Wiping the smile off your base (correlation curve)

Ed Parcell, Derivative Fitch*

James Wood, Derivative Fitch

21 June 2007

Abstract

We discuss problems with interpolating and extrapolating base correlation curves and examine the pricing of CDO tranches with non-standard subordination levels. We introduce an alternative risk measure, the expected loss of equity tranches. We calculate upper and lower boundaries on "base EL", and set out the behaviours that base EL must obey. We investigate interpolation schemes that best avoid model arbitrage. We also look at the calculated prices and sensitivities for tranches using these different methods.

This paper is intended to be a practical explanation of methodologies used by Derivative Fitch in its RAP CD synthetic CDO pricing tool. For our introduction to the underlying concept of looking at the expected loss of equity tranches, we are grateful for discussions with a number of market participants, notably Jon Gregory.

1 Introduction

In 2003, Andersen, Sidenius and Basu[2] introduced the 1-factor model for calculating prices and sensitivities of synthetic CDO tranches. This allowed a move away from time-consuming Monte Carlo methods, which in turn caused correlation to be viewed as an implied parameter, rather than an exogenously set parameter.

The initial approach, "tranche correlation", implied a single correlation parameter for each tranche in the iTraxx and CDX indices. One problem with the use of tranche correlation is that it is not obvious how to interpolate between tranche correlations - given correlations for

*The views, thoughts and opinions expressed in this paper are those of the authors in their individual capacity. The authors would gratefully receive any comments on this material. They can be reached at ed.parcell@derivativefitch.com and james.wood@derivativefitch.com. James Wood asserts that any errors in this paper are Ed Parcell's. Ed Parcell asserts the inverse.

3%-6% and 6%-9% tranches, it is hard to justify what correlation should be used to price a 5.4%-7.0% tranche, say. Worse, during the May 2005 correlation crisis, correlations for some mezzanine tranches ceased to exist - no correlation parameter would give correct market prices.

Base correlation, introduced by McGinty et al[6], implies correlations for various base (i.e. equity) tranches, pricing mezzanine and senior tranches as the difference of two base tranches. Correlations at standard strikes (e.g. 3%, 6%, 9%, 12%, 22% on iTraxx Europe[1]) can be found as a bootstrap. The approach makes pricing non-standard tranches easier. Using the example above, the 5.4% and 7.0% base correlations can be interpolated from the 3%, 6% and 9% base correlations implied from market prices.

However, some care is needed with this interpolation. The price of a tranche is very sensitive to the slope of the base correlation curve. In some interpolation schemes, the slope can change rapidly around control points. So even though we know base correlation at 6%, the relative value of a 5.9%-6.0% tranche compared to a 6.0%-6.1% tranche will be very sensitive to the interpolation method chosen, and may even generate arbitrages in the model.

Extrapolation outside the 3%-22% range on the iTraxx index or the 3%-30% range on the CDX index is even more dangerous. Different extrapolation methods can give very different prices for junior tranches, and it is easy to accidentally generate arbitrages in senior tranches.

Some researchers have looked at various skew models. Examples include the NIG model[4], random factor loading[1] and stochastic correlation models[3]. Skew models are designed to attempt to allow a single set of inputs that fit market prices at all strikes. Typically these parameterizations are multi-dimensional, and, unlike base correlations, it is not possible to solve for them via a bootstrap. They do not necessarily fit well to market prices, and often it is only possible to fit to a subset of the 5 available prices - the extra degrees of freedom do not greatly increase the set of attainable tranche prices. These models are arbitrage-free in the sense that more senior tranches are guaranteed to have lower risk. Their downside is that there is not one unique distribution that best fits the market, so day-to-day consistency of pricing remains a problem.

A separate question is how to use base correlations implied for indices to value bespoke asset pools. This is a methodology that lies outside the scope of a model. Commonly used techniques include expected loss mapping and spread matching. We do not deal with it further in this paper.

[1]The two best known credit indices have different tranches – 0%-3%, 3%-6%, 6%-9%, 9%-12%, 12%-22% on the iTraxx, and 0%-3%, 3%-7%, 7%-10%, 10%-15%, 15%-30% on the CDX NA IG. Rather than clumsily attempt to refer to both, we shall use only the iTraxx Europe strikes in this paper, and understand that the same analysis applies to the CDX NA IG index.

The remainder of this paper is structured as follows: In section 2 we briefly describe the 1-factor copula, and describe the relation between base correlation and expected loss of base tranches. In section 3 we give results for various methods of interpolating/extrapolating base correlation, and their advantages and disadvantages. In section 4 we outline the base expected loss approach, and give results for various methods of interpolating the base expected loss curve. In section 5 we look at the effect on price and sensitivities of different methods. Section 6 is a conclusion. In Appendix A we give details of our shape-preserving interpolation method.

2 The 1-factor Gaussian copula

We use the standard formulation of the 1-factor Gaussian copula. In this model, the default of each obligor is conditionally independent given a level of state risk, M, with default occurring if the latent variable

$$X_i = \sqrt{\rho} M + \sqrt{1-\rho}\epsilon_i$$

is less than some threshold, constructed to fit market CDS prices. M and each of the ϵ_i are taken to be independent normally distributed random variates.

Andersen, Sidenius and Basu[2] and Gregory and Laurent[5] give detailed descriptions of calculation of thresholds, cashflows and PVs. For completeness, and to fix notation, we summarize some of the values that can be obtained from the model. The model allows calculation of conditional, independent default probability for each obligor i at time t, $q_i(t|M)$. These values are used to calculate the probability distribution of loss (as a fraction of the pool notional) occurring at time t, given state risk:

$$\mathbb{P}\left(\text{Loss}(t) = \frac{l \times LU}{PN} \middle| M\right) \tag{1}$$

where LU is a loss unit chosen so that potential losses can reasonably be represented as integer numbers of loss units of loss, and PN is the pool notional, the sum of notionals of the underlying obligors. Parcell[7] gives details of relaxing the restriction on losses on default so that deals with inhomogenous notional and recovery rate can be priced without losses in model speed.

2.1 Valuing CDO tranches

The loss distribution $\mathbb{P}(\text{Loss}(t) = \bullet|M)$ allows calculation of the expected loss of a 0-x base tranche at any time t:

$$\mathbb{E}[\min(\text{Loss}(t), x)] = \int_{-\infty}^{\infty} \phi(m) \sum_i \mathbb{P}\left(\text{Loss}(t) = \frac{i \times LU}{PN} \middle| M = m\right) \min\left[\frac{i \times LU}{PN}, x\right] dm \quad (2)$$

By summing over the coupon dates $t_1, t_2, \ldots, t_M$, we can calculate the expected (discounted) loss paid by the protection seller on an a-b tranche having correlation ρ, $EL(a, b; \rho)$ and the value of a premium of 1 basis point paid by the protection buyer, $Prem01(a, b; \rho)$, and hence the present value of buying protection on an a-b tranche, $PV(a, b)$:

$$EL(0, x; \rho_x) = \sum_j (E[\min(\text{Loss}(t_{j+1}), x)] - E[\min(\text{Loss}(t_j), x)]) \times \text{df}\left(\frac{t_j + t_{j+1}}{2}\right) \quad (3)$$

$$Prem01(0, x; \rho_x) = 0.0001 \times \sum_j (x - E[\min(\text{Loss}(t_j), x)]) \times (t_j - t_{j-1}) \times \text{df}(t_j) \quad (4)$$

$$\begin{aligned} PV(a, b) &= EL(0, b; \rho_b) - EL(0, a; \rho_a) \\ &\quad - PREM \times [Prem01(0, b; \rho_b) - Prem01(0, a; \rho_a)] - \text{Upfront} \end{aligned} \quad (5)$$

where $PREM$ is the premium in basis points paid by the protection buyer and $Upfront$ is the value of any payment made at the start of the deal.

2.2 Boot-strapping base correlations

One of the attractions of the base correlation approach is that we can boot-strap base correlations. By that we mean that from the market we can calibrate a base correlation for each strike in turn, rather than have to solve for all of them simultaneously using a multi-dimensional solver as in a skew model. We outline the calibration process here.

We look at the 0%-3% tranche first. We know that the PV when entering a fair swap is 0. Also, we know that the expected discounted loss and premium leg of the 0%-0% base tranche is 0. So equation (5) reduces to:

$$0 = EL(0, 3\%; \rho_{3\%}) - PREM \times Prem01(0, 3\%; \rho_{3\%}) - \text{Upfront} \quad (6)$$

The premium and upfront can be observed in the market, so the only unknown in this equation is $\rho_{3\%}$, and we can solve for it numerically, using a 1-dimensional solver.

Having solved for $\rho_{3\%}$, we can look at the 3%-6% tranche. Again, we know that the PV of the fair swap in the market is 0. So equation (5) gives:

$$\begin{aligned} 0 &= EL(0, 6\%; \rho_{6\%}) - EL(0, 3\%; \rho_{3\%}) \\ &\quad - PREM \times [Prem01(0, 6\%; \rho_{6\%}) - Prem01(0, 3\%; \rho_{3\%})] - \text{Upfront} \end{aligned} \tag{7}$$

We already know $\rho_{3\%}$ from solving equation (6), and we know the market premium and upfront. Therefore the only unknown in this equation is $\rho_{6\%}$, and we can solve for it numerically.

We continue like this up the capital structure, solving for $\rho_{9\%}$, $\rho_{12\%}$ and $\rho_{22\%}$. At the end of this process we have a set of base correlations calibrated to the market tranche prices of an index pool at a particular maturity. Shelton[9] gives details on extending this to make correlation a function of time, to fit multiple maturities simultaneously.

3 Interpolating the base correlation curve

We have used boot-strapping to solve for the base correlation at 5 strikes. We now turn our attention to pricing tranches with arbitrary attachment point, a, and detachment point, b. The market appears to imply that implied correlation for equity tranches increases smoothly with the detachment point of the tranche. This suggests that a sensible way to proceed is to interpolate or extrapolate from our 5 known points to determine ρ_a and ρ_b. We look at the effect of using different methods for doing this. We use this approach in our model.

3.1 Linear interpolation

The simplest method of interpolation is to use linear interpolation to determine ρ_x on tranches between 3% and 22%. To determine ρ_x outside this region, there are many potential choices of extrapolation strategy. Here are a few:

- Continue the gradient of the last line segment i.e. linearly extrapolate from the last two points
- Set the correlation at the 0% strike to be 0%, and the 100% strike to be 100%.
- Set the correlations at 0% and 100% to be some other value.

Each extrapolation strategy has difficulties. It is hard to defend the arbitrary choice of correlations at 0% and 100%. On the other hand, continuing the gradient of the last line segment can give correlations below 0% for junior tranches or above 100% for senior tranches. Capping and flooring ρ_x to keep it in the range $[0\%, 100\%]$ introduces a sudden change in the gradient of ρ_x at an arbitrary point.

Here we illustrate the first strategy – linearly extrapolating from the last two points. Formally, given ρ at a set of strikes $x_0, x_1, \ldots, x_k$, we determine ρ at an arbitrary point using:

$$\rho_x = \begin{cases} \rho_{x_0} + \frac{x-x_0}{x_1-x_0}(\rho_{x_1} - \rho_{x_0}) & \text{if } 0 \leq x < x_0 \\ \rho_{x_i} + \frac{x-x_i}{x_{i+1}-x_i}(\rho_{x_{i+1}} - \rho_{x_i}) & \text{if } x_i \leq x < x_{i+1} \\ \rho_{x_{k-1}} + \frac{x-x_{k-1}}{x_k-x_{k-1}}(\rho_{x_k} - \rho_{x_{k-1}}) & \text{if } x \geq x_k \end{cases} \tag{8}$$

This gives us the base correlation curve shown in figure 1. This curve, and all examples in this paper, were calculated using data for the iTraxx 5 year index on 8 May 2007.

Figure 2 shows the fair spreads of thin tranches at various points in the capital structure implied by the linearly interpolated/extrapolated base correlation curve. Two indications of arbitrage can be seen.

First, at 9%, and again at 12%, the fair spread increases. That is, tranches with greater subordination are considered riskier by this model. Clearly this is inconsistent with the fact that any increase in subordination decreases risk. The reason that this occurs is that pricing is very sensitive to the gradient of the base correlation curve. The sudden changes in gradient given by linear interpolation around control points such as 9% and 12% gives a sudden change in the price, or fair spread, of tranchlets. Where the gradient change causes an increase in fair

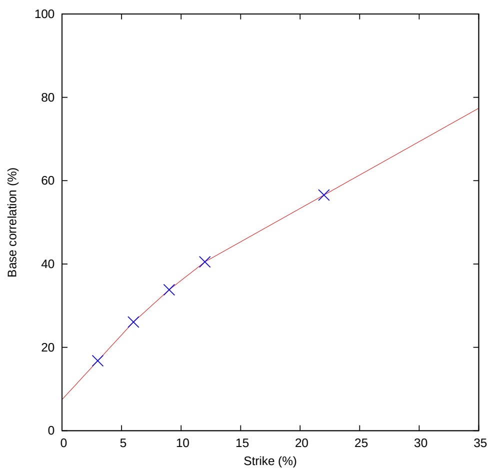

Figure 1: Linearly interpolated/extrapolated base correlation curve

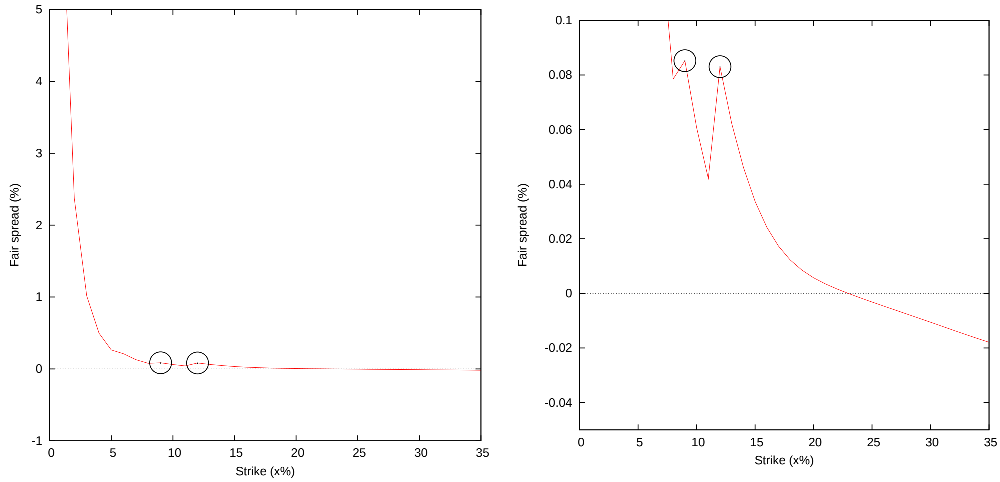

Figure 2: Fair spreads for tranchlets with width 0.5% at various points on the capital structure.

spread larger than the decrease due to increasing subordination, there will be arbitrages in the model. This is not a trading opportunity; the model is simply inconsistent with itself between strikes. We can reduce this arbitrage by using an interpolation method that is smoother around control points and has equal gradient on either side of them.

Second, above 23% the fair spread becomes negative. This is clearly impossible, and indicates an arbitrage in the model. Again, the reason this happens is that the change in base correlation is moving prices more than the increase in subordination is; the base correlation curve is too steep above 22%. This can be fixed by choosing a different extrapolation method for base correlation, but we must still be careful, because when extrapolating base correlation it is not clear whether it will result in such arbitrages.

Linear extrapolation is very easy to set up. On the downside it has serious flaws: mispricing very senior tranches, odd relative prices of some mezzanine tranchlets, and possible problems with invalid correlations. A better method is obviously needed.

3.2 Cubic spline interpolation

The requirement for smooth interpolation at control points suggests using a spline for interpolation.

A cubic spline is a piecewise cubic, constructed to fit a set of control points, so that it is continuous and has continuous first and second derivatives everywhere. It has two degrees of freedom – the second derivatives at each end point. Setting both of these equal to zero gives the natural spline. For more information on splines, we refer to Numerical Recipes in C++[8].

Outside the 3%-22% region, it is still not clear what base correlations should be used. Here we show the effects of extrapolating linearly, using the gradients of the spline at 3% and 22%. This keeps the derivative of base correlation smooth, avoiding some arbitrage problems, but we still have the possibility of moving outside the 0-100% range for correlation.

Figure 3 shows the base correlation curve and fair spreads using this interpolation method. We note that in this example (which has fairly large tranchlets of width 0.5%), we avoid the arbitrages of having tranchlets with negative value, or more value than more junior tranchlets. However, we cannot guarantee this will always be the case without actually pricing all tranchlets and detecting this problem.

Spline interpolation is about as far as it is possible to go with interpolation on base correlation curves. There are undoubtedly fancier methods of interpolating, but the core problem remains that it is not possible to distinguish an arbitrage-free base correlation curve from an arbitragable one without checking every possible tranchlet price. This is difficult even for a single set of index prices on a single day, let alone to show for a particular interpolation method in the general case. For that, we need to look at interpolating on a fundamentally different risk-measure than base correlation.

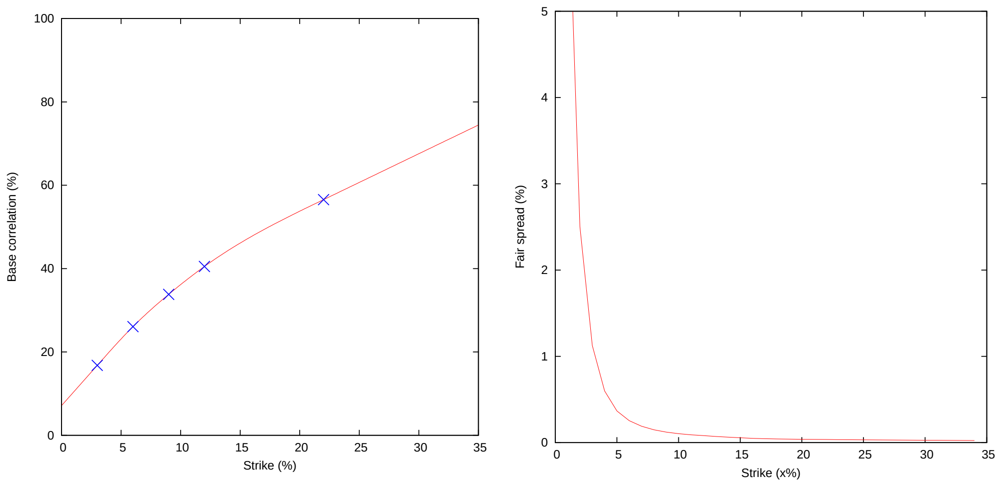

Figure 3: Left: Base correlation curve using a cubic spline for interpolation, and linear extrapolation, based on the gradient of the spline at 3% and 22% Right: Fair spreads for tranchlets with width 0.5% at various points on the capital structure, using this base correlation curve.

4 Base expected loss

Base correlation is preferable to tranche correlation because it allows interpolation and extrapolation from market-implied correlations and thus allows CDOs with non-standard subordinations to be priced. The reason for this is that the base correlation approach only deals with correlations of base tranches, making ρ a function of only one variable. By contrast, tranche correlation requires ρ to be considered as a function of both the attachment point and detachment point of a tranche. In either approch, interpolation and extrapolation have to be performed from only 5 data points - on the iTraxx $\rho_{3\%}$, $\rho_{6\%}$, $\rho_{9\%}$, $\rho_{12\%}$ and $\rho_{22\%}$ for base correlation and $\rho_{0\%,3\%}$, $\rho_{3\%,6\%}$, $\rho_{6\%,9\%}$, $\rho_{9\%,12\%}$ and $\rho_{12\%,22\%}$ for tranche correlation. Clearly, given such sparse data it is easier to use base correlation and interpolate a function of one variable than a function of two.

However, base correlation still has problems. The base correlation curve is not very intuitive. Worse, it is easy to implement this approach in a way that permits model arbitrage. Given base correlations at 3% and 6%, say, it's not clear what base correlations are allowable at 4.5%. Setting base correlation for 4.5% too high would cause the expected loss of the 0%-4.5% tranche to be lower than that of the 0%-3% tranche, while too low a base correlation causes the 0%-4.5% tranche no have higher expected loss than the 0%-6% tranche. It is not clear whether a particular interpolation method gives a base correlation curve free of such problems at every point.

We would prefer to retain the "base" approach, dealing only with equity tranches, to allow us to construct a 'pricing curve' to value tranches with non-standard subordinations, but use a risk measure that makes it easy to keep our interpolation/extrapolation arbitrage-free. One possible measure is $EL(0, x; \rho_x)$, the expected discounted loss of each equity tranche, which we call the *base EL* curve. We can use boot-strapped base correlations at 3%, 6% etc. to calculate $EL(0, x; \rho_x)$ at each of these points.

As we will explain below, it is easy to tell whether a base EL curve is arbitrage-free, so by ensuring that our interpolation maintains certain key properties, we can avoid arbitrages.

4.1 What do we know about base expected loss?

We know that $EL(0, 0; \rho_0) = 0$ and $EL(0, 100\%; \rho_{100\%})$ is just the expected loss of the asset pool, which is determined by CDS spreads and recovery rates (and not correlation). This means that calculating $EL(0, 2\%; \rho_{2\%})$, for example, is an interpolation, rather than an extrapolation.

This should give us some extra assurance when pricing tranches with strikes below 3%.

Any base correlation curve implies a base EL curve. This relationship is one-to-one; different base correlation curves imply different base EL curves. However, not all valid expected loss surfaces have an associated base correlation curve - there are some arbitrage-free base EL curves that cannot be generated with the base correlation approach and some base correlation curves imply base EL curves with arbitrages.

We can perform our interpolation/extrapolation on either the base correlation curve, or on the expected loss curve $EL(0,\cdot)$. It is not required that our interpolation/extrapolation be expressable as a base correlation curve (although it might be helpful in some contexts).

Note that $EL(0,x;\rho_x)$ decreases as ρ_x increases - intuitively this is because correlation does not affect the expected loss on the pool, but increases the probability of extreme large losses which are truncated.

We can use the fact that $EL(0,x;\rho_x)$ can be calculated from the distribution of discounted loss. Considering discounted loss as a continuous quantity we have:

$$EL(0,x;\rho_x) = \int_0^x \mathbb{P}(\text{Loss} > t)dt \tag{9}$$

$$\therefore \quad \mathbb{P}(\text{Loss} > x) = \frac{\partial EL(0,x;\rho_x)}{\partial x} \tag{10}$$

$$\therefore \quad f_{\text{Loss}}(x) = -\frac{\partial^2 EL(0,x;\rho_x)}{\partial x^2} \tag{11}$$

Any probability has to be non-negative. The constraint that $\mathbb{P}(\text{Loss} > x) > 0$ gives us that the first derivative of $EL(0,x;\rho_x)$ must be positive i.e.the base EL curve must be monotonically increasing. Violating this constraint gives rise to negative probability densities on the loss distribution, and hence tranchlets with negative spreads.

Similarly, the constraint $f_{\text{Loss}}(x) > 0$ translates into the requirement that the second derivative of the base EL curve must be non-positive i.e. the base EL curve must be convex to avoid the fair spreads of tranchlets increasing with seniority.

4.1.1 Boundaries on interpolated base EL curves

Once we have calculated expected losses at each of the liquid strikes (and also at 0 and 100%), the fact that base EL curve have to be monotonic and convex can be used to produce constraints on expected loss interpolations. The bounds we give are not strict, and it is possible for an interpolation to lie within the bounds but not be arbitrage-free. However, violation of the bounds indicates that there is certainly a model arbitrage; tranche values outside these bounds are not possible.

4.1.2 Lower bound

There are a set of expected losses at various strikes which are increasing and convex, starting from 0% and going to 100%. These are our control points for interpolation.

Clearly the curve, $\mathcal{C}$, formed by straight lines joining each control point is also increasing and convex. We will demonstrate that this curve is also the "lowest" convex curve joining these points.

Consider some continuous interpolated base EL curve that is lower at some point than $\mathcal{C}$. That is, at some strike x, the expected loss on this curve is below the straight line joining two control points. Call the strikes of those control points a and b, with $a < b$. The curve segment joining x to the control point at a has lower average gradient than the curve segment joining x to the control point at b. No matter how the curve is shaped, somewhere between a and x it must have gradient not larger than the average gradient between a and x, say at strike A. Similarly, somewhere between x and b it must have gradient not smaller than the average gradient between x and b. We have that A is smaller than B, and the gradient at A is smaller than the gradient at B, which is not possible if the curve is convex. So any base EL curve that is anywhere lower than $\mathcal{C}$ is not convex, and hence that $\mathcal{C}$ is a lower bound for all base EL interpolations.

Simplistically, any curve that goes below $\mathcal{C}$ will have to bend upwards to hit the next control point, and so will not be convex.

4.1.3 Upper bound

To work out an upper bound, we have to combine several upper bounds. Our base EL curve is not allowed to violate any of them, so the actual upper bound at each strike is the lowest of the upper bounds given by these rules:

- The most obvious upper bound is that the expected loss of a tranche can never exceed the expected loss of the pool, so $EL(0, x; \rho_x) \leq EL(0, 100\%; \rho_{100\%})$ for all strikes x.
- Also,the loss on a tranche cannot exceed the width of the tranche, and so neither can the expected loss, giving $EL(0, x; \rho_x) \leq x$.
- Other bounds are more complex. Consider three control points, with strikes a, b and c. The gradient of the base EL curve must be at least the gradient of the straight line

joining b and c – if it were lower the base EL curve would fail to be convex in the same way is the previous section. The gradient of the base EL curve between a and b cannot be smaller than its gradient at b, so the smallest it can be is the gradient of the line joining b and c. This gives an upper bound composed by continuing the straight line segment joining each consecutive pair of control points from the more junior of those control points.

- Using similar reasoning, another upper bound is the continuation of the straight line segment joining each consecutive pair of control points for strikes above the more senior control point.

Unlike the lower bound, which is convex, the upper bound is not, and so is not a feasible base EL curve. However, it is the envelope of a set of convex functions.

4.1.4 Algebra

Given a set of boot-strapped base ELs, $(x_0 = 0, y_0 = 0)$, (x_0, y_0), (x_1, y_1), …, (x_{n-1}, y_{n-1}), $(x_n = 1, y_n)$, with $EL(0, x_i; \rho_{x_i}) = y_i$ for $i = 0, 1, \ldots, n$ that are monotonic and convex, the convexity requirement gives us constraints on the interpolated base EL curve:

$$EL(0, x; \rho_x) > y_j + (y_{j+1} - y_j)\frac{x - x_j}{x_{j+1} - x_j} \qquad \text{if } x_j < x < x_{j+1}, 0 \le j < n \qquad (12)$$

$$EL(0, x; \rho_x) < y_{j-1} + (y_j - y_{j-1})\frac{x - x_{j-1}}{x_j - x_{j-1}} \qquad \text{if } x_j < x < x_{j+1}, 1 \le j < n \qquad (13)$$

$$EL(0, x; \rho_x) < y_{j+1} + (y_{j+2} - y_{j+1})\frac{x - x_{j+1}}{x_{j+2} - x_{j+1}} \qquad \text{if } x_j < x < x_{j+1}, 0 \le j < n-1 \qquad (14)$$

4.1.5 Example

Figure 4 shows the bounds on base EL given by market prices on the iTraxx 5 year on 8 May 2007. Figure 5 shows the implied bounds on base correlation given by the base EL bounds. Note that a base correlation curve contained by these bounds may still fail to be arbitrage-free, for example if it has a very fast changing gradient at some points.

- The bounds on base correlation in the range 3%-22% are fairly tight. So any interpolation method that remains arbitrage-free should give similar prices on any tranche in this range.

- The bounds on base correlation or expected loss below the 3% strike are very wide. This means that there is a lot of uncertainty when pricing or risk-managing deals with strikes more junior than this. This could be alleviated by the formation of a liquid tranchlet market below 3%.

- At a strike of about 1.3% the lower bound on base correlation hits 0%. If a base correlation below 0% occurred, and we believe there is no theoretical reason why it could not, we might interpret this as meaning that it is possible that purely idiosyncratic defaults actually reduce the possibility of systematic defaults. A real life example of this might be the improvements in corporate governance after Enron's bankruptcy.
- The bounds on base correlation and expected loss above 22% are reasonably tight, although widening as we move away from 22%.
- The bounds on base correlation above 22% show that a linearly extrapolated base correlation is not feasible.

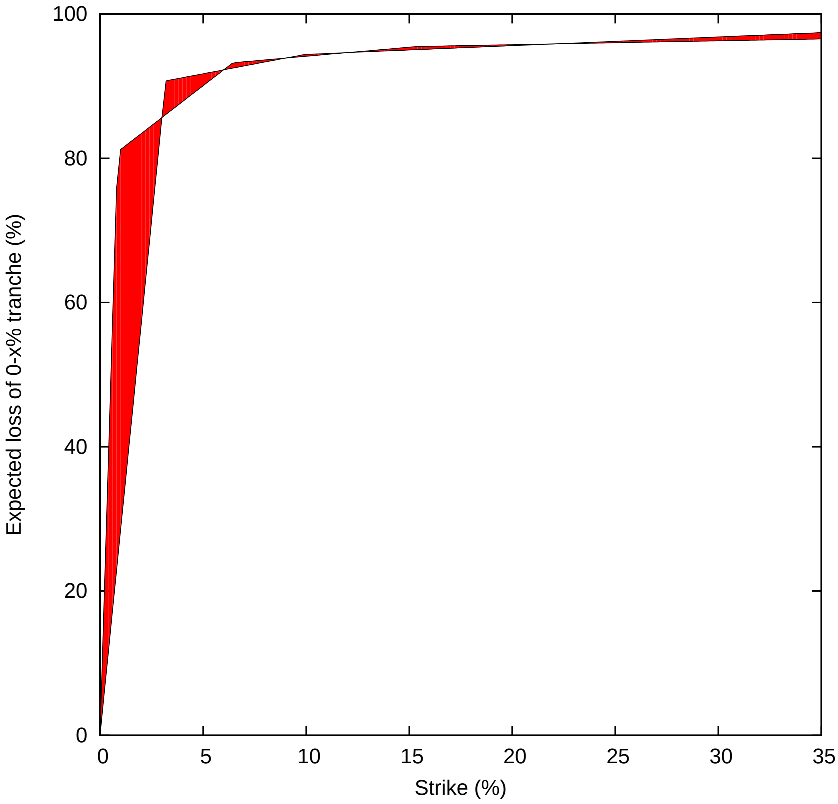

Figure 4: Bounds on base EL implied by monotonicity and convexity.

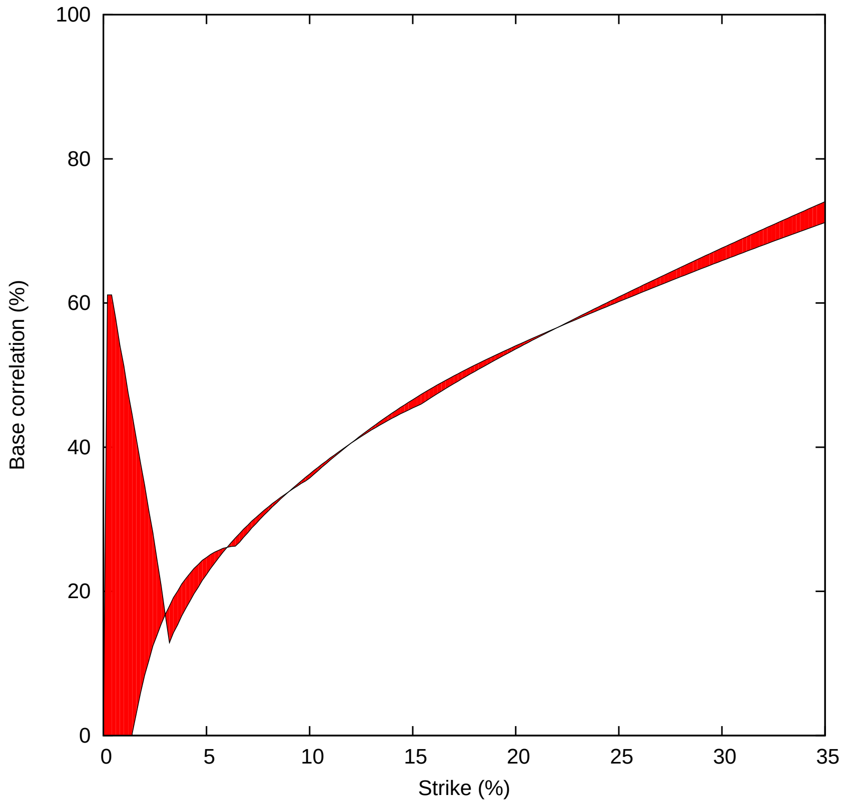

Figure 5: Base correlation bounds implied by bounds on base EL.

4.2 Interpolating the base expected loss curve

The linear interpolations on base EL above give boundaries for base correlation (see figure 5) and tranche prices. In practice, we would expect the base EL to be a smooth curve. The reason for this is that loss is (approximately) a continuous random variable, and hence so is discounted loss, so the cumulative density function never jumps, and equation (10) gives that the first derivative of the base EL curve is continuous. In addition, any value of loss has a non-zero probablity density, so from equation (11) we should expect the base EL curve to have decreasing gradient everywhere.

4.2.1 Spline interpolation

Since we would like to keep our interpolated base EL curve smooth, spline interpolation is an obvious choice. Figure 6 shows the interpolated base EL curve, and the base correlation curve implied by the base EL curve, using a natural spline.

It can immediately be seen that the base EL curve is oscillatory, and definitely is not

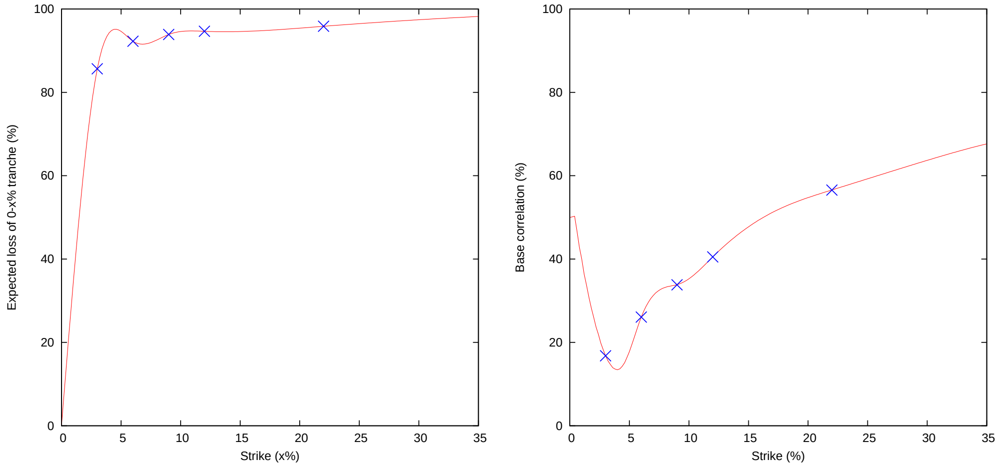

Figure 6: Left: Base ELinterpolated using spline interpolation. Right: Base correlation curve implied with this interpolation method.

monotonically increasing. It will generate significant arbitrages for a wide range of tranches.

It has been said that spline interpolation tends to work well, except when it does not. To be fair, splines do not guarantee to maintain the monotonicity or convexity of their control points, and here they do not as the stiffness that gives splines their accuracy forces the curve to oscillate to hit the control points.

4.2.2 Steffen's monotonic interpolation

Steffen[10] introduced an interpolation method that guarantees that monotonic sets of control points will generate a monotonic interpolating curve. The method produces smooth curves,

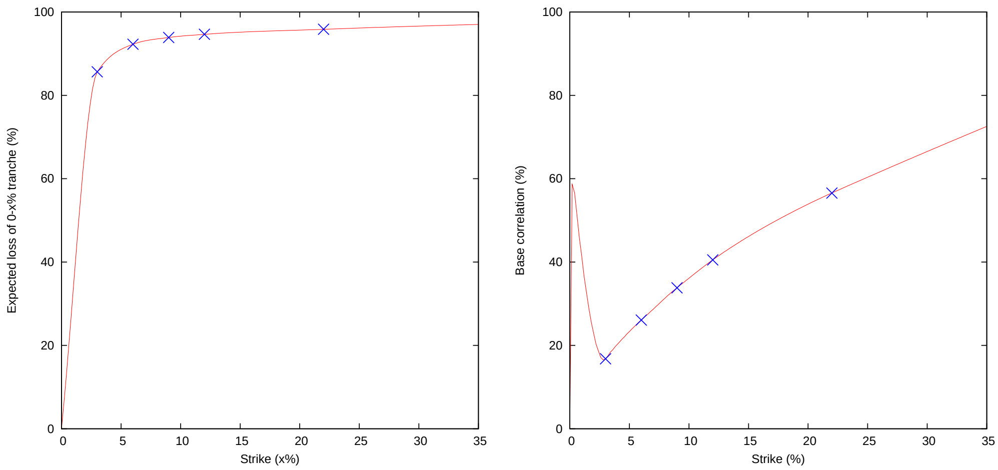

Figure 7: Left: Expected loss interpolated using Steffen interpolation. Right: Base correlation curve implied with this interpolation method.

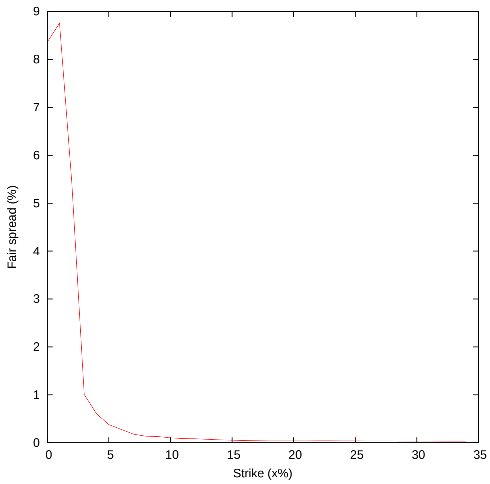

Figure 8: Fair spreads for tranchlets with width 0.5% at various points on the capital structure, using Steffen interpolation on the BaseEL curve.

and has better locality properties than cubic interpolation - small perturbations of control points tend not to introduce large changes in the interpolation function far away from that control point. In practice, we have found that the interpolation has a slightly lower order of accuracy than a cubic spline on various test-cases, but for many uses in financial modelling this is more than made up for by its preservation of monotonicity.

Figure 7 shows the interpolated base EL curve, and the base correlation curve implied by the base EL curve, using Steffen interpolation. Compared to using a spline for interpolation, the expected loss and base correlation curves are smooth, and avoid oscillation.

Figure 8 shows the fair spreads of thin tranches at various points in the capital structure implied by this interpolated base EL curve. The method works well over most strikes, however, there is an arbitrage between the fair spreads of the 0%-0.5% and 1%-1.5% tranches - in the model, the latter pays more, despite having extra subordination. The reason for this is that, despite its appearance in 7, the interpolated base EL curve is slightly concave near the 0% strike, as convexity in the control points is not preserved by the Steffen interpolation.

4.2.3 Piecewise quadratic interpolation

We have also looked at other schemes for interpolation to preserve monotonicity and convexity. We outline one here, giving further details in appendix A.

Our approach is to fit a piece-wise quadratic so it goes through each of the control points. We choose the gradient at the right extreme of the right-most piece to be half the gradient between the final two control points. This fixes the quadratic used for interpolation between

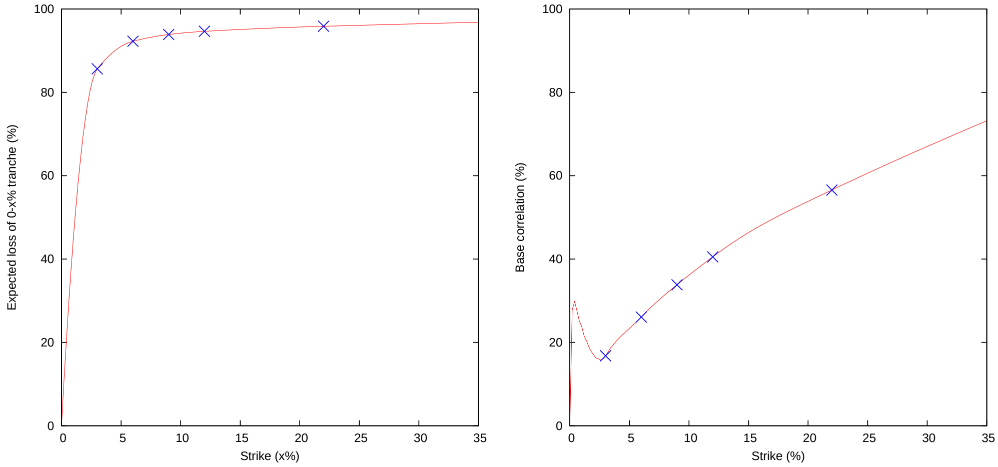

Figure 9: Left: Expected loss interpolated using piecewise quadratic interpolation. Right: Base correlation curve implied with this interpolation method.

the final two control points. We then choose the gradient of the right extreme of the second right-most piece to match the left extreme of the right-most piece, which fixes the quadratic for use between the penultimate two control points.

In our example we found that the curve produced was monotonic and convex, though it is not guaranteed to be. If the quadratic interpolation function is not monotonic or convex, this can be detected by the fact that the gradient on the left extreme between two control points would be smaller than the gradient of the line between those control points. In that case, we reject that segment, and the one to the right, replace the one to the right with a different quadratic, and re-start the algorithm. The first derivative will not be continuous at

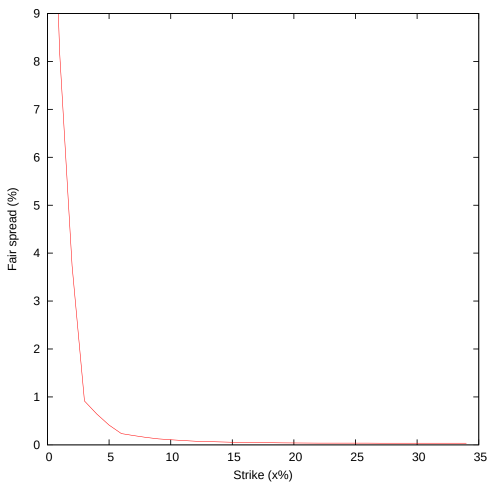

Figure 10: Fair spreads for tranchlets with width 0.5% at various points on the capital structure, using piecewise quadratic interpolation on the BaseEL curve.

this control point, but this is not a condition for the expected loss curve to be arbitrage free.

There is typically a range of choices of gradient at the right extreme that will produce monotonic and convex base EL curves, without requiring re-starts. Our choice of gradient is arbitrary, but usually lies in this range. While different choices in the range will give different EL curves, in practice we find there is little difference between the EL curves produced.

Figure 9 shows the base EL curve given by this method, and the implied base correlation curve, and figure 10 shows the fair spreads of various tranchlets. The base EL curve is smooth, as is the base correlation curve. The base correlation curve turns upwards as it moves from the 3% strike to 0%, which has been observed on some traded tranchlets. The calculated trancelet prices are monotonically decreasing with strike, indicating the absence of model arbitrages. In short, we believe this method fixes the various deficiencies in the other methods we have looked at in this paper.

One problem with fitting a base EL curve below 3% is that the expected loss on the entire pool is only around 1%, and all our market information comes from significantly higher strikes. In other words we are applying information from unlikely levels of loss, where systematic risk is the main driver, to strikes at levels of loss that are very likely to be incurred, where idiosyncratic risk is more important than systematic risk. In addition, the granularity of defaults may cause the expected loss curve to be less well approximated by a continuous curve. The upshot is that market participants who are able to observe tranchlet levels are likely to be significantly better informed than those who cannot. In absence of this information, one approach is to look at the range of tranche prices at the 1% level and pick a tranche price based on "trader's" perception of idiosyncratic risk. This extra price can the be used to incorporate an extra control point into the base EL interpolation.

5 Price and sensitivities of a non-standard tranche

Table 1 shows the fair spreads calculated for a 1%-2% tranche and a 4%-5% on the iTraxx using market data from 8 May 2007. It can be seen that there is a large variation in the fair spread of the 1%-2% tranche, depending on which extrapolation method is used - the largest fair spread is 72% greater than the smallest. The variation on the 4%-5% tranche is lower, but still significant. It is worth noting that each of the 3 different base EL interpolation methods produce higher fair spreads for the 1%-2% tranche than methods based on base correlation. This seems to indicate (but not prove) that base correlation should naturally increase below

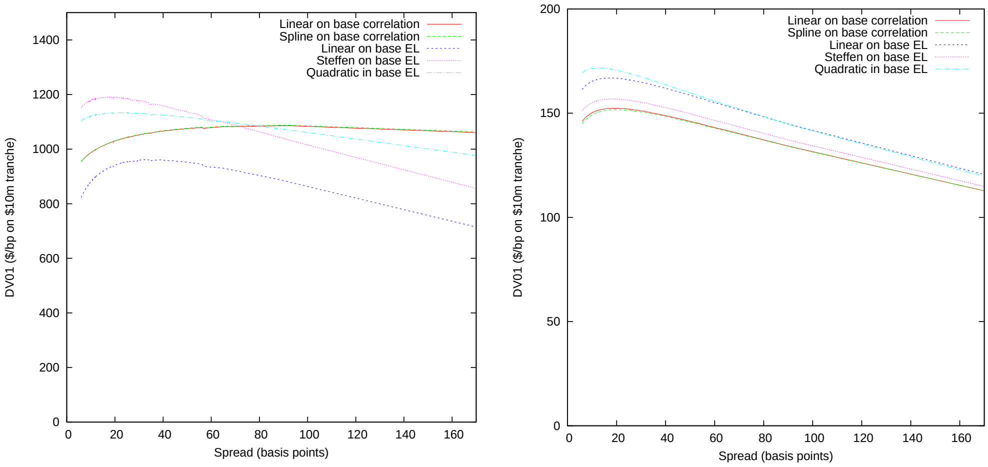

Figure 11: Deltas for different spread names. Left: 1%-2% tranche. Right: 4%-5% tranche.

3%.

Method	1%-2% tranche	4%-5% tranche
Linear interpolation on base correlation	523	42.9
Spline interpolation on base correlation	520	42.7
Linear interpolation on base EL	742	50.5
Steffen interpolation on base EL	899	45.0
Piecewise quadratic interpolation on base EL	727	50.5

Table 1: Fair spreads (in basis points) of tranches using different extrapolation methods

Figure 11 shows the credit spread dv01 on each name in the iTraxx pool as a function of spread for the 1%-2% and 4%-5% tranches as a function of each obligor's CDS spread. Each dv01 is calculated as the change in PV of a $10m tranche with a parallel shift CDS spread movement of 1 basis point. We use a contractual spread of 5% for the 1%-2% tranche, and 40 basis points for the 4%-5% tranche.

The deltas calculated for the 1%-2% tranche show some variation, particularly for very low or high spread names, again highlighting the danger of extrapolating outside the 3%-22% region. By comparison, the deltas for the 4%-5% tranche are reasonably stable to the choice of interpolation method used.

6 Conclusion

In this paper we examined various methods of constructing a base correlation curve through five points observable from the market. We demonstrated that it is easy to generate model arbitrages when using base correlation as the fundamental measure of tranche risk, and that base correlation curves do not directly indicate whether such arbitrages exist (without pricing tranchlets). We examined a more intuitive risk measure for tranches - the expected loss of equity tranches - in which model arbitrage can instantly be observed, and showed the boundaries and behaviours which this function must obey. For the iTraxx Europe 8 May 2007 we constructed base EL curves using three interpolation schemes and for each tested whether the function behaved acceptably. We noted that quadratic interpolation met all our criteria for an acceptable base EL curve. Finally we looked at prices of 2 tranchlets using base correlation and base EL and noted that for the equity tranche, each of the 3 different base EL interpolations implied that base correlation increases below 3%.

7 Acknowledgements

The authors would like to thank Jon Gregory, for first bringing the base EL approach to our attention, and Simon Greaves for his help.

References

[1] L. Andersen and J. Sidenius. Extensions to the gaussian copula: Random recovery and random factor loadings. 2004.

[2] L. Andersen, J. Sidenius, and S. Basu. All your hedges in one basket. *RISK*, November 2003.

[3] X. Burtshell, J. Gregory, and J.-P. Laurent. Beyond the gaussian copula: Stochastic and local correlation. 2005.

[4] A. Kalemanova, B. Schmid, and R. Werner. The normal inverse gaussian distribution for synthetic cdo pricing. 2005.

[5] J.-P. Laurent and J. Gregory. Basket default swaps, cdo's and factor copulas, 2003.

[6] L. McGinty and R. Ahluwalia. A model for base correlation calculation. 2004.

[7] E. Parcell. Loss unit interpolation in the collateralized debt obligation pricing model. 2006.

[8] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling. *Numerical Recipes in C++: The Art of Scientific Computing.* Cambridge University Press, Cambridge (UK) and New York, 2nd edition, 2002.

[9] D. Shelton. Base correlation: The term structure dimension. 2005.

[10] M. Steffen. A simple method for monotonic interpolation in one dimension. *Astronomy and Astrophysics*, 239:443–+, 1990.

A Shape-preserving piecewise quadratic interpolation

A set of points $(x_0, y_0), (x_1, y_1), \ldots, (x_N, y_N)$ are given with $x_0 < x_1 < \cdots < x_N$. For $i = 1, 2, .., N-1$, let

$$\Delta_i = \frac{y_{i+1} - y_i}{x_{i+1} - i_k}.$$

For any choice of z_0, the function

$$s(x) = y_i + z_i(x - x_i) + \frac{z_{i+1} - z_i}{2(x_{i+1} - x_i)}(x - x_i)^2 \quad \text{for } x_i < x < x_{x+1} \tag{15}$$

$$\text{with} \quad z_{i+1} = 2\Delta_i - z_i \tag{16}$$

satisfies $s(x_i) = y_i$ for $i = 0, 1, \ldots, N$, and is continuous and has a continuous first derivative on $[x_0, x_N]$.

In addition the points satisfy a strict monotonicity constraint,

$$y_0 < y_1 < \cdots < y_N$$

and a convexity constraint,

$$x_i < x_j < x_k \Rightarrow y_j > y_i + \frac{x_j - x_i}{x_k - x_i}(y_k - y_i).$$

These convexity and monotonicity constraints are equivalent to

$$0 < \Delta_{N-1} < \Delta_{N-2} < \cdots < \Delta_1 < \Delta_0.$$

We require

$$x < y \Rightarrow s(x) < s(y) \tag{17}$$

$$x < y < z \Rightarrow s(y) > s(x) + \frac{y - x}{z - x}[s(y) - s(x)] \tag{18}$$

which is equivalent to

$$0 < z_N < z_{N-1} < \cdots < z_1 < z_0$$

This is not necessary possible. E.g. consider $(0, 0), (1, 1), (2, 2 - \epsilon), (3, 2 + \epsilon)$ for some small

$\epsilon > 0$. Then we have

$$z_0 > z_1 = 2 - z_0 \therefore z_0 > 1 \tag{19}$$

$$z_3 = 6\epsilon - z_0 > 0 \therefore z_0 < 6\epsilon \tag{20}$$

Clearly it is not possible to simultaneously satisfy these inequalities if $\epsilon < \frac{1}{6}$.

We use the following heuristic to generate s:

- Set $z_N = \frac{1}{2}\Delta_{N-1}$.
- Calculate z_k for $k = N-1, N-2, \ldots$.
- If, for some j, $z_j > \Delta_{j-1}$ or $z_j < \Delta_j$, we use the reduced data set (x_0, y_0), (x_1, t_1), …, (x_{j+1}, y_{j+1}), with
$$z_{j+1} = \frac{\Delta_j + \Delta_{j+1}}{2}$$
to generate the curve s for $x \leq x_{j+1}$. We retain s for $x > x_{j+1}$. This will cause a discontinuity in the first derivative of s at $x = x_{j+1}$, but ensures s is monotonic and convex.
- If this choice of z_{j+1} still causes concavity, we try again with the reduced data set (x_0, y_0), (x_1, t_1), …, (x_{j+2}, y_{j+2}), and repeatedly try increasing the reduced data set until convexity is restored.