb{P}$. In this paper, we do not address how the market price of risk is determined/specified, for example, by using standard or stressed scenarios in combination with historical calibration (time series analysis) or market-implied calibration. Consequently, for the purposes of benchmarking our algorithms, we shall perform all expectations under the risk-neutral measure. In other words, we shall assume that $\mathbb{P}$ and $\mathbb{Q}$ are the same.

3 Numerical Algorithms

Having specified a general definition for DIM in (2), we now consider the slightly less general case where DIM is computed at the present time,

$$\mathsf{DIM}_{0,t,\delta} = \mathbb{E}^{\mathbb{P}}[\mathsf{IM}_{t,\delta} | \mathscr{F}_0]. \tag{6}$$

Note, as seen in (1), the quantile associated with the initial margin is conditioned on $\mathscr{F}_t$.

3.1 Nested Monte Carlo

A straightforward approach for estimating DIM is traditional nested Monte Carlo simulation. Here, we shall assume that the system of SDEs governing our state variables is Markovian.[9] Thus, when conditioning on the filtration $\mathscr{F}_t$ we can rather condition on the state variables at time t. To produce an estimate for (6), given the state at time 0 (x_0), we proceed as follows. An outer Monte Carlo experiment simulates realizations $x_t|x_0 \sim X_t$, from which realizations of the portfolio value at time t, denoted $v_t|x_0$, are computed. Conditional on each one of the outer realizations of x_t, the i^{th} realization denoted by x_t^i, an inner (nested) Monte Carlo experiment simulates realizations of the state variables $x_{t+\delta}|x_t^i$, with corresponding $v_{t+\delta}|x_t^i$ values, from which realisations of $\Delta v_t|x_t^i$ may be computed. From each inner Monte Carlo sample, a sample quantile, $\widehat{Q}_\alpha(\Delta v_t|x_t^i)$, is computed. The estimate for DIM is then computed as

$$\mathsf{DIM}_{0,t,\delta} = \mathbb{E}^{\mathbb{P}}\left[Q_\alpha(\Delta V_t|\mathscr{F}_t)|\mathscr{F}_0\right] \approx \tfrac{1}{N}\sum_{i=1}^{N} \widehat{Q}_\alpha(\Delta v_t|x_t^i),$$

where N is the number of outer Monte Carlo samples. It is clear that the nested Monte Carlo algorithm converges to the correct solution, but, owing to the nesting, the computational complexity quickly becomes intractable for a reasonable number of inner and outer simulations. Therefore, we take a different approach.

[9] This is not an egregious assumption as the models generally used in practice are Markovian, although they may rely on state variables that are not directly observable and must be estimated using market data.

3.2 Gaussian Least-Squares Monte Carlo (GLSMC)

It is well known that conditional expectations may be estimated using least-squares Monte Carlo (LSMC), obviating the need for nested Monte Carlo. This is easily implemented when there is a single state variable X_t. When multiple state variables are present, it is convenient to use the portfolio value as the regressor. Thus we wish to find the quantiles conditional on V_t rather than X_t. This requires a projection of the filtration $\mathscr{F}_t$ on to the smaller filtration $\mathscr{F}_t^V$ generated by V_t.

For simplicity, we denote the resulting quantile by $Q_\alpha(\Delta V_t|V_t)$. This implies that there exists a measurable function $f_\alpha : \mathbb{R} \to \mathbb{R}$ such that, for all $v \in \mathbb{R}$, we have $f_\alpha(v) = Q_\alpha(\Delta V_t|V_t = v)$, and thus

$$\mathbb{E}^{\mathbb{P}}[Q_\alpha(\Delta V_t|V_t)|\mathscr{F}_0] = \mathbb{E}^{\mathbb{P}}[f_\alpha(V_t)|\mathscr{F}_0].$$

The function $f_\alpha(v)$ may be estimated using least-squares Monte Carlo (Andersen et al. 2016; Anfuso et al. 2017; Caspers et al. 2017). It is common practice in the literature (including by the aforementioned authors) to assume that portfolio changes, $\Delta V_t|V_t$, are normally distributed with mean zero. In this case, it is enough to use LSMC to find a (parametric) functional form for the second moment of conditional portfolio changes.

To see this, consider the measurable function, $M_2 : \mathbb{R} \to \mathbb{R}$, that represents the functional form of the conditional second moment of portfolio changes

$$M_2(v) = \mathbb{E}^{\mathbb{P}}[(\Delta V_t)^2|V_t = v].$$

With knowledge of this function, the conditional quantile may be approximated as

$$f_\alpha(v) \approx \Phi^{-1}(1-\alpha)\sqrt{M_2(v)} \tag{7}$$

under the above normality assumptions. The LSMC algorithm is the mechanism to provide the approximate parametric functional form for the conditional second moment, which through (7) then yields the approximate quantile. We shall refer to this approach as Gaussian least-squares Monte Carlo (GLSMC).

Of course, any quantile estimate based on this assumption will only be accurate if the underlying portfolio changes are indeed close to normal. This is often not the case in many common situations, with certain asset classes exhibiting non-normal and fat-tailed price changes over short periods. For this reason, we now propose a novel approach that makes use of higher order moments.

3.3 The Johnson Least-Squares Monte Carlo Algorithm (JLSMC)

We now present an approach that uses higher moment information to find a better approximation for $f_\alpha(v)$. The idea is to find parametric functional forms for the first four moments, conditional on V_t, given by

$$M_i(v) = \mathbb{E}^{\mathbb{P}}[(\Delta V_t)^i|V_t = v], \qquad \text{for } i = \{1,2,3,4\}.$$

With these, conditional values of mean, standard deviation, skewness and kurtosis may be estimated on a stratified sample of V_t. Johnson (1949) curves are fitted using these conditional

values from which the conditional quantiles may be computed in closed form.[10] Using these conditional quantiles, we then approximate f_α via regression using a parametric function $f_\alpha : \mathbb{R}^{R+1} \times \mathbb{R} \to \mathbb{R}$, with $R+1$ parameters specified by the vector $\boldsymbol{\beta}^{(\alpha)} \in \mathbb{R}^{R+1}$. That is

$$f_\alpha(\boldsymbol{\beta}^{(\alpha)}, v) \approx Q_\alpha(\Delta V_t | V_t = v).$$

This function is specified using orthogonal polynomials, with the parameters determined by regression, as described in Appendix A.2. The DIM may then be estimated, using a Monte Carlo sample of size n, as

$$\widehat{\mathsf{DIM}}_{0,t,\delta} = \frac{1}{n} \sum_{i=1}^{n} f_\alpha(\boldsymbol{\beta}^{(\alpha)}, v_i), \tag{8}$$

where $v_i \sim V_t$.

The detailed algorithm now follows.

1. Generate n Monte Carlo sample pairs, $(V_t, \Delta V_t | V_t)$, and let $\boldsymbol{X}$ be the vector of the V_t values.
2. For $i \in \{1, 2, 3, 4\}$:
   2.1. Let $\boldsymbol{Y}^{(i)}$ be the vector of values $(\Delta V_t)^i | V_t$.
   2.2. Let $M_i : \mathbb{R}^{S+1} \times \mathbb{R} \to \mathbb{R}$ be a parametric function, specified using $S+1$ parameters in vector $\boldsymbol{\beta}^{(i)} \in \mathbb{R}^{S+1}$, which approximates the i-th raw conditional moment of ΔV_t, i.e.,
   $$M_i(\boldsymbol{\beta}^{(i)}, v) \approx \mathbb{E}[(\Delta V_t)^i | V_t = v].$$
   Perform regression on $\boldsymbol{Y}^{(i)}$ and the function $M_i(\boldsymbol{\beta}^{(i)}, \boldsymbol{X})$ to estimate $\boldsymbol{\beta}^{(i)}$, see Appendix A.2 for details.
3. Compute $q_j = Q_{j/N}(\boldsymbol{X})$, for $0 < j < N$, being evenly-spaced sample quantiles of V_t. Optionally: Add N' additional points in the tails of the distribution.
4. For each $0 < j < N + N'$:
   4.1. Compute the conditional raw moments $\mu'_{i,j} = M_i(\boldsymbol{\beta}^{(i)}, q_j)$, $i \in \{1, 2, 3, 4\}$.
   4.2. Using raw moments $\mu'_{i,j}$, compute mean, variance, skewness and kurtosis, and fit Johnson curve j using these, see Appendix A.3 for details.
   4.3. Compute $Q_{\alpha|j}$, being the α-quantile of Johnson curve j.
5. Let $\boldsymbol{Q}$ be the vector of q_j values and $\boldsymbol{Y}$ the vector of corresponding $Q_{\alpha|j}$ values. Perform least-squares regression on $\boldsymbol{Y}$ and $f_\alpha(\boldsymbol{\beta}^{(\alpha)}, \boldsymbol{Q})$ to estimate $\boldsymbol{\beta}^{(\alpha)}$.
6. Compute $\widehat{\mathsf{DIM}}_{0,t,\delta}$ using (8) with the samples $f_\alpha(\boldsymbol{\beta}^{(\alpha)}, \boldsymbol{X})$.

Here, like f_α, the functions M_i, determining the first four raw conditional moments in the algorithm, are specified using orthogonal polynomials with weights determined using least-squares regression. As with any Monte Carlo method, a suitably large sample size should

[10] Incidentally, instead of computing conditional quantiles, it should be possible to compute expected shortfall, since the functional forms are known.

be used to produce reasonable estimates, i.e., with low variance. This is especially true for reliable estimation in the tails of the conditional moments, where a small sample size may mean that outliers have an undue influence on the conditional moments. Empirically, we have found that sample sizes, n, of at least 100 000 are needed to produce reliable estimates. Another consideration is the number, N, of evenly spaced quantiles. Since the payoffs of derivatives may substantially vary with changes in price, it is necessary to select the number of Johnson curve fits in such a way that changes in the conditional moments are well represented, with a modest value of N around 100 performing well. Because the polynomial fits to these quantiles (step 5 above) may severely deviate from the points fitted at the left and right edges, we suggest adding N' extra additional quantiles in the tails. Here we suggest that N' should be around 10% of N.

4 Numerical Results

4.1 Geometric Brownian Motion

In order to confirm the accuracy of the JLSMC algorithm we consider a DIM estimation based on a portfolio consisting of a single put option on a stock driven by geometric Brownian motion.

We generate geometric Brownian motion realizations under the standard SDE

$$dS_t = rS_t\,dt + \sigma S_t\,dW_t, \qquad S_0 = s_0,$$

where $r = 0.05$ is the constant interest rate, $\sigma = 0.3$ is the volatility and $s_0 = 100$ is the initial stock price. The realizations are generated at times $t = 1/12$ and $t+\delta$, where $\delta = 1/24$ is the MPOR (specified in the same units as t, i.e., years). The put option expires at $T = 1$ year and is struck at 95.

Figure 1 shows four panels that depict the conditional sample moments based on least-squares Monte Carlo. In this figure, the light blue points indicate the samples (sample size $n = 200\,000$). The red points indicate the fitted parametric forms for each of the four moments, which were selected as second-order Laguerre polynomials.[11]

Based on the parametric functions describing these raw moments, Johnson distributions are fitted to the quantiles of the portfolio value at t, $q_j = Q_{j/N}(\boldsymbol{X})$, $0 < j < N$, where $N = 100$ was used. To more accurately represent the tail dynamics, we augmented the set $\{q_j\}$ with outlying points determined by subdividing the 1% and 99% quantiles of $\boldsymbol{X}$ into 5 evenly spaced quantiles each.

Using the conditional quantiles implied by the 110 fitted Johnson distributions, shown as blue crosses in the left of Figure 2, f_α may be computed ($\alpha = 0.99$). This was done by a fitting a forth-order Laguerre polynomial and is displayed as a blue line. For comparison, a $1000 \times 100\,000$ uniformly stratified nested Monte Carlo simulation was performed, with the samples displayed as black circles. In the same panel, the functional form of the GLSMC algorithm is a red line. Furthermore, because the portfolio is monotonic in the single underlying, it is possible to numerically integrate an analytic expression for the DIM. This is displayed as a black line.

[11] While we have chosen to use Laguerre polynomials, other sets may be used, for example Chebyshev, Gegenbauer, Hermite, Jacobi and Legendre polynomials, amongst others.

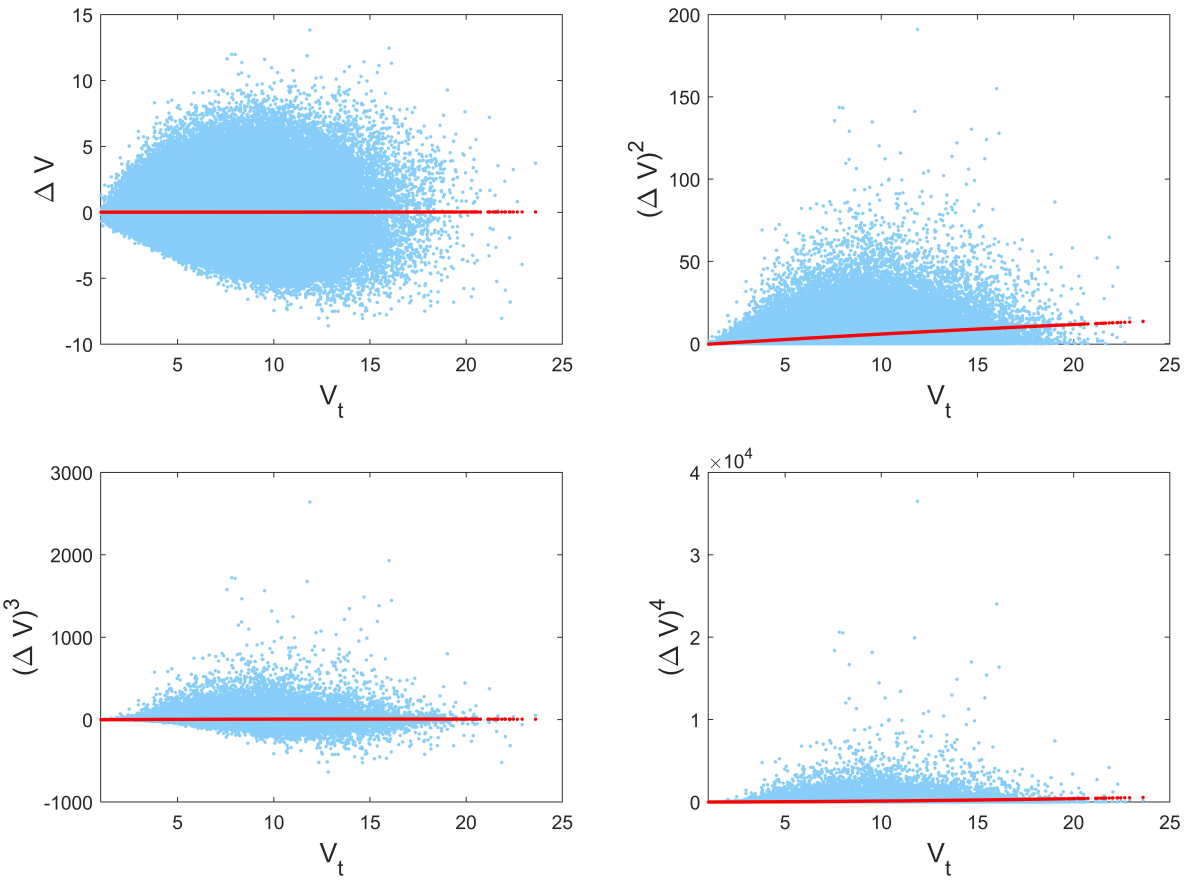

Figure 1: LSMC estimates for the first four conditional raw moments of changes in value of a put option over the MPOR. As a function of value V_t, samples, being powers of changes in value $(\Delta V_t)^i$, are shown in blue with conditional raw moment estimates, $\mathbb{E}[(\Delta V_t)^i | V_t]$, shown in red.

Because of the existence of an analytic solution, the accuracy of the algorithms can be directly compared. The error in the JLSMC algorithm is 0.65%, whereas for the normal approximation it is 13.26%.[12] For this stratified, nested Monte Carlo simulation the error is 0.0004%.

We note here that while the nested Monte Carlo simulation is very accurate it takes approximately 15 times longer to compute than the JLSMC algorithm, in this example. This is one of the key benefits of using least-squares Monte Carlo approaches, they obviate the need for the computational inefficiencies of nested Monte Carlo methods. We shall highlight this advantage in the more sophisticated examples explored later in the paper (see Table 1 for further details).

The right panel of Figure 2 shows the conditional densities fitted at $X_t = 100$, or correspondingly, $V_t = 6.875$. The blue line indicates the resultant Johnson density, with the red the conditional normal distribution. The dashed magenta line is a kernel-smoothed density

[12] Errors are the absolute difference between the quoted value and the analytical value expressed as a percentage of the analytical value.

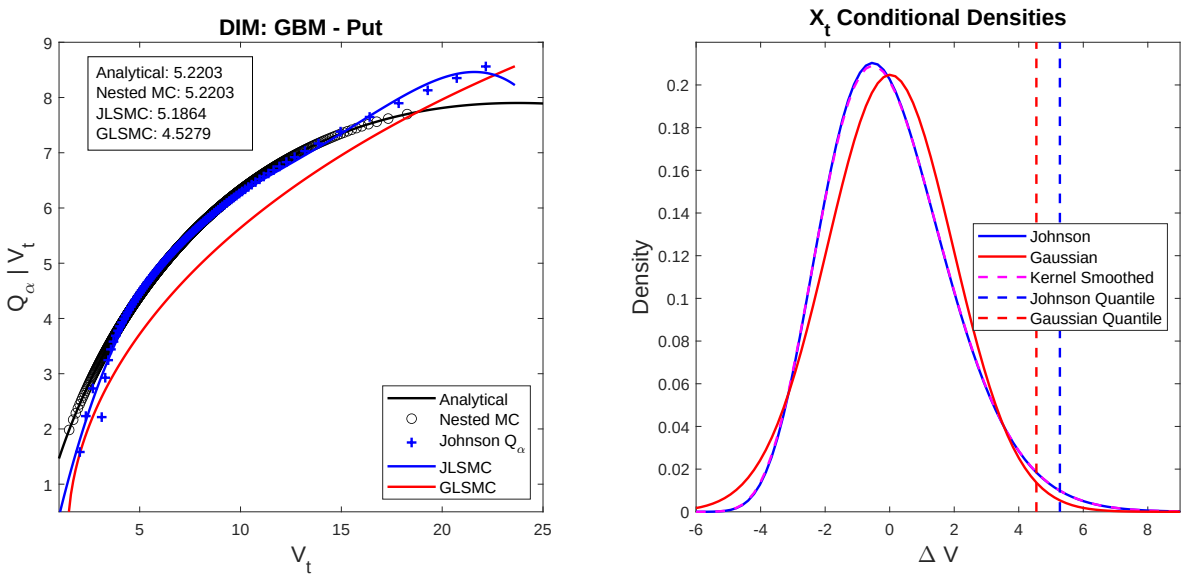

Figure 2: The left panel shows the conditional quantile, $Q_\alpha|V_t$, of changes in value of a put option over an MPOR of $1/12$ for $\alpha = 0.99$, as a function of the realised value, V_t, using different methods (JLSMC, GLSMC, nested Monte Carlo and analytical quantiles). The top legend gives the DIM values computed using the different methods. The right panel shows the conditional densities at $X_t = 100$ ($V_t = 6.875$), being the probability of a realization of ΔV, with conditional quantiles shown as dashed lines.

fit to the nested Monte Carlo samples.

The vertical dashed blue and red lines indicate the α-quantiles implied by the fitted Johnson distribution and fitted Gaussian distribution respectively. Thus, the distance on the x-axis between these two lines is the same as the vertical distance between the blue and red lines at $V_t = 6.875$ in the left panel of the figure. It is clear that the normal density is not a good fit to this conditional distribution, as both the fitted Johnson curve and kernel-smoothed density display a fatter right tail.

In the above example we have used an MPOR of two weeks, but there are cases where the MPOR may be significantly longer. In the left panel of Figure 3, using the same put option as before, we show the conditional densities for an MPOR of $\delta = 1/12$ (one month). This shows that the situation gets worse when the MPOR is increased, which is generally the case—if the Gaussian approach performs poorly for a short MPOR, it performs worse for a longer MPOR.

While up to this point we have tackled the relatively 'vanilla' case of a put option, the results carry over to more exotic options. For example, in the right panel of Figure 3, we show the results for a down-and-out call option, using a strike of 110 and lower knock-out level of 75, with $\delta = 1/24$.

For the remainder of Section 4, the parameters used for the JLSMC and GLSMC approaches remain the same as above. In summary, the number of samples is $n = 200\,000$, the order of Laguerre polynomials for four conditional moments is $S = 2$, the number of sample quantiles of V_t is $N = 100$, with $N' = 10$ additional tail quantiles, and the order of Laguerre polynomials specifying the conditional quantile is $R = 4$.

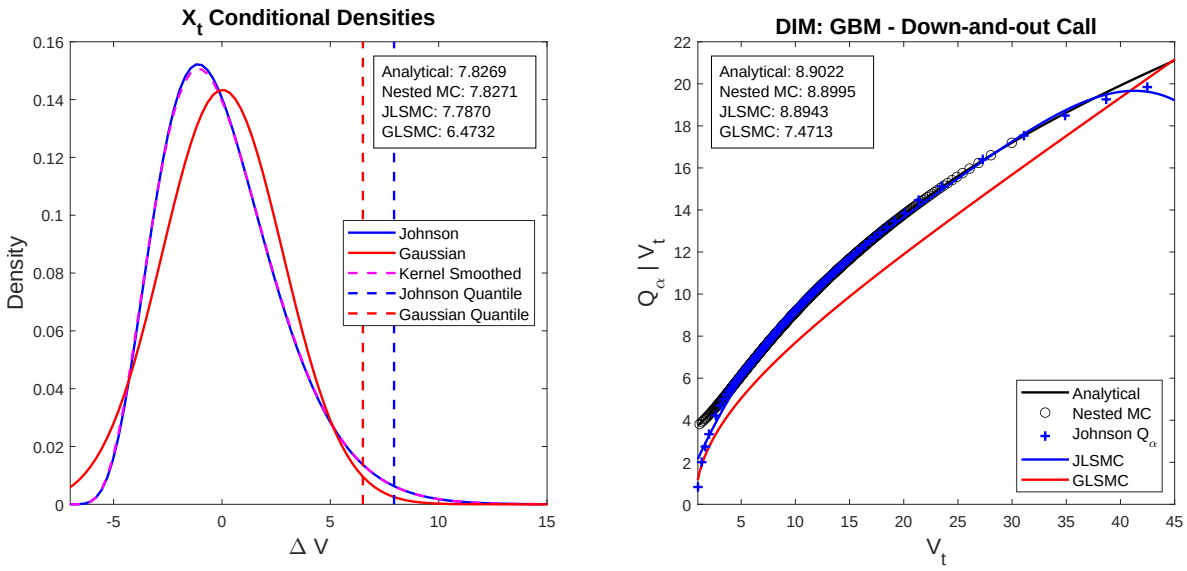

Figure 3: The left panel shows the conditional densities at $X_t = 100$ ($V_t = 6.875$), being the probability of a realization of ΔV for the put option over an MPOR of $1/12$, with conditional quantiles shown as dashed lines. The right panel shows the conditional quantile, $Q_\alpha|V_t$, of changes in value of a down-and-out call option over an MPOR of $1/24$ for $\alpha = 0.99$, as a function of the realised value, V_t, using different methods. In both panels the top legends give the DIM values computed.

4.2 The Hull-White Model

We now consider the Hull and White (1990) model, specifically the version analysed in Hull and White (1994), with only one time-varying parameter,

$$dr_t = (\theta(t) - ar_t)\,dt + \sigma\,dW_t, \tag{9}$$

with a and σ positive constants and $\theta(t)$ specified to exactly fit the market-observed term-structure of interest rates. To ensure that our results are replicable, we specify a parametric form for the initial yield curve given by $y(t) = C_1 + C_2\exp(C_3 t)$. Note that the initial short rate, r_0, is specified by this initial term structure. For our examples, we used $a = 0.015$, $\sigma = 0.01$, $C_1 = 0.05$, $C_2 = -0.03$ and $C_3 = -0.18$.

The Hull-White model is a useful introduction because closed-form expressions exist for common instruments, like caps, floors, bond options and swaptions Brigo and Mercurio (2007). In the left panel of Figure 4 we consider a caplet with a maturity of 9 months, written on the 3-month (or quarterly) rate. The notional is 10 000 and the strike is set at 2.5%.

The 99%-quantile, conditional on the portfolio value one month from the present, is displayed, with the MPOR set at 10 days. For this example, the nested Monte Carlo simulation can generate exact samples of both the short rate and the corresponding discount factor (see Ostrovski (2013)) at any future time. The outer Monte Carlo simulation (to determine the portfolio values one month from now) used 1000 samples, whereas the inner Monte Carlo simulation (to determine the conditional quantiles) used 10 000 samples per outer path. Uniform stratification was used for both the outer and inner paths to reduce the variance of the

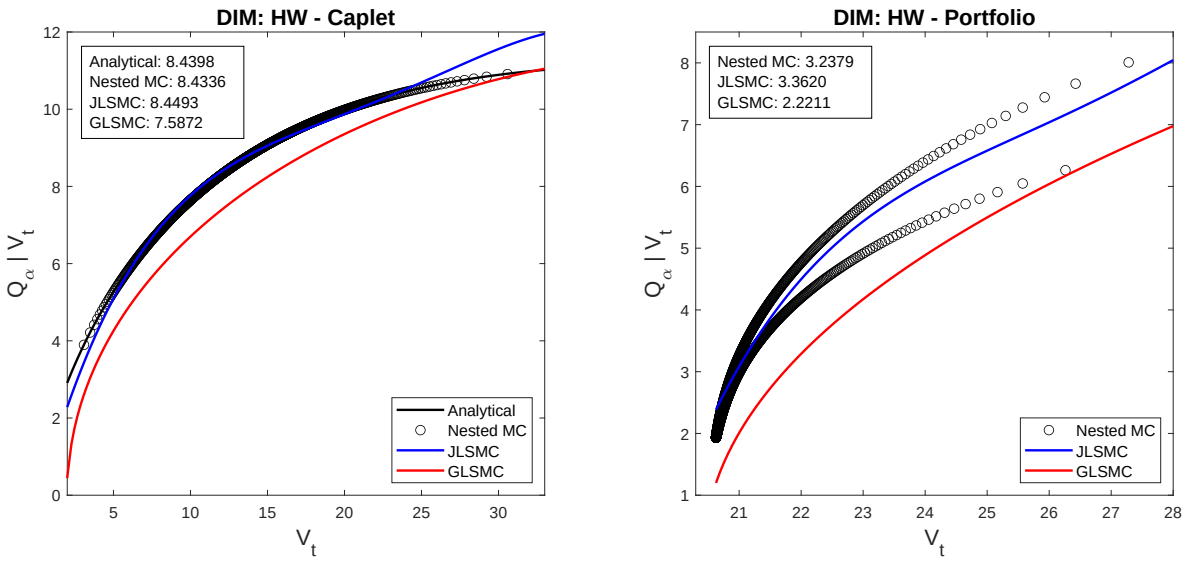

Figure 4: DIM results for the Hull-White model when considering a caplet (left panel), and a portfolio of a caplet and floor (right panel). Both panels show the conditional quantile, $Q_\alpha|V_t$, of changes in value, ΔV, over an MPOR of 1/12 for $\alpha = 0.99$, as a function of the realised value, V_t, using different methods. DIM values are reported in the top legends.

estimate.

Since this a single factor model, and the value of a caplet is monotonic with respect to the underlying short rate, an analytic DIM value is available for comparison. It is clear that both the JLSMC algorithm, with an error of 0.11%, and the nested Monte Carlo algorithm, with an error of 0.073%, yield very accurate DIM results, with the JLSMC algorithm being an order of magnitude faster to compute. The GLSMC algorithm deviates significantly from the true value, with an error of 10.10%, indicating that the Gaussian assumption is suspect for even a very simple portfolio of a single caplet.

In the right panel of Figure 4, we consider a portfolio consisting of the sum of the same caplet and a floor. The floor commences a year from the present, lasts for one year and is written on the quarterly rate. Thus there are four payments, made in arrears. The floor is struck at-the-money. The portfolio has a notional of 10 000 with 75% invested in the caplet and 25% invested in the floor.

This panel is interesting for two reasons: first, the portfolio value is no longer monotonic in the underlying rate, so an analytical DIM is unavailable, and second, there are now clearly two states of the underlying short rate that yield the same portfolio value in the future. Thus, as can be seen from the nested Monte Carlo simulation, there are two quantile values corresponding to each of the two states of the underlying that yield the same portfolio value. When considering the GLSMC or JLSMC approaches, there is, however, only a single quantile value specified by the functional form for the quantile *conditional on the portfolio value*. Thus, in this instance, the JLSMC approach gives a conditional quantile that falls between the two realised by nested Monte Carlo. We will describe this phenomenon further in the next section when we look at the Heston model. Note that in this example the JLSMC is two orders of magnitude faster to compute than the nested Monte Carlo simulation.

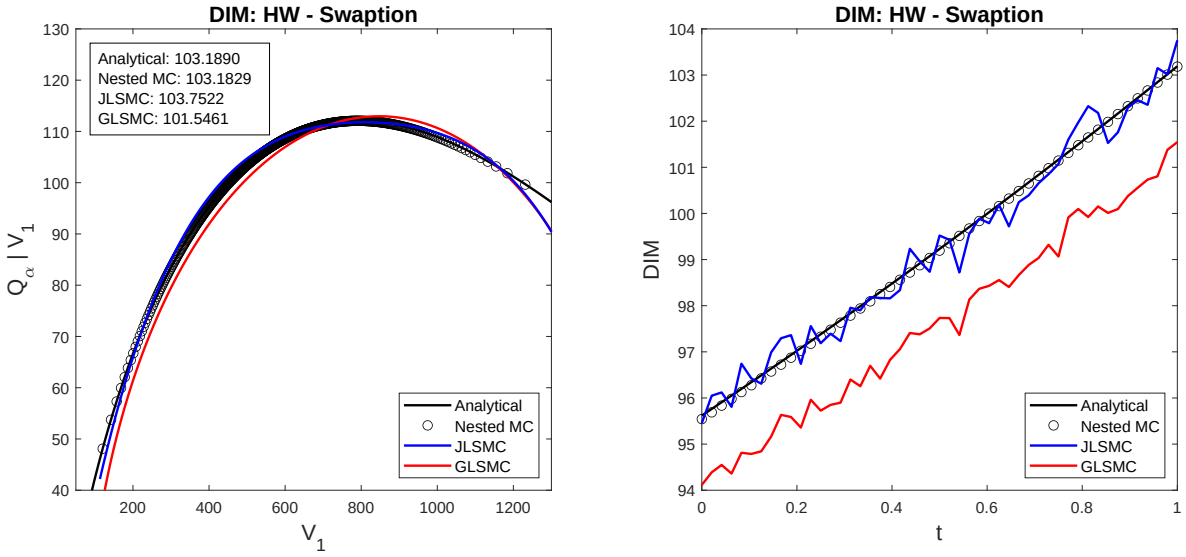

Figure 5: The left panel shows the conditional quantile, $Q_\alpha|V_t$, of changes in value of a swaption under the Hull-White model over an MPOR of 1/12 for $\alpha = 0.99$ and $t = 1$, as a function of realised value, V_t, using different methods. DIM values are reported in the top legend. The right panel shows DIM values for the swaption as a function of t.

Figure 5 displays the time-evolution of DIM for a swaption in the Hull-White model. The swaption matures in 5 years and the underlying swap has a 5-year term, written on the quarterly rate. This 5×5 swaption has a notional of 10 000 and a strike of 4%. DIM is computed at weekly increments out to one year with the MPOR kept at 10 days. The nested Monte Carlo simulation used 500 outer paths and 10 000 inner paths.

The left panel displays the conditional quantiles for the final DIM evaluation. The right panel shows the evolution of the DIM value over time. It is this curve that would be integrated to obtain an MVA estimate. Note that again, the portfolio is monotonic in the state variable, and thus analytic values can be obtained for both the DIM and MVA.

4.3 The Heston Model

In this section we consider computing DIM for various products under the Heston (1993) stochastic volatility model. The standard dynamics are given by

$$\begin{aligned} dS_t &= rS_t\,dt + S_t\sqrt{V_t}\,dW_t^s, & S_0 &= s_0, \\ dV_t &= \kappa(\theta - V_t)\,dt + \sigma\sqrt{V_t}\,dW_t^v, & V_0 &= v_0, \end{aligned}$$

where

$$\langle dW_t^s, dW_t^v \rangle = \rho\,dt.$$

Here, r is the interest rate, κ the mean reversion rate of the variance, θ the long term variance, σ the volatility of variance and the Brownian motions are correlated with correlation parameter $\rho \in [0, 1]$. The initial stock price and initial variance are given by s_0 and v_0 respectively. These parameters were set to $s_0 = 100$, $v_0 = 0.035$, $r = 5\%$, $\kappa = 3$, $\theta = 0.1$,

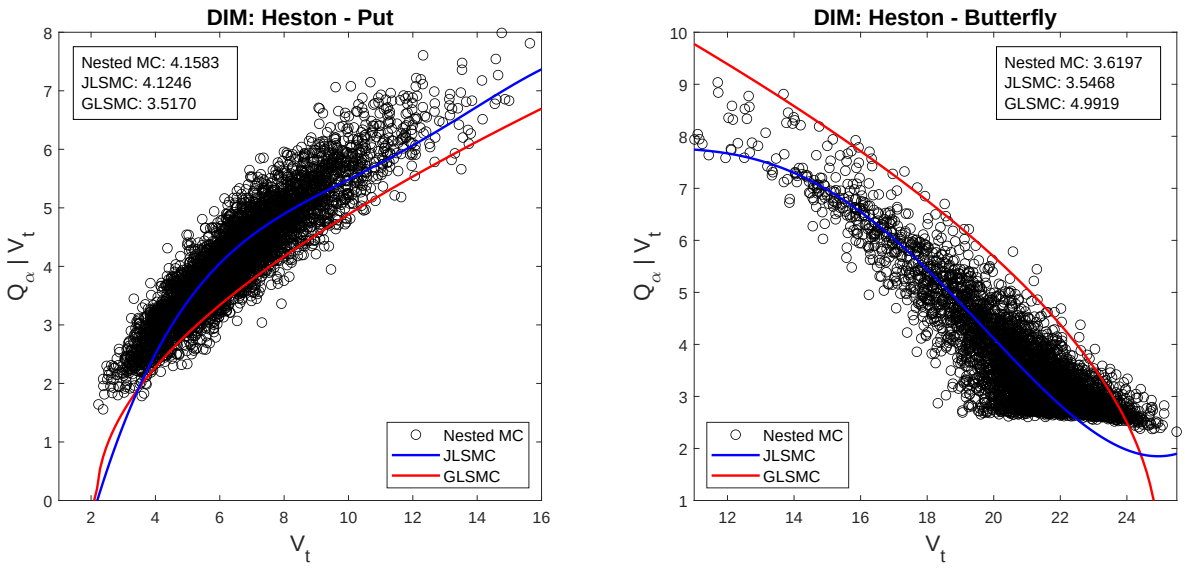

Figure 6: The DIM results for the Heston model when considering a put (left), and a butterfly spread (right). Both panels show the conditional quantile, $Q_\alpha|V_t$, of changes in value, ΔV, over an MPOR of $1/12$ for $\alpha = 0.99$, as a function of the realised value, V_t, using different methods. DIM values are reported in the top legends.

$\sigma = 0.25$ and $\rho = -0.4$.

The left panel of Figure 6 shows the conditional 99%-quantiles for a 1-year put option with strike 95 computed one month from now with the MPOR fixed at 10 days. The nested Monte Carlo simulation used 5 000 outer paths and 5 000 inner paths. To simulate paths of the Heston model, we used the quadratic exponential scheme from Andersen (2008), which is neatly summarized in Rouah (2013). As this is a short step method, we used 10 steps per week or 2 per day. To price the European options considered under the Heston model we used the second volatility of volatility series expansion of Lewis (2000).

Note that now the conditional quantiles produced by the nested Monte Carlo simulation no longer lie on a single curve, but rather form a scattered cloud of points. This is because the portfolio value, V_t, is exclusively determined by the stock price, while the distribution of a portfolio change over the MPOR is determined by both the stock price and the realised variance. Thus, portfolios with the same value on the x-axis, i.e., conditioned on V_t, can produce many different estimates of a 99%-quantile, since these quantiles are dependent on both underlying state variables. It is encouraging that the JLSMC curve is fitted through this cloud of points, greatly outperforming the curve provided by the normality assumption.

If we assume that the nested Monte Carlo simulation provides the correct DIM estimate, then the JLSMC has an error of 0.81% as opposed to the error made by the normal assumption of 15.42%.

In the right panel of Figure 6, we consider a call butterfly spread. It is a portfolio of call options constructed by purchasing one in-the-money call option struck at 90, one out-the-money call option struck at 110, and shorting two at-the-money call options. The notional invested in the portfolio is 10 and all the options mature in six months. Again, the JSLMC DIM is very accurate compared to the normal assumption. This is the first conditional

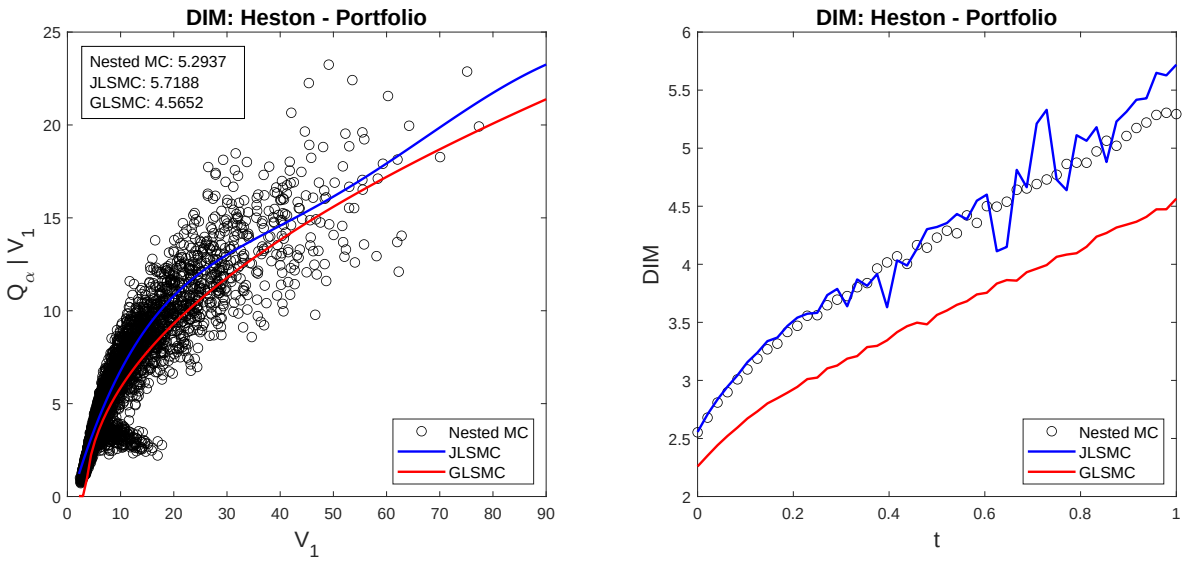

Figure 7: The left panel shows the conditional quantile, $Q_\alpha|V_t$, of changes in value of a European option portfolio under the Heston model over an MPOR of 1/12 for $\alpha = 0.99$ and $t = 1$, as a function of realised value, V_t, using different methods. DIM values are reported in the top legend. The right panel shows DIM values for the portfolio as a function of t.

quantile curve we see that decreases as the portfolio value increases. If we again assume that the nested Monte Carlo simulation provides the correct DIM estimate, then the JLSMC has an error of 2.01% and the GLSMC an error of 37.91%.

In Figure 7 the time-evolution of DIM is displayed for a portfolio far out-the-money options under the Heston model. The portfolio consists of a call option, with maturity 2 years and strike 120, a put option, with maturity 1.5 year and strike 80. The portfolio notional is 1, with 60% invested in the call, 40% in the put. The nested Monte Carlo simulation consisted of 5 000 outer paths and 4 000 inner paths.

It is clear for Figure 7 that the time-series of DIM values estimated using the JLSMC algorithm is more accurate than the Gaussian approach. This is because of the far from normal portfolio changes, which mean that DIM estimates not incorporating higher order moment information are inaccurate. While the accuracy of the nested Monte Carlo algorithm may be considered the gold standard, as it is an unbiased estimator, the computational effort used to produce such estimates may be prohibitive. Even though we have used relatively modest sample sizes for the nested Monte Carlo (in examples which are relatively simple), the computational effort is large in comparison with both least-squares Monte Carlo methods. This can be seen in Table 1, which shows how many times slower nested Monte Carlo is when compared with the JLSMC algorithm.

5 Conclusion

In the present paper, we have introduced a methodology to estimate DIM based on regression using Johnson-type distributions—the JLSMC algorithm. The advantage of this approach is that it makes use of readily available quantities that are needed to price the underlying

Fig. 2 (left)	15.0
Fig. 3 (right)	18.5
Fig. 4 (left)	12.7
Fig. 4 (right)	13.2
Fig. 5 (left and right)	12.1
Fig. 6 (left)	18.7
Fig. 6 (right)	20.2
Fig. 7 (left and right)	16.2

Table 1: Speed-up factor computed as the time taken by the nested Monte Carlo approach divided by the time taken by the JLSMC approach.

portfolio. To provide a benchmark for the approach, we have compared the JLSMC approach to a nested Monte Carlo implementation and analytical solutions. The results of the nested Monte Carlo and JLSMC approaches are shown to be coherent when modelling European options under the dynamics of geometric Brownian motion.

We have also compared the DIM estimates determined using JLSMC with the standard Gaussian regression-based approach (GLSMC) of Anfuso et al. (2017), which makes the assumption that the portfolio changes are normally distributed with zero mean. The benchmarking exercise shows that making the Gaussian assumption leads to less accurate DIM values, as compared with JLSMC. When the empirical profit and loss distribution is compared to the parametric distribution implied by GLSM, it is noted that there is significant deviation from normality. However, the Johnson distribution, when fitted using all four moments of the data, provides a much better fit. Furthermore, these results carry over for exotic options.

In further numerical experiments, we have considered the Hull-White and Heston models for interest rates and equity, and estimated DIM values for portfolios of caplets, floorlets, swaptions and equity options using the GLSMC and JLSMC approaches. As expected the DIM estimates derived from these approaches are quite different, with JLSMC providing significantly more accurate results, when compared with nested Monte Carlo estimates. Again, this is to be expected since the normal assumption made by GLSMC does not adequately capture the distributional properties of portfolio profits and losses.

While we have focused on introducing a more accurate method for estimating DIM, there are several ancillary issues that could be explored—these affect all the methods explored in this paper. For example, we have only built simple portfolios with static weights and have not investigated the effect of any dynamic hedging or offsetting of risks within asset classes. We have also not investigated any portfolios that incorporated cash flows (see, for example, Andersen et al. (2016), which provides various modelling approaches for incorporating cash flows).

Since the proposed JLSMC method is agnostic to the underlying model, it can be used as an alternative to DIM estimation based on forward sensitivities, while providing an efficient computational alternative to the full nested Monte Carlo approach.

References

Andersen L. (2008). Simple and efficient simulation of the Heston stochastic volatility model. *Journal of Computational Finance* **11**(3), 1–42. doi:10.21314/JCF.2008.189.

Andersen L., Pykhtin M., and Sokol A. (2016). Credit exposure in the presence of initial margin URL https://ssrn.com/abstract=2806156.

Andersen L., Pykhtin M., and Sokol A. (2017). Does initial margin eliminate counterparty risk? *Risk Magazine* **(May)**. URL https://www.risk.net/risk-management/5209111/does-initial-margin-eliminate-counterparty-risk.

Anfuso F., Aziz D., Giltinan P., and Loukopoulos K. (2017). A sound modelling and backtesting framework for forecasting initial margin requirements URL https://ssrn.com/abstract=2716279.

Antonov A. (2017). Algorithmic differentiation for callable exotics URL https://ssrn.com/abstract=2839362.

Antonov A., Issakov S., Konikov M., McClelland A., and Mechkov S. (2017a). PV and XVA Greeks for callable exotics by algorithmic differentiation URL https://ssrn.com/abstract=2881992.

Antonov A., Issakov S., and McClelland A. (2017b). Efficient SIMM-MVA calculations for callable exotics URL https://papers.ssrn.com/sol3/abstract=3040061.

Bank of International Settlements (2017). Global OTC derivatives market semi-annual statistics. Technical report, BIS. URL http://stats.bis.org/statx/srs/table/d5.1.

Basel Committee on Banking Supervision (2015). Margin requirements for non-centrally cleared derivatives. Technical report, BCBS. URL https://www.bis.org/bcbs/publ/d317.html.

Brigo D. and Mercurio F. (2007). *Interest Rate Models – Theory and Practice: With Smile, Inflation and Credit*. Springer Science & Business Media. doi:10.1007/978-3-540-34604-3.

Caspers P., Giltinan P., Lichters R., and Nowaczyk N. (2017). Forecasting initial margin requirements: A model evaluation. *Journal of Risk Management in Financial Institutions* **10**(4), 365–394. doi:10.2139/ssrn.2911167.

Caspers P. and Lichters R. (2018). Initial margin forecast – Bermudan swaption methodology and case study URL https://ssrn.com/abstract=3132008.

Fries C. (2017a). Automatic backward differentiation for American Monte-Carlo algorithms (conditional expectation) URL https://ssrn.com/abstract=3000822.

Fries C. (2017b). Fast stochastic forward sensitivities in Monte-Carlo simulations using stochastic automatic differentiation (with applications to initial margin valuation adjustments (MVA)) URL https://ssrn.com/abstract=3018165.

Fries C. (2017c). Stochastic automatic differentiation: Automatic differentiation for Monte-Carlo simulations URL https://ssrn.com/abstract=2995695.

George F. and Ramachandran K. (2011). Estimation of parameters of Johnson's system of distributions. *Journal of Modern Applied Statistical Methods* **10**(2), 9. doi:10.22237/jmasm/1320120480.

Heston S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* **6**(2), 327–343. doi:10.1093/rfs/6.2.327.

Hill I., Hill R., and Holder R. (1976). Algorithm AS 99: Fitting Johnson curves by moments. *Journal of the Royal Statistical Society. Series C (Applied statistics)* **25**(2), 180–189. doi:10.2307/2346692.

Hull J. and White A. (1990). Pricing interest-rate-derivative securities. *The Review of Financial Studies* **3**(4), 573–592. doi:10.1093/rfs/3.4.573.

Hull J. and White A. (1994). Branching out. *Risk* **7**(7), 34–37.

International Swaps and Derivatives Association (2016). ISDA SIMM (TM) methodology version R1.0. Technical report, ISDA. URL https://www.isda.org/a/xwEDE/isda-simm-vr1-0-public.pdf.

Johnson N. (1949). Systems of frequency curves generated by methods of translation. *Biometrika* **36**(1/2), 149–176. doi:10.1093/biomet/36.1-2.149.

Jones D. (2014). Johnson curve toolbox for Matlab: Analysis of non-normal data using the Johnson family of distributions. URL https://www.usf.edu/marine-science/research/matlab-resources/johnson-system-of-distributions-johnson-curves.aspx.

Klenke A. (2013). *Probability Theory: A Comprehensive Course*. Springer Science & Business Media. doi:10.1007/978-1-4471-5361-1.

Lewis A. (2000). *Option Valuation under Stochastic Volatility with Mathematica Code*. Finance Press: Newport Beach, CA. ISBN 0-9676372-0-1.

Nazneen S. (2017). Capital savings from new IM regime elude dealers. *Risk Magazine* **(Aug)**. URL https://www.risk.net/derivatives/5315426/capital-savings-from-new-im-regime-elude-dealers.

O'Halloran S., Nowaczyk N., and Gallagher D. (2017). Big data and graph theoretic models: Simulating the impact of collateralization on a financial system. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, 1056–1064. ACM. doi:10.1145/3110025.3120989.

O'Halloran S., Nowaczyk N., and Gallagher D.and Subramaniam V. (2019). *A Data Science Approach to Predict the Impact of Collateralization on Systemic Risk*, 171–192. Springer International Publishing. doi:10.1007/978-3-030-11286-8_8.

Ostrovski V. (2013). Efficient and exact simulation of the Hull-White model URL https://papers.ssrn.com/abstract=2304848.

Rouah F.D. (2013). *The Heston Model and Its Extensions in Matlab and C#*. John Wiley & Sons. doi:10.1002/9781118656471.

Storer R.H. (1987). Adaptive estimation by maximum likelihood: Fitting of Johnson distributions. Technical report, Unpublished Ph.D. thesis, School of Industrial and Systems Engineering, Georgia Institute of Technology.

Tuenter H. (2001). An algorithm to determine the parameters of SU-curves in the Johnson system of probability distributions by moment matching. *Journal of Statistical Computation and Simulation* **70**(4), 325–347. doi:10.1080/00949650108812126.

Wheeler R. (1980). Quantile estimators of Johnson curve parameters. *Biometrika* 725–728. doi:10.2307/2335153.

A Appendix

A.1 Conditional Distributions and Quantiles

While conditional expectations and their properties are widely known, conditional quantiles are a little less standard. We collect the relevant facts in this appendix for convenient reference. Recall that if $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space and $\mathscr{G} \subseteq \mathscr{F}$ is a σ-sub-algebra, the *conditional expectation* $\mathbb{E}[X \mid \mathscr{G}]$ of any real random variable X given $\mathscr{G}$ is defined uniquely $\mathbb{P}$-almost surely by the equation $\mathbb{E}[\mathbb{E}[X \mid \mathscr{G}]1_A] = \mathbb{E}[X1_A]$ for all $A \in \mathscr{G}$. For any $B \in \mathscr{F}$, the *conditional probability of B given $\mathscr{G}$* is then defined as $\mathbb{P}[B \mid \mathscr{G}] := \mathbb{E}[1_B \mid \mathscr{G}]$. By the regular version of the conditional distribution, see for example (Klenke 2013, Thm. 8.63), there exists a function $\kappa = \kappa_{X|\mathscr{G}} : \Omega \times \mathscr{B}_{\mathbb{R}} \to [0,1]$ such that for each $B \in \mathscr{B}_{\mathbb{R}}$, the function $\Omega \to [0,1]$, $\omega \mapsto \kappa(\omega, B)$, is $\mathscr{F}$-measurable and for each $\omega \in \Omega$, the function $\mathscr{B}_{\mathbb{R}} \to [0,1]$, $B \mapsto \kappa(\omega, B)$, is a probability measure on $(\mathbb{R}, \mathscr{B}_{\mathbb{R}})$ (here $\mathscr{B}_{\mathbb{R}}$ denotes the Borel algebra). For any $\alpha \in]0,1[$, the function

$$Q_\alpha(X \mid \mathscr{G}) : \Omega \to \mathbb{R}, \qquad \omega \mapsto Q_\alpha(\kappa_{X|\mathscr{G}}(\omega, \_)),$$

where Q_α denotes the α-quantile of a distribution, is called the *conditional α-quantile of X given $\mathscr{G}$*.

A.2 Regression Procedure

In performing the least-squares optimization, we shall assume the functional form

$$f(\boldsymbol{\beta}, x) = \sum_{r=0}^{R} \beta_r \phi_r(x)$$

in terms of the basis functions $\phi_r(x)$, $0 \le r \le R$, and $\boldsymbol{\beta} = [\beta_0, \beta_1, \ldots, \beta_R]^T \in \mathbb{R}^{R+1}$. There are many choices applicable for the basis functions including Hermite, Legendre, Chebyshev, Gegenbauer and Jacobi polynomials. We choose to use the Laguerre polynomials, which are

defined as

$$\begin{aligned}
\phi_0(x) &:= 1, \\
\phi_1(x) &:= 1 - x, \\
\phi_2(x) &:= 1 - 2x + x^2/2, \\
&\vdots \\
\phi_r(x) &:= \frac{e^x}{r!}\frac{d^r}{dx^r}(x^r e^{-x}).
\end{aligned}$$

In terms of the observations $x_i \sim X$ and $y_i \sim Y$, for $1 \leq i \leq n$, we wish to minimize the mean square error

$$\mathbb{E}\left[\left(Y - \sum_{r=0}^{R} \beta_r \phi_r(X)\right)^2\right] \approx \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \sum_{r=0}^{R} \beta_r \phi_r(x_i)\right)^2$$

with respect to the coefficients β_r. Thus, differentiating this expression with respect to β_s and equating to zero we have

$$\sum_{i=1}^{n} \phi_s(x_i) \sum_{r=0}^{R} \phi_r(x_i)\beta_r = \sum_{i=1}^{n} \phi_s(x_i) y_i,$$

for $0 \leq s \leq R$. This solution may be written in matrix form as

$$\boldsymbol{\beta} = (\boldsymbol{F}\boldsymbol{F}^T)^{-1}\boldsymbol{F}\boldsymbol{Y} \quad \text{with} \quad f(\boldsymbol{\beta}, \boldsymbol{X}) = \boldsymbol{F}^T\boldsymbol{\beta},$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ are written as column vectors of observations x_i and y_i, and $\boldsymbol{F}$ is the matrix

$$\boldsymbol{F} = \begin{bmatrix} \phi_0(x_1) & \phi_0(x_2) & \cdots & \phi_0(x_n) \\ \phi_1(x_1) & \phi_1(x_2) & \cdots & \phi_1(x_n) \\ \vdots & \vdots & & \vdots \\ \phi_R(x_1) & \phi_R(x_2) & \cdots & \phi_R(x_n) \end{bmatrix}.$$

A.3 The Johnson Curves

Johnson (1949) proposes a set of translations that transform a continuous random variable X, with an unknown distribution suggested either by empirical data or generated realizations from a Monte Carlo simulation, into a standard normal random variable Z. These translations are known as Johnson translations and have the general form

$$Z = a + bJ\left(\frac{X - c}{d}\right), \tag{10}$$

where a and b are shape parameters, c is a location parameter, d is a scale parameter and

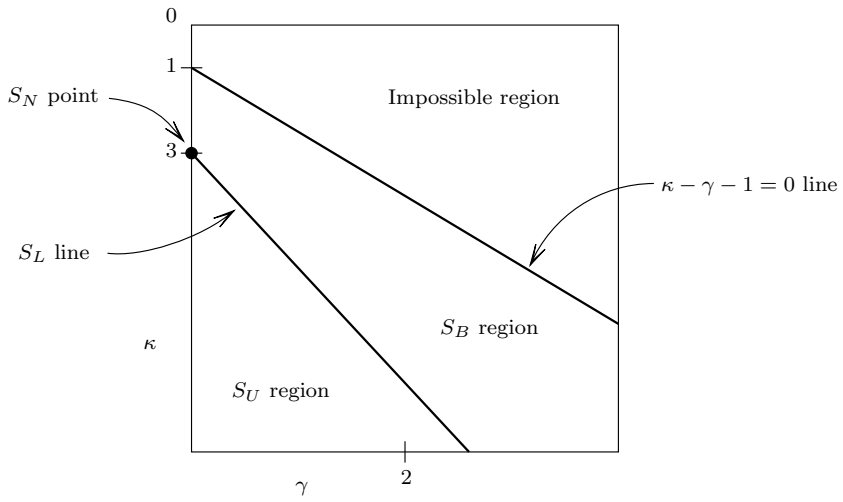

Figure 8: Identification of Johnson subsystem illustrating the method from Johnson (1949).

$J(\cdot)$ is a function defining the families of distributions in the Johnson system:

$$J(u) = \begin{cases} \ln(u) & S_L\text{: log-normal family} \\ \ln(u + \sqrt{u^2 + 1}) & S_U\text{: unbounded family} \\ \ln(u/(1-u)) & S_B\text{: bounded family} \\ u & S_N\text{: normal family.} \end{cases} \tag{11}$$

Given the random variable X, the first four central moments are defined as

$$\begin{aligned} \mu_1 &:= \mathbb{E}[X] = \mu_1', \\ \mu_2 &:= \mathbb{E}\left[X - \mu_1\right]^2 = \mu_2' - \mu_1^2, \\ \mu_3 &:= \mathbb{E}\left[X - \mu_1\right]^3 = \mu_3' - 3\mu_1\mu_2' + 2\mu_1^3, \\ \mu_4 &:= \mathbb{E}\left[X - \mu_1\right]^4 = \mu_4' - 4\mu_1\mu_3' + 6\mu_1^2\mu_2' - 3\mu_1^4, \end{aligned} \tag{12}$$

where $\mu_i' = \mathbb{E}\, X^i$, for $i \in \{1, 2, 3, 4\}$, are the first four raw moments. Then, the skewness and kurtosis of X are defined by

$$\gamma := \frac{\mu_3}{\mu_2^{3/2}} \qquad \text{and} \qquad \kappa := \frac{\mu_4}{\mu_2^2}. \tag{13}$$

Johnson (1949) showed that the system of curves described by (11) can accommodate all attainable points on the (γ, κ) plane, see Figure 8. The impossible region represents all combinations of skewness and kurtosis that cannot occur theoretically, i.e., it is always true that $\kappa \geq \gamma^2 + 1$.

The S_N family is represented by the point $(0, 3)$. Recall that the skewness and kurtosis of the normal distribution are zero and three respectively. The S_L family is represented by a curve—for the definition of this curve we refer to (Johnson 1949, p. 155). The S_B family is represented by the region between the S_L line and the limiting line $\kappa - \gamma - 1 = 0$, beyond

which is the unattainable area. The remainder of the (γ, κ) plane corresponds to the S_U family. Therefore, the skewness and kurtosis uniquely identify the appropriate form of the Johnson translation function $J(\cdot)$.

Once $J(\cdot)$ is identified, the distribution can be approximated by finding values for the parameters a, b, c and d that match the theoretical distribution to the sample distribution. In the literature there are three distinct approaches used for finding these parameters:

1. Moment matching, see e.g. Hill et al. (1976);
2. Matching of quantiles, see e.g. Wheeler (1980); and
3. Maximum likelihood estimation, see e.g. George and Ramachandran (2011); Storer (1987).

For the numerical examples provided in the next section, we used the Matlab toolbox described by Jones (2014). The toolbox is based on the moment matching approach in Hill et al. (1976), which, given the first four moments of a sample distribution, provides an algorithm to find the appropriate $J(\cdot)$ function and corresponding parameters. A brief summary of the algorithm for the moment matching approach is as follows:

Moment Matching Algorithm

1. Compute the sample analogues of Equations (12) and (13), i.e., $\hat{\mu}_1$, $\hat{\mu}_2$, $\hat{\mu}_3$, $\hat{\mu}_4$, $\hat{\gamma}$ and $\hat{\kappa}$.
2. Identify $J(\cdot)$ using $(\hat{\gamma}, \hat{\kappa})$.
3. Find a, b, c and d by equating the first four sample moments to the corresponding theoretical moments of the chosen Johnson distribution.

This algorithm is efficient but does not guarantee convergence as outlined in Hill et al. (1976). However, failure to converge occurs only in very few cases as observed in our numerical experiments. Alternative robust approaches are available in Wheeler (1980); Tuenter (2001); George and Ramachandran (2011); Storer (1987).

The coefficients defining a Johnson distribution consist of two shape (a, b), a location (c), and a scale (d) parameter. This allows a unique distribution to be derived for any combination of mean, standard deviation, skewness, and kurtosis that occurs for a given set of observed data. Once a variate is appropriately transformed, probability densities and percentage points may be derived based on the standard normal curve. To illustrate the behaviour, we have produced examples of the different shapes of the distributions by varying the shape parameters (a, b) in Figures 9 and 10.

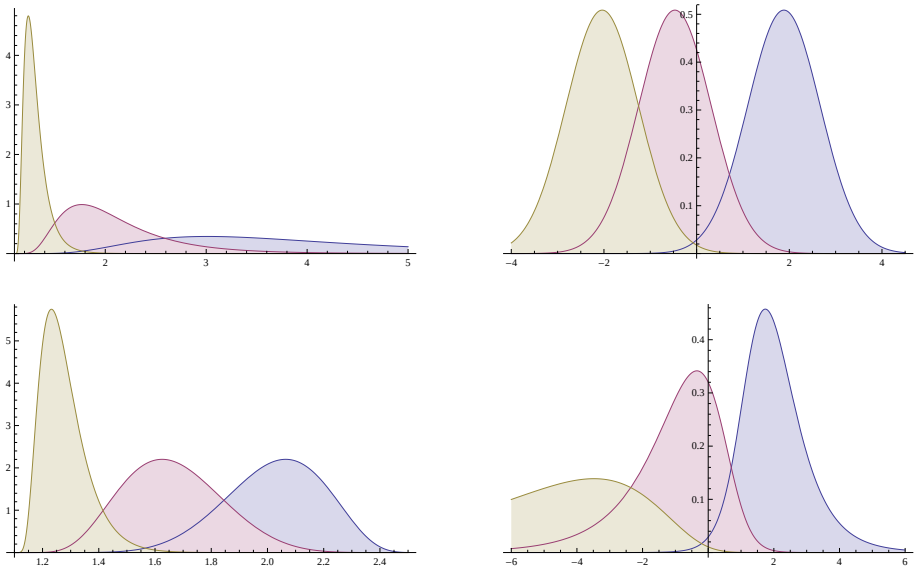

Figure 9: Different shapes of the Johnson distribution for the shape parameter a, for S_L (top left), S_N (top right), S_B (bottom left) and S_U (bottom right).

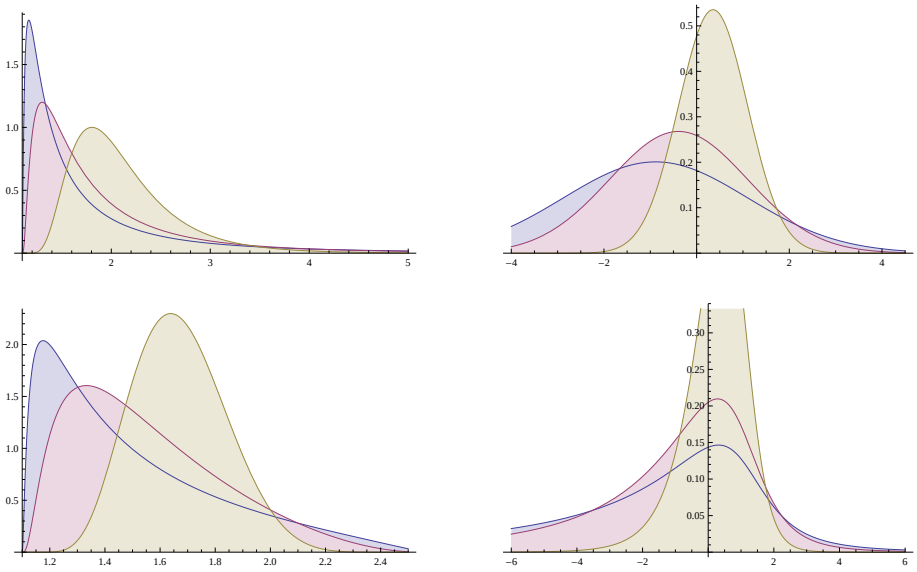

Figure 10: Different shapes of the Johnson distribution for the shape parameter b, for S_L (top left), S_N (top right), S_B (bottom left) and S_U (bottom right).