A NONPARAMETRIC COPULA MODEL FOR CDO TRANCHE PRICING

ED PARCELL

April 16, 2026

Abstract. The one-factor Gaussian copula is the standard model for pricing synthetic CDO tranches, but no single correlation parameter can reproduce the prices of all liquid tranches simultaneously—the correlation smile. We introduce the *arbitrary normal transform* (ANT) copula, in which the market factor follows a nonparametric distribution defined by a monotone function mapping normal quantiles to those of the desired distribution. The model nests the Gaussian copula as a special case and retains its structural advantages: a single correlation parameter, single-name default probabilities, and the efficient Andersen–Sidenius–Basu recursion for the loss distribution. Calibration to CDX IG 5Y tranche data from 2004–2005 achieves fits to within approximately 1 bp on all liquid tranches. The calibrated market factor distribution exhibits a fat left tail, consistent with the market's pricing of elevated systemic stress risk. The approach was first attempted by the author in 2006 but was impractical with the computational resources available at the time; modern hardware and numerical methods make it viable.

1. Introduction

The one-factor Gaussian copula became the market standard model for pricing synthetic CDO tranches following the work of Andersen, Sidenius and Basu [1], Hull and White [10], and Gregory and Laurent [7]. By conditioning on a single common market factor, the model reduces the joint default problem to a set of conditionally independent Bernoulli trials, and the loss distribution can be built by a fast recursion. Given a loss distribution, the fair spread of any tranche can be computed in closed form.

The model has a single free parameter: the pairwise default correlation ρ. No single choice of ρ is able to reproduce the prices of all liquid tranches simultaneously—the well-known *correlation smile*. The market response was to adopt *base correlation*, in which a different ρ is used for each "base" tranche (one with an attachment point of zero). A non-base tranche is priced as the difference of two base tranches, each with its own correlation. Base correlation replaced the earlier *tranche correlation* convention, which suffered from non-existence at certain market levels during the May 2005 correlation crisis and produced poor hedge ratios. See Parcell and Wood [13] for a detailed treatment, and Garcia and Goossens [5] for an arbitrage-free interpolation framework based on expected tranche loss.

Base correlation is a quoting convention, not a model. It does not provide a single consistent distribution of portfolio losses: each base tranche implies a different loss

The author thanks James Wood for his support and encouragement, and for early discussions on the ANT concept in 2006.

distribution, and the two distributions used to price a non-base tranche are not generally compatible. This creates problems for interpolation, extrapolation, and risk management.

A natural goal is a model that fits all liquid tranche prices with a single, internally consistent loss distribution, while nesting the Gaussian copula as a special case. This is analogous to the role played by the SABR model [8] for interest rate volatility smiles: a parsimonious extension of the standard model (Black–Scholes in that case, Gaussian copula here) that captures the smile with a single calibration. Several parametric extensions to the Gaussian copula were proposed:

- **Double-t copula** (Hull and White [10]): the market factor and idiosyncratic factor follow scaled Student t distributions. The degrees-of-freedom parameters are constrained to be integers, leaving very little room to fit the smile.
- **NIG copula** (Kalemanova, Schmid and Werner [11]): the factors follow Normal Inverse Gaussian distributions. Convolution stability requires tight constraints between the parameters of the market and idiosyncratic distributions, limiting the model's flexibility.
- **Random factor loading** (Andersen and Sidenius [2]): the factor loading $\sqrt{\rho}$ is made a function of the market factor. A piecewise-constant specification can fit tranche prices, but generates discontinuities in the state-risk distribution and admits multiple solutions with different hedge ratios.
- **Generalised Poisson loss (GPL) model** (Brigo, Pallavicini and Torresetti [3]): a dynamical loss model that directly specifies the portfolio loss process rather than modelling individual defaults. Tractable and flexible enough to fit market prices well, but at the cost of abandoning the single-name structure—the model does not provide individual default probabilities or single-name hedge ratios.

Burtschell, Gregory and Laurent [4] provide a comparative survey of the factor copula approaches.

In this paper, we take a different approach. Rather than choosing a parametric family for the factor distributions, we define them nonparametrically using what we call the *arbitrary normal transform* (ANT). Each factor distribution is specified by a monotone function that maps normal quantiles to the quantiles of the desired distribution. The function is represented by a set of freely moving control points, interpolated using Steffen's monotonic scheme [15]. The Gaussian copula is nested as the trivial case where both ANT functions are the identity.

The idea of fitting the correlation smile by calibrating a nonparametric market factor distribution dates to 2006, when the author, then working at a credit derivatives consultancy, attempted calibration by direct specification of the ANT function at discrete points. The approach was sound in principle but impractical with the tools available at the time: a single calibration run took 36 hours, and the optimiser exploited non-monotonicity in the interpolant to achieve fits via negative probability densities. Twenty years of improvements in computation, numerical optimisation, and monotonic interpolation methods have made the approach viable. We present the method in detail and demonstrate calibration to CDX IG 5Y tranche data from 2004–2005, achieving fits to within approximately 1 bp on all liquid tranches.

2. The one-factor copula model

This section describes the one-factor copula model and the procedure for pricing CDO tranches within it. We aim to give a complete, implementable description: a reader with access to a numerical library and this section alone should be able to build a working pricer.

2.1. **Default probabilities from CDS spreads.** Consider a portfolio of N credits. For each credit i, the risk-neutral default probability is implied by the credit default swap (CDS) market. A CDS is a contract in which the protection buyer pays a periodic spread s on a notional I until default or maturity, and in return receives $I(1-R)$ if the reference credit defaults, where R is the recovery rate.

Let $\lambda_i(t)$ be the hazard rate (instantaneous conditional default intensity) for credit i. The survival probability to time t is:

$$q_i(t) = \exp\left(-\int_0^t \lambda_i(u)\,du\right) \tag{1}$$

and the cumulative default probability is $p_i(t) = 1 - q_i(t)$.

The present value of the CDS premium leg, paid at dates $T_1, \ldots, T_n$, is:

$$\mathrm{PV}^{\mathrm{CDS}}_{\mathrm{prem}} = s \cdot I \sum_{j=1}^{n} D(T_j)\,\alpha_j\,q_i(T_j) \tag{2}$$

where $D(t)$ is the risk-free discount factor and α_j is the accrual fraction for period j. For simplicity we omit the accrued premium at default; including it would add a small correction but does not affect the structure of the model.

The present value of the CDS protection leg is:

$$\mathrm{PV}^{\mathrm{CDS}}_{\mathrm{prot}} = I(1-R)\sum_{j=1}^{n} D(\bar{T}_j)\left(q_i(T_{j-1}) - q_i(T_j)\right) \tag{3}$$

where $\bar{T}_j = \frac{1}{2}(T_{j-1}+T_j)$ approximates the expected default time within each period.

The par CDS spread is the value of s at which the two legs are equal. Given a quoted spread, we can invert this relationship to find the hazard rate. For a flat spread and piecewise-constant hazard rate, the calculation is a straightforward bootstrap. In the homogeneous case where all credits share the same spread (as we assume when using the index level), the hazard rate is simply:

$$\lambda = \frac{s}{1-R} \tag{4}$$

and $p(t) = 1 - e^{-\lambda t}$. Throughout this paper we assume $R = 40\%$.

2.2. **Latent variables and default.** Default of credit i by time t is modelled by a latent random variable A_i. Credit i defaults if and only if A_i falls below a threshold calibrated to match the observed default probability:

$$A_i < F_A^{-1}(p_i(t))$$

where F_A is the cumulative distribution function of A_i.

The latent variable is decomposed into a systematic (market) component shared across all credits and an idiosyncratic component specific to each credit:

$$A_i = \sqrt{\rho}\,M + \sqrt{1-\rho}\,\varepsilon_i \tag{5}$$

where M and each ε_i are independent with mean 0 and variance 1, and $\rho \in (0,1)$ is the pairwise default correlation. In the standard Gaussian copula, M and ε_i are standard normal, A_i is also standard normal, and $F_A^{-1} = \Phi^{-1}$.

2.3. **Conditional default probability.** Given a realisation $M = m$ of the market factor, the credits are conditionally independent. The probability that credit i defaults by time t, conditional on $M = m$, is:

$$p_i(t \mid m) \;=\; \Phi\!\left(\frac{\Phi^{-1}(p_i(t)) - \sqrt{\rho}\, m}{\sqrt{1-\rho}}\right) \tag{6}$$

When m is large and negative—a bad market outcome—the argument to Φ is pushed to the right, and the conditional default probability rises for every credit simultaneously. This is how the model generates correlated defaults.

In the generalised setting of Section 3, where M and ε_i may be non-normal, equation (6) is replaced by a more general expression. We present the Gaussian version here because it is the standard case and the starting point for any implementation.

2.4. **Building the conditional loss distribution.** Conditional on $M = m$, defaults are independent Bernoulli trials. The distribution of portfolio losses can therefore be built efficiently using a recursion due to Andersen, Sidenius and Basu [1].

We discretise losses into units. In the simplest case—a homogeneous pool where every credit has the same notional and recovery rate—each default produces one unit of loss, and the maximum number of loss units is N. (For heterogeneous pools, see Parcell [12] on loss unit interpolation.)

Let $L_j^{(k)}$ denote the probability that exactly j units of loss have occurred after considering the first k credits. Initialise with the empty portfolio:

$$L_j^{(0)} = \begin{cases} 1 & j = 0 \\ 0 & j > 0 \end{cases}$$

Then for each credit $k = 1, \dots, N$, letting $q_k = p_k(t \mid m)$ and l_k be the number of loss units for credit k:

$$L_j^{(k)} = \begin{cases} (1-q_k)\, L_j^{(k-1)} & j < l_k \\ (1-q_k)\, L_j^{(k-1)} + q_k\, L_{j-l_k}^{(k-1)} & j \geq l_k \end{cases} \tag{7}$$

After processing all N credits, the vector $L^{(N)}$ is the conditional loss distribution. The computation is $O(N \times l_{\max})$ where $l_{\max}$ is the total number of loss units.

This recursion is run once for each time step at which a premium payment is due (typically quarterly), using the cumulative default probabilities $p_i(t)$ for that horizon.

2.5. **Pricing a CDO tranche.** A CDO tranche is defined by an attachment point L_a and a detachment point L_d, expressed as fractions of the portfolio notional. For example, the equity tranche of the CDX IG index has $L_a = 0$, $L_d = 0.03$.

The tranche has two legs:

Default (protection) leg. The protection seller pays the buyer for any losses that fall within the tranche $[L_a, L_d]$. Let $E(t)$ denote the expected tranche loss at time t:

$$E(t) = \mathbb{E}[\min(\Lambda(t),\, L_d) - \min(\Lambda(t),\, L_a)]$$

where $\Lambda(t)$ is the cumulative portfolio loss at time t, expressed as a fraction of portfolio notional. Given the discretised loss distribution, $E(t)$ is computed by summing over loss states.

Let $D(t)$ be the risk-free discount factor to time t, and let $T_1, \ldots, T_n$ be the premium payment dates, with midpoints $\bar{T}_i = \frac{1}{2}(T_{i-1} + T_i)$ and $T_0 = 0$. The present value of the default leg is:

$$\mathrm{PV}_{\mathrm{def}} = \sum_{i=1}^{n} D(\bar{T}_i)\,\big(E(T_i) - E(T_{i-1})\big) \tag{8}$$

Premium (fee) leg. The protection buyer pays a periodic coupon at rate s on the outstanding tranche notional, which decays as losses accumulate. The present value of the premium leg is:

$$\mathrm{PV}_{\mathrm{prem}}(s) = s \sum_{i=1}^{n} D(T_i)\,\alpha_i\,\big(L_d - L_a - \tfrac{1}{2}(E(T_i) + E(T_{i-1}))\big) \tag{9}$$

where α_i is the day-count fraction for the ith period.
Fair spread. The fair spread s^* is the coupon rate at which the two legs have equal present value:

$$s^* = \frac{\mathrm{PV}_{\mathrm{def}}}{\sum_{i=1}^{n} D(T_i)\,\alpha_i\,\big(L_d - L_a - \frac{1}{2}(E(T_i) + E(T_{i-1}))\big)} \tag{10}$$

Equity tranche. For the equity tranche (0–3%), the market convention during the period covered by our data was to quote an upfront percentage plus a fixed running spread of 500 bp. The present value of the tranche at the market quote is:

$$\mathrm{PV} = \mathrm{PV}_{\mathrm{def}} - \mathrm{PV}_{\mathrm{prem}}(500\,\mathrm{bp}) - U \cdot (L_d - L_a) \tag{11}$$

where U is the upfront percentage. At the market quote, $\mathrm{PV} = 0$.

2.6. **Integrating over the market factor.** Everything above is conditional on a realisation $M = m$. To obtain the unconditional expected tranche loss $E(t)$, we integrate over the distribution of M:

$$E(t) = \int_{-\infty}^{\infty} E(t \mid m)\; f_M(m)\, dm \tag{12}$$

In the Gaussian copula, $f_M = \phi$ and this integral is naturally suited to Gauss–Hermite quadrature. Gauss–Hermite evaluates integrals of the form $\int_{-\infty}^{\infty} g(x)\, e^{-x^2} dx$ using K abscissae $\{u_k\}$ and weights $\{w_k\}$. After the change of variable $m = \sqrt{2}\, u$, the integral (12) becomes:

$$E(t) \approx \frac{1}{\sqrt{\pi}} \sum_{k=1}^{K} w_k\, E(t \mid \sqrt{2}\, u_k) \tag{13}$$

In practice, $K = 100$ Gauss–Hermite points provide more than adequate precision. In Section 3 we show how to preserve this efficient quadrature when M has a non-Gaussian ANT distribution.

2.7. **Implementation summary.** For reference, here is the complete pricing procedure:

(1) **Bootstrap hazard rates.** From the index spread (or individual CDS spreads), compute the hazard rate using (4) and the cumulative default probability $p(t) = 1 - e^{-\lambda t}$ at each premium date.
(2) **Set up quadrature.** Compute Gauss–Hermite abscissae and weights. Transform the abscissae to market factor values $m_k = \sqrt{2}\, u_k$.
(3) **For each quadrature point m_k:**
   (a) For each premium date T_i, compute the conditional default probability $p(T_i \mid m_k)$ using (6).
   (b) Run the recursion (7) to build the conditional loss distribution at each T_i.
   (c) Compute the conditional expected tranche loss $E(T_i \mid m_k)$ for each tranche and each T_i.
   (d) Compute the conditional PVs of the default and premium legs for each tranche.
(4) **Integrate.** Compute unconditional leg PVs by weighted sum: $\mathrm{PV} = \frac{1}{\sqrt{\pi}} \sum_k w_k \,\mathrm{PV}(m_k)$.
(5) **Compute fair spread** using (10), or the tranche PV at the market-quoted spread using (11).

3. The ANT copula

In this section we introduce the *arbitrary normal transform* (ANT) copula, in which the distributions of the market factor and idiosyncratic factor are defined nonparametrically.

3.1. The ANT distribution.

Definition 1. *Let $Z \sim N(0,1)$ and let $h : \mathbb{R} \to \mathbb{R}$ be a strictly monotonically increasing function. Define a random variable X by its cumulative distribution function:*

$$F_X(x) = \Phi(h(x))$$

Then X has an arbitrary normal transform distribution*, written $X \sim \mathcal{ANT}(h)$, and h is the* ANT function *of X.*

The interpretation is direct: h maps each value of X to the corresponding quantile in the standard normal distribution. When h is the identity, X is standard normal. When h grows faster than the identity in the tails—$h(x) > x$ for large $|x|$—the distribution of X has thinner tails than the normal. Conversely, $h(x) < x$ in the tails gives a fat-tailed distribution.

Lemma 1. *Any continuous distribution can be represented as an ANT distribution.*

Proof. Let X be any continuous random variable. Take $h(x) = \Phi^{-1}(F_X(x))$, and let $Y \sim \mathcal{ANT}(h)$. Then $F_Y(x) = \Phi(h(x)) = \Phi(\Phi^{-1}(F_X(x))) = F_X(x)$, so Y and X have the same distribution. □

For example, the ANT function of a Student t distribution with ν degrees of freedom (rescaled to unit variance) is $h(x) = \Phi^{-1}\Big(F_{t_\nu}(x\sqrt{(\nu-2)/\nu})\Big)$. These functions are smooth, gently curving away from the identity, and are well-suited to interpolation.

3.2. **The ANT copula model.** In the most general form, we replace the standard Gaussian assumption with ANT distributions for both factors:

$$M \sim \mathcal{ANT}(h_M) \tag{14}$$

$$\varepsilon_i \sim \mathcal{ANT}(h_\varepsilon) \quad \text{i.i.d.} \tag{15}$$

with the latent variable defined as before:

$$A_i = \sqrt{\rho}\, M + \sqrt{1-\rho}\, \varepsilon_i$$

In practice, calibration experiments show that the idiosyncratic distribution has negligible leverage on tranche prices (see Section 5), and we fix h_ε as the identity (i.e., $\varepsilon_i \sim N(0,1)$) throughout. Only h_M is calibrated. The framework supports a general h_ε without modification.

For ρ to be the pairwise default correlation, we require M and ε_i to have mean 0 and variance 1. Since the ANT function is unconstrained, the raw distribution defined by h_M will not in general satisfy this. We therefore *standardise*: after evaluating the ANT function, we compute the mean μ and standard deviation σ of the resulting distribution by direct numerical integration, and work with the standardised variable $(X-\mu)/\sigma$. This is done once per calibration step at negligible cost. The standardisation eliminates the location and scale degeneracy that would otherwise exist in the parameterisation, and ensures that ρ retains its interpretation as the default correlation.

3.3. **Computing F_A.** Since M and ε_i are no longer Gaussian, A_i is no longer Gaussian, and F_A must be computed numerically. Expressing F_A as an integral over the market factor:

$$F_A(x) = \int_{-\infty}^{\infty} F_\varepsilon\left(\frac{x - \sqrt{\rho}\, m}{\sqrt{1-\rho}}\right) f_M(m)\, dm \tag{16}$$

where $F_\varepsilon(x) = \Phi(h_\varepsilon(x))$ and $f_M(m) = \phi(h_M(m))\, h_M'(m)$. The integrand involves only the CDF of ε (not its density), making the integral well-behaved.

Using the same change of variable as in Section 3.5—substituting $u = h_M(m)$—this becomes a standard expectation under the normal distribution, evaluated by Gauss–Hermite quadrature.

Since F_A depends only on h_M, h_ε, and ρ—not on any particular credit or time horizon—we precompute it once per calibration step.

We evaluate (16) on a fixed grid of $\sim$2000 points via Gauss–Hermite quadrature, then interpolate F_A using Steffen's method. The inverse F_A^{-1} is computed by root-finding. The interpolation error converges as $O(h^4)$, consistent with Steffen's cubic scheme; at 2001 grid points the maximum tranche pricing error from discretisation is below 0.01 bp (see Appendix D).

An equivalent route is via the density convolution:

$$f_A(x) = \frac{1}{\sqrt{1-\rho}} \int_{-\infty}^{\infty} f_\varepsilon\left(\frac{x - \sqrt{\rho}\, m}{\sqrt{1-\rho}}\right) f_M(m)\, dm \tag{17}$$

The scaled densities can be represented as Chebyshev polynomial approximations on a truncated domain and convolved using the Hale–Townsend algorithm [6], with F_A obtained by indefinite integration and F_A^{-1} by root-finding.[1]

[1]The `chebpy` library (M. Richardson, `github.com/chebpy/chebpy`), a Python implementation of the Chebfun system [16], provides these operations.

3.4. **Conditional default probability in the ANT copula.** The default threshold for credit i at horizon t is $c_i(t) = F_A^{-1}(p_i(t))$, obtained from the precomputed interpolant. The conditional default probability is:

$$p_i(t \mid m) = F_\varepsilon\left(\frac{c_i(t) - \sqrt{\rho}\, m}{\sqrt{1-\rho}}\right) = \Phi\left(h_\varepsilon\left(\frac{c_i(t) - \sqrt{\rho}\, m}{\sqrt{1-\rho}}\right)\right) \tag{18}$$

This reduces to the standard formula (6) when h_M and h_ε are both the identity.

3.5. **Gauss–Hermite quadrature with ANT distributions.** A key practical advantage of the ANT formulation is that Gauss–Hermite quadrature extends naturally to non-Gaussian market factors. The integral over the market factor is:

$$\int_{-\infty}^{\infty} g(m)\, f_M(m)\, dm = \int_{-\infty}^{\infty} g(m)\, \phi(h_M(m))\, h_M'(m)\, dm$$

Substituting $u = h_M(m)$, so that $m = h_M^{-1}(u)$:

$$= \int_{-\infty}^{\infty} g\big(h_M^{-1}(u)\big)\, \phi(u)\, du \tag{19}$$

This is a standard integral against the normal density and can be evaluated by Gauss–Hermite quadrature in the usual way. The only modification is that the standard Gauss–Hermite abscissae u_k are transformed through h_M^{-1} to obtain the market factor values:

$$m_k = h_M^{-1}(\sqrt{2}\, u_k) \tag{20}$$

The weights are unchanged. The inverse h_M^{-1} is computed by root-finding on the Steffen interpolant for h_M.

This change-of-variable approach avoids the need for adaptive quadrature, which would introduce discontinuities in the objective function as quadrature points appear or disappear during calibration. Such discontinuities are problematic for any optimiser: gradient-based methods compute meaningless derivatives at the jumps, while gradient-free methods may converge slowly or to spurious minima. With Gauss–Hermite, the number and weights of quadrature points are fixed throughout the optimisation; only the locations m_k change smoothly as the ANT function evolves. We note that Gauss–Hermite convergence for the ANT integrand is algebraic rather than spectral, owing to the non-smoothness introduced by h_M^{-1}; see Appendix E for a convergence study.

3.6. **Representing the ANT function.** Each ANT function is represented by a set of control points $\{(x_j, y_j)\}_{j=0}^{n}$ with $x_0 < x_1 < \cdots < x_n$ and $y_0 < y_1 < \cdots < y_n$, interpolated using Steffen's monotonic piecewise cubic method [15]. Outside the range of the control points, the ANT function is extended linearly with slope 1. That is, for $x < x_0$, $h(x) = y_0 + (x - x_0)$, and similarly for $x > x_n$. This means the distribution is exactly normal in the far tails: any departure from normality is confined to the region covered by the control points. If the optimiser needs to control tail behaviour in a particular region, it places control points there.

Steffen's method determines the slope at each control point from local data only (the three-point parabola through the point and its two neighbours, clamped to ensure monotonicity). This guarantees that the interpolant is monotonically increasing, has continuous first derivatives, and depends only on nearby control points. These properties are essential for our application: monotonicity is required

for h to define a valid distribution, locality ensures that moving one control point does not cause oscillations elsewhere, and C^1 continuity keeps the density $f_X(x) = \phi(h(x))\,h'(x)$ continuous.

3.7. **Summary of changes to the pricing procedure.** Relative to the standard Gaussian implementation (Section 2), the ANT copula requires three modifications:

(1) **Precompute F_A and F_A^{-1}.** Evaluate equation (16) on a grid and construct Steffen interpolants (Section 3.3).
(2) **Replace the conditional default probability.** Use equation (18) instead of (6). The default threshold $c_i(t)$ is obtained from the precomputed F_A^{-1} interpolant.
(3) **Transform the Gauss–Hermite abscissae.** Map the standard abscissae through h_M^{-1} using (20). The weights are unchanged.

The recursion for the loss distribution and the leg PV calculations are unchanged.

4. Calibration

4.1. **Parameters.** The model parameters are:

- The correlation ρ.
- The y-coordinates of the control points defining h_M, with x-coordinates fixed.
- The y-coordinates of the control points defining h_ε, with x-coordinates fixed.

For each ANT function, the x-coordinates are fixed at $n+1$ equally spaced points spanning a range $[x_{\min}, x_{\max}]$ (we use $[-6, 6]$, covering six standard deviations on each side). The y-coordinates are the free parameters. The total number of free parameters is $1 + n_M + n_\varepsilon$.

4.2. **Unconstrained parameterisation.** The y-coordinates must be strictly ordered: $y_0 < y_1 < \cdots < y_n$, with the endpoints pinned at $y_0 = x_{\min}$ and $y_n = x_{\max}$ so that h is the identity in the far tails. To achieve this with unconstrained parameters, we parameterise the *gaps* between adjacent y-values using a softmax transformation. Given n unconstrained weights $\omega_1, \ldots, \omega_n$:

$$g_j = \frac{e^{\omega_j}}{\sum_{k=1}^{n} e^{\omega_k}}, \qquad y_j = x_{\min} + (x_{\max} - x_{\min}) \sum_{k=1}^{j} g_k, \quad j = 1, \ldots, n \tag{21}$$

with $y_0 = x_{\min}$. Since the softmax outputs are positive and sum to one, the y-coordinates are automatically ordered and span $[x_{\min}, x_{\max}]$. When all weights are equal, $y_j = x_j$ and h is the identity.

The same scheme is applied independently to both h_M and h_ε, giving a single flat vector of unconstrained parameters that can be passed to any gradient-free or gradient-based optimiser.

4.3. **Focused x-grid.** The fixed x-grid need not be equally spaced. We allow the optimiser to concentrate knots in regions where the distribution departs most from normality by parameterising the x-positions as quantiles of a mixture distribution:

$$F_{\text{grid}}(x) = (1 - s)\,F_{\text{unif}}(x) + s\,F_{\text{TN}}(x; \mu_f, \sigma_f) \tag{22}$$

where F_{unif} is the uniform CDF on $[x_{\min}, x_{\max}]$ and F_{TN} is a truncated normal with mean μ_f (the *focus*) and standard deviation σ_f (the *tightness*), truncated

to $[x_{\min}, x_{\max}]$. The parameter $s \in [0,1]$ (the *strength*) blends between uniform spacing ($s = 0$) and concentrated spacing ($s = 1$). The $n-1$ interior x-positions are obtained by evaluating F_{grid}^{-1} at equally spaced quantile levels.

The focus, tightness and strength are optimiser parameters, allowing the model to adaptively place resolution where it is most needed. In practice, the calibrated focus tends to lie in the left tail ($\mu_f \approx -3$ to -4), reflecting the fact that the correlation smile is driven by the probability of adverse market outcomes.

4.4. **Bump basis.** A practical challenge with the softmax parameterisation is that perturbing a single weight ω_j moves only the jth gap in the y-sequence, producing a highly localised change in h. For gradient-based optimisers, this means each parameter has very little leverage on the tranche prices, resulting in near-zero gradients and poor convergence.

To address this, we map the raw parameters through a *bump basis matrix* B before applying the softmax:

$$\tilde{\omega}_j = \sum_k B_{jk}\,\omega_k, \qquad B_{jk} = \exp\left(-\frac{(\bar{x}_j - \bar{x}_k)^2}{2\beta^2}\right) \tag{23}$$

where $\bar{x}_j$ is the midpoint of the jth x-interval and β is a bandwidth in x-space. We define a base bandwidth $\bar{d} = \frac{1}{2}$ times the average spacing between adjacent x-knots. The matrix B is symmetric, positive definite, and banded: perturbing a single raw parameter ω_k creates a smooth Gaussian bump across several softmax inputs. Because the kernel operates in x-space rather than index space, the coupling between knots reflects their actual proximity in the distribution—knots that are close together in a focused region are more tightly coupled than knots that are far apart in a sparse region.

4.5. **Objective function.** For each tranche k, the model produces a present value V_k at the market-quoted coupon (or, for the equity tranche, at 500 bp running plus the quoted upfront). At the correct model parameters, $V_k = 0$ for all k. We minimise:

$$\mathcal{L} = \sum_{k=1}^{K} V_k^2 + \lambda\,\mathcal{R} \tag{24}$$

where $\mathcal{R}$ is a regularisation term and $\lambda > 0$ controls its strength.

Working with present values rather than fair spreads avoids mixing units (upfront percentage for equity, running spread for other tranches) and eliminates the non-linear transformation from PV to spread in the objective. In practice, we scale the PVs to a notional of \$100 million before squaring, so that the pricing error terms are $O(10^{10})$ rather than $O(10^{-6})$; this improves numerical visibility for gradient-based polishing steps.

4.6. **Entropy regularisation.** The normal distribution has the maximum differential entropy among all continuous distributions with a given variance. When the ANT functions are the identity, the model reduces to the Gaussian copula. We use the *negentropy*—the difference between the entropy of the normal and the entropy of the ANT distribution—as a regulariser, encouraging the solution to stay close to the Gaussian copula unless the data demands otherwise.

For a random variable $X \sim \mathcal{ANT}(h)$ with density $f_X(x) = \phi(h(x))\, h'(x)$, the differential entropy is:

$$H(X) = -\int_{-\infty}^{\infty} f_X(x) \ln f_X(x)\, dx \tag{25}$$

The differential entropy of a standard normal is $H_N = \frac{1}{2}\ln(2\pi e)$. The regularisation term is:

$$\mathcal{R} = \big(H_N - H(M)\big) + \big(H_N - H(\varepsilon)\big) \tag{26}$$

Both terms are non-negative (by the maximum entropy property of the normal after standardisation to unit variance), and are zero if and only if the corresponding distribution is normal.

The entropy integral (25) is evaluated by direct numerical integration (trapezoidal rule) on a fine grid. The integrand $-f(x)\ln f(x)$ tends to zero naturally as $f(x) \to 0$, so regions where the density vanishes contribute nothing and require no special treatment.

4.7. **Choice of optimiser.** The objective function is non-convex, and the relationship between the raw parameters and tranche PVs is mediated by numerical integration, interpolation, and the loss distribution recursion. We use *differential evolution* [14], a population-based global optimiser that evaluates a population of candidate solutions at each generation. The candidate evaluations are independent and can be distributed across multiple CPU cores, providing a near-linear speedup.

4.8. **Multi-stage bandwidth annealing.** The bump bandwidth β in (23) controls the trade-off between exploration (broad coupling, smooth deformations) and exploitation (tight coupling, localised control). We exploit this by running the optimiser in three stages with decreasing bandwidth:

1. **Coarse** ($\beta = 6\bar{d}$, 30% of generations): each parameter moves a broad region of h, enabling large-scale shape changes. The optimiser finds the right basin.
2. **Medium** ($\beta = 2\bar{d}$, 30%): moderate coupling for refinement.
3. **Fine** ($\beta = \bar{d}$, 40%): tight coupling for final tuning of individual knot positions.

where $\bar{d}$ is half the average spacing between x-knots.

At each stage transition, the population must be transformed to the new bump basis to preserve continuity. If the previous stage used bump matrix B_{old} and the new stage uses B_{new}, the raw weights are transformed as:

$$\boldsymbol{\omega}_{\text{new}} = B_{\text{new}}^{-1}\, B_{\text{old}}\, \boldsymbol{\omega}_{\text{old}} \tag{27}$$

This ensures that the softmax inputs—and therefore the distribution h_M—are unchanged across the transition. Without this transformation, the stage transition scrambles the population and can cause the optimiser to regress.

The initial population is seeded around the Gaussian copula: all ANT control points on the line $y = x$, with ρ seeded at several values spanning $[0.05, 0.5]$ to ensure diverse starting points.

4.9. **Identifiability.** The model has more parameters than tranche prices, so the calibration problem is underdetermined. This is by design: we are fitting a distribution, not just a small set of scalar parameters. The entropy regulariser resolves the ambiguity by selecting, among all distributions that fit the data, the one closest to normal in an information-theoretic sense. This is analogous to Tikhonov regularisation in inverse problems or the maximum entropy principle in statistical mechanics.

A subtler identifiability issue arises between ρ and the tail behaviour of h_M: fatter tails in the market factor can produce similar tranche prices to a higher ρ with thinner tails. The entropy regulariser addresses this by penalising departure from normality, favouring a moderate ρ with mildly non-normal tails over an extreme ρ with heavily deformed ANT functions.

To guard against the optimiser finding a local minimum rather than the global solution, we recommend running the calibration from multiple starting points—for example, varying the initial ρ across a range of values—and checking that the results are consistent.

The strength parameter λ controls the trade-off between fit and parsimony. A large λ pulls the solution towards the Gaussian copula; a small λ allows more departure from normality at the cost of potential overfitting. In practice, λ is chosen so that the pricing errors are within bid–ask spreads.

5. Results

We calibrate the ANT copula to CDX IG 5Y tranche data from two dates: August 4, 2004 and August 30, 2005. These dates represent different market conditions: August 2004 has wider credit spreads (index at 63 bp) while August 2005 reflects the tighter post-rally environment (index at 50 bp). For each date, we also compute the standard base correlation curve for comparison.

We use 20 softmax weights for h_M (giving $n + 1 = 21$ control points including endpoints), with the focused x-grid, bump basis, entropy regularisation ($\lambda = 0.01$), and multi-stage bandwidth annealing with basis change. The idiosyncratic factor ε is fixed as standard normal throughout: calibration experiments show that the idiosyncratic distribution has negligible leverage on tranche prices, consistent with the correlation smile being driven by the market factor.

5.1. **Base correlation.** As a benchmark, Table 1 shows the bootstrapped base correlations for each date. The characteristic upward slope from equity to super-senior—the correlation smile—is present on both dates. The 2005 date exhibits a steeper smile, with base correlation ranging from 12% at the 3% detachment point to 68% at 30%.

Table 1. Bootstrapped base correlations.

Detachment	2004-08-04	2005-08-30
3%	19.9%	11.8%
7%	27.0%	27.2%
10%	30.3%	35.4%
15%	37.8%	45.9%
30%	57.8%	68.6%

The base correlation curves are plotted together in Figure 5 (Appendix C).

5.2. **ANT calibration: fitted tranche spreads.** Table 2 compares the market-quoted tranche spreads with the ANT copula model prices for each date. On both dates, the model fits all five liquid tranches to within a few basis points—well within typical bid–ask spreads, and dramatically better than the Gaussian copula at any single correlation.

Table 2. ANT copula calibration: market vs model tranche spreads. The equity tranche is quoted as upfront percentage; all others in basis points.

	2004-08-04 ($\rho = 0.135$)			2005-08-30 ($\rho = 0.086$)		
Tranche	Market	Model	Error	Market	Model	Error
0–3%	41.8%uf	41.8%uf	+0.0	40.0%uf	40.0%uf	+0.0
3–7%	347.0	346.7	−0.3	127.0	127.3	+0.3
7–10%	135.5	135.1	−0.4	35.5	34.8	−0.7
10–15%	47.5	47.4	−0.1	20.5	20.1	−0.4
15–30%	14.5	14.1	−0.4	9.5	13.0	+3.5

The August 4, 2004 calibration achieves a near-perfect fit: all five tranches match to within 0.4 bp, with a calibrated $\rho = 0.135$. This demonstrates that the ANT copula can reproduce the full tranche capital structure with a single, consistent loss distribution.

The August 30, 2005 calibration fits four of five tranches to within 1 bp, with the 15–30% super-senior tranche approximately 3.5 bp wide. The calibrated $\rho = 0.086$ is lower than for 2004, reflecting the tighter spread environment: the non-normal market factor distribution, rather than a high correlation parameter, is doing the work of generating the correlation smile.

5.3. **The calibrated market factor distribution.** Figures 1 and 3 (Appendices A and B) show the calibrated ANT functions h_M and the implied market factor densities for each date. Several features are consistent across both calibrations:

- The ANT function lies below the identity ($h_M(x) < x$) for $x \ll 0$, creating a fatter left tail than the normal distribution. This means the model assigns higher probability to extreme adverse market outcomes than the Gaussian copula.
- The ANT function lies above the identity for moderate negative x, concentrating probability mass in a region that produces conditional default rates consistent with the observed tranche prices.
- The resulting density is left-skewed and more peaked than the normal, with a heavier left tail and a thinner right tail.

The full calibration charts, showing h_M, the implied density, and log-density plots, are presented in Appendix A (August 30, 2005) and Appendix B (August 4, 2004).

This shape is economically intuitive. The correlation smile tells us that the market prices a higher probability of systemic stress events (many simultaneous defaults) than the Gaussian copula implies. The ANT copula captures this through the fat left tail of M: in bad market states, the conditional default probability is higher than the Gaussian model would predict, driving up the price of protection on equity and mezzanine tranches.

The calibrated ρ values are notably lower than the mezzanine base correlations (0.086–0.135 vs 0.27–0.29). This reflects a fundamental difference in how the two approaches generate correlation: base correlation uses a high ρ with a normal distribution, while the ANT copula uses a lower ρ with a non-normal distribution that has more probability mass in the tails. The latter provides a single consistent model—the same ρ and the same loss distribution apply to all tranches simultaneously, with the distribution shape accounting for the smile.

5.4. **Comparison with the Gaussian copula.** To illustrate the improvement over the standard model, we compare the ANT copula fit with the Gaussian copula at several fixed correlations. No single correlation can reproduce the market: at $\rho = 0.08$ (near the ANT calibrated value), the equity tranche is reasonably priced but the mezzanine is approximately 240 bp vs the market's 127 bp, and the senior tranches are near zero. Increasing ρ improves the senior fit but worsens the equity and mezzanine.

The ANT copula resolves this trade-off by allowing the distribution shape, rather than the correlation parameter alone, to account for the variation across tranches.

TABLE 3. Gaussian copula at various ρ vs market (August 30, 2005). No single ρ can fit all tranches simultaneously.

	0–3%	3–7%	7–10%	10–15%	15–30%
Market	40.0%uf	127.0	35.5	20.5	9.5
$\rho = 0.08$	43.7%uf	238.7	24.6	2.4	0.0
$\rho = 0.15$	37.1%uf	294.3	66.1	15.4	0.9
$\rho = 0.20$	33.0%uf	315.6	92.9	29.0	2.9
$\rho = 0.30$	25.7%uf	332.9	134.0	58.1	10.7
ANT ($\rho = 0.086$)	40.0%uf	127.3	34.8	20.1	13.0

5.5. **Sensitivity analysis.**

Idiosyncratic distribution. In preliminary calibrations, we allowed the idiosyncratic factor ε to have an ANT distribution alongside the market factor. The optimiser consistently returned h_ε indistinguishable from the identity, confirming that the idiosyncratic distribution has negligible effect on tranche prices. This is expected: ε affects each name independently and therefore influences individual default probabilities but not the correlation structure that drives the tranche smile. We fix ε as standard normal in all results reported here.

Number of knots. We tested calibrations with 8 and 20 interior softmax weights. With 8 weights, the model achieves good fits (all tranches within 3–4 bp on the best dates) but has limited ability to shape the distribution finely. With 20 weights and multi-stage bandwidth annealing, the fit improves substantially: the August 4, 2004 date achieves sub-basis-point precision on all tranches. The additional resolution allows the optimiser to shape the left tail of M more precisely, which is critical for fitting the mezzanine and senior tranches simultaneously.

Regularisation strength. The entropy regularisation parameter λ controls the trade-off between fit quality and proximity to the Gaussian copula. At $\lambda \geq 0.1$, the regularisation dominates and the calibrated distribution is indistinguishable from normal. At $\lambda = 0.01$, the model departs sufficiently from normality to capture the smile while remaining smooth. We use $\lambda = 0.01$ throughout.

Bump bandwidth and multi-stage annealing. The multi-stage bandwidth schedule (Section 4.8) is important for convergence. Without it, the optimiser either finds only broad, Gaussian-like solutions (at high bandwidth) or gets stuck in local minima with jagged distributions (at low bandwidth). The coarse stage identifies the right basin of attraction; the medium and fine stages refine within it. The basis change at stage transitions (27) is critical: without it, the fine stage can regress, producing worse fits than the medium stage found.

6. Conclusion

In the mid-2000s, a great deal of effort was devoted to finding a copula model that could fit all liquid CDO tranche prices simultaneously, much as the SABR model [8] had done for interest rate volatility smiles. Parametric extensions to the Gaussian copula—Student t, NIG, random factor loading—each offered a few extra parameters, but none could consistently reproduce the full capital structure. The implied copula [9] and GPL model [3] achieved exact fits by abandoning the structural one-factor framework, but at the cost of the single correlation parameter and single-name hedge ratios that made the Gaussian copula useful in practice.

The ANT copula takes a different path. By replacing the Gaussian distribution of the market factor with a nonparametric distribution—specified by a monotone function that maps normal quantiles to those of the desired distribution—the model retains the full one-factor copula structure while gaining the flexibility to fit the smile. The calibrated ρ and h_M together provide a complete, interpretable model of default dependence: ρ is still a single correlation number, and the shape of h_M reveals the market's implied view of systemic risk in a way that base correlation cannot.

We have demonstrated calibration to CDX IG 5Y tranche data from 2004 and 2005, achieving fits to within approximately 1 bp on all liquid tranches. The calibrated market factor distribution consistently exhibits a fat left tail, reflecting the market's pricing of elevated systemic stress risk. The multi-stage bandwidth annealing procedure, with basis change at stage transitions, is essential for reliable convergence.

The techniques developed here—nonparametric factor distributions, monotonic interpolation, bump-basis parameterisation, focused grids—are not specific to CDO tranches. They apply to any setting where a one-factor model of correlated defaults must be calibrated to observed prices: CLO tranches, mortgage portfolios, bank loan books. The broader principle is that nonparametric distributional flexibility, combined with modern optimisation, can solve calibration problems that seemed intractable with the computational resources and numerical methods available twenty years ago.

A reference implementation, including the code used to produce the figures in this paper, is available at `https://github.com/edparcell/numerical-cdo-copula-paper`.

AI Disclosure. The core methodological idea—using a nonparametric market factor distribution to fit the CDO correlation smile within the standard one-factor copula framework—was developed by the author circa 2006, and an implementation and partial draft were created at that time. This paper was completed in 2026 with substantial assistance from Claude (Anthropic), which was used to expand the exposition, implement the numerical methods, produce the figures, and survey related literature. The author provided the mathematical framework, designed key aspects of the parameterisation and calibration approach, and reviewed all technical content. The choice to publish with AI assistance, rather than leave the idea incomplete indefinitely, was a deliberate one.

References

[1] L. Andersen, J. Sidenius, and S. Basu. All your hedges in one basket. *Risk*, pages 67–72, November 2003.

[2] L. Andersen and J. Sidenius. Extensions to the Gaussian copula: random recovery and random factor loadings. *Journal of Credit Risk*, 1(1), 2004/05.

[3] D. Brigo, A. Pallavicini, and R. Torresetti. Calibration of CDO tranches with the dynamical generalized-Poisson loss model. *Risk*, May 2007.

[4] X. Burtschell, J. Gregory, and J.-P. Laurent. A comparative analysis of CDO pricing models under the factor copula framework. *Journal of Derivatives*, 16(4):9–37, 2009.

[5] J. Garcia and S. Goossens. Base expected loss explains Lévy base correlation smile. Working paper, SSRN 1305398, 2007.

[6] N. Hale and A. Townsend. An algorithm for the convolution of Legendre series. *SIAM Journal on Scientific Computing*, 36(3):A1207–A1220, 2014.

[7] J. Gregory and J.-P. Laurent. I will survive. *Risk*, pages 103–107, June 2003.

[8] P. Hagan, D. Kumar, A. Lesniewski, and D. Woodward. Managing smile risk. *Wilmott Magazine*, pages 84–108, September 2002.

[9] J. Hull and A. White. Valuing credit derivatives using an implied copula approach. *Journal of Derivatives*, Fall 2006.

[10] J. Hull and A. White. Valuation of a CDO and an nth to default CDS without Monte Carlo simulation. *Journal of Derivatives*, 12(2):8–23, 2004.

[11] A. Kalemanova, B. Schmid, and R. Werner. The normal inverse Gaussian distribution for synthetic CDO pricing. *Journal of Derivatives*, 14(3):80–93, 2007.

[12] E. Parcell. Loss unit interpolation in the collateralized debt obligation pricing model. Working paper, 2006.

[13] E. Parcell and J. Wood. Wiping the smile off your base (correlation curve). Working paper, 2007.

[14] R. Storn and K. Price. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359, 1997.

[15] M. Steffen. A simple method for monotonic interpolation in one dimension. *Astronomy and Astrophysics*, 239:443–450, 1990.

[16] L. N. Trefethen. *Approximation Theory and Approximation Practice*. SIAM, Philadelphia, 2013.

Appendix A. Calibration results: August 30, 2005

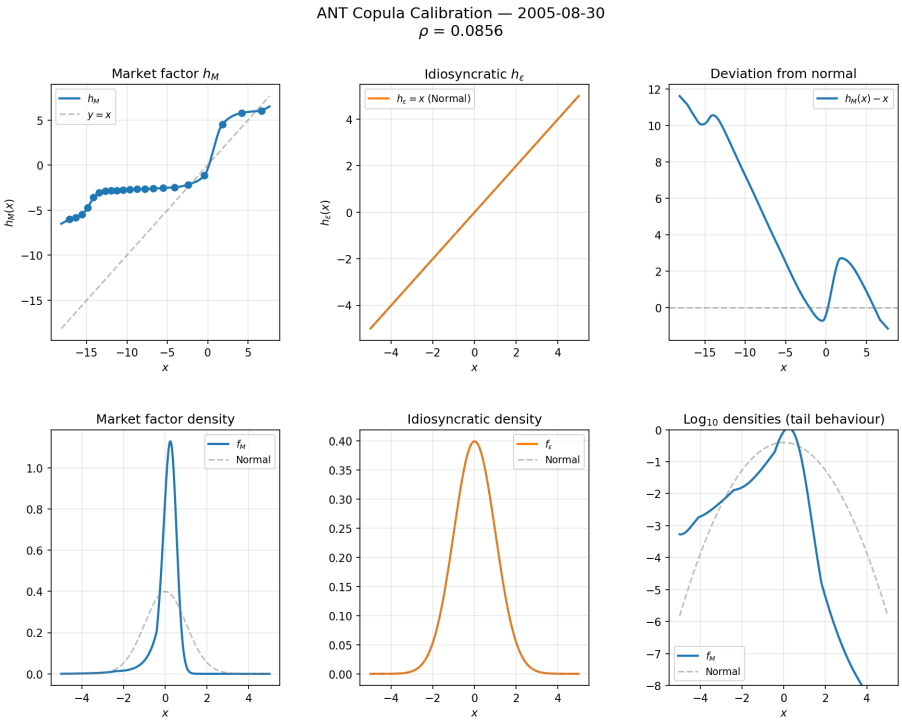

Figure 1. ANT copula calibration for CDX IG 5Y, August 30, 2005. Top row: calibrated ANT function h_M, idiosyncratic function (fixed as identity), and deviation $h_M(x) - x$. Bottom row: market factor density, idiosyncratic density (standard normal), and log-scale densities showing tail behaviour. $\rho = 0.086$.

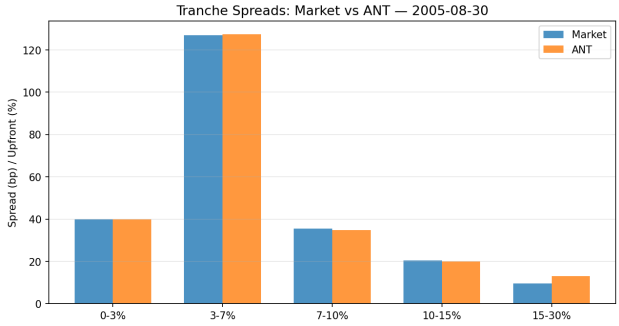

Figure 2. Tranche spread comparison: market vs ANT copula model, August 30, 2005.

Appendix B. Calibration results: August 4, 2004

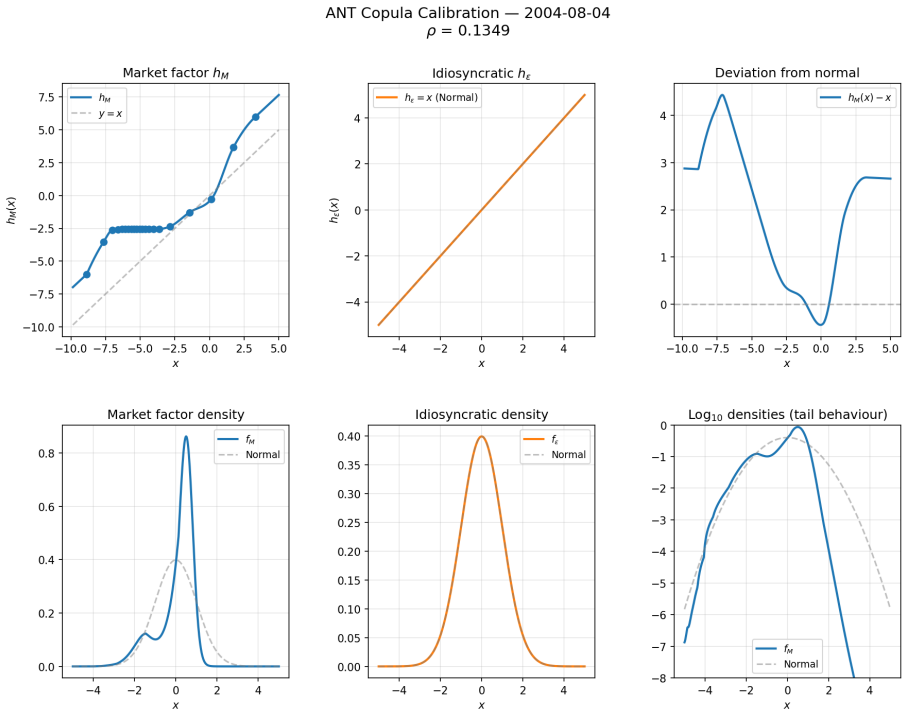

Figure 3. ANT copula calibration for CDX IG 5Y, August 4, 2004. Top row: calibrated ANT function h_M, idiosyncratic function (fixed as identity), and deviation $h_M(x) - x$. Bottom row: market factor density, idiosyncratic density (standard normal), and log-scale densities showing tail behaviour. $\rho = 0.135$.

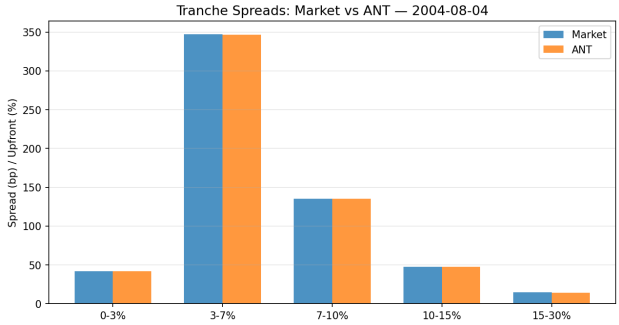

Figure 4. Tranche spread comparison: market vs ANT copula model, August 4, 2004.

Appendix C. Base correlation curves

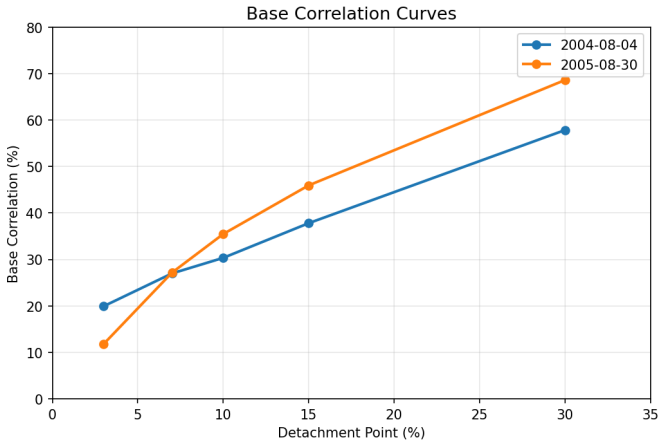

Figure 5. Bootstrapped base correlation curves for CDX IG 5Y on August 4, 2004 and August 30, 2005.

Appendix D. Numerical convergence of F_A

The distribution function F_A is evaluated on a uniform grid and interpolated using Steffen's monotonic cubic method (Section 3.3). Figure 6 shows convergence as the number of grid points increases, using the calibrated ANT copula for August 30, 2005.

The left panel measures the supremum interpolation error $\sup_x |F_A(x) - F_A^{\text{ref}}(x)|$, where the reference is computed on a grid of 50,001 points. The right panel measures the maximum error across all five tranche fair spreads. Both converge as $O(h^4)$, consistent with Steffen's piecewise cubic interpolation. At the 2001 grid points used in this paper, the interpolation error is $\sim 2 \times 10^{-8}$ and the maximum tranche pricing error is ~ 0.002 bp—well below the ~ 1 bp fitting tolerance.

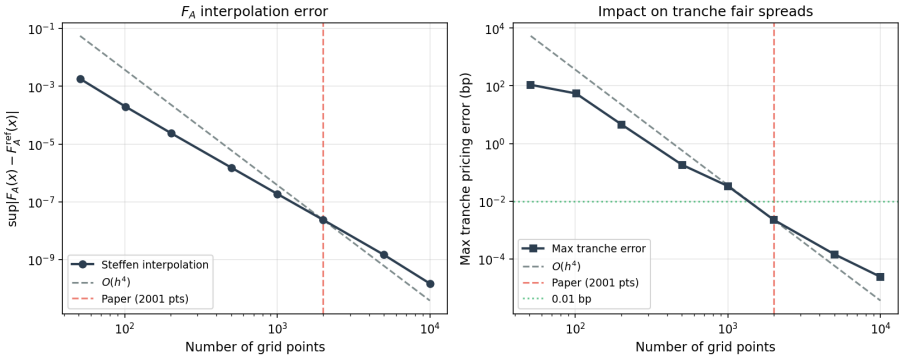

Figure 6. Convergence of F_A grid discretisation. Left: supremum interpolation error vs. reference. Right: maximum tranche fair spread error. Dashed grey line shows $O(h^4)$ rate; vertical red line marks the 2001 grid points used in this paper.

Appendix E. Gauss–Hermite quadrature convergence

The convolution integral (16) is evaluated by Gauss–Hermite quadrature after the change of variable described in Section 3.5. For the Gaussian copula (where h_M is the identity), the integrand after substitution is entire and Gauss–Hermite converges spectrally. For the ANT copula, the substitution introduces h_M^{-1}, whose derivatives may grow rapidly in the tails—particularly where h_M has steep gradients.

Figure 7 shows convergence as the quadrature order increases, with the reference computed by trapezoidal integration on 100,000 points. The F_A grid is fixed at 2001 points throughout. Two features are notable:

(1) Convergence is *algebraic*, not exponential: the supremum F_A error decreases roughly as $O(n^{-1})$, far slower than the spectral convergence expected for smooth integrands.
(2) The tranche pricing errors are *non-monotone*, with local increases at orders around 100 and 300. At high quadrature orders the outermost nodes sample extreme quantiles of the ANT distribution, where h_M^{-1} relies on linear extrapolation beyond the knot boundaries, introducing oscillatory artefacts.

Crucially, this does not invalidate the calibration results. The optimiser searches for h_M and ρ that price all tranches correctly *under the chosen quadrature rule*. The calibrated parameters are therefore self-consistent at 100 points: any systematic bias in F_A relative to the exact integral is absorbed into the calibrated h_M and ρ. A recalibration at a different quadrature order would yield slightly different parameters but comparably good tranche fits. In this sense the calibrated ANT function should be understood as "the distribution that, convolved via 100-point Gauss–Hermite quadrature, reproduces all market tranche prices" rather than a pointwise-converged estimate of the true market factor distribution.

The 100-point quadrature is a practical choice that balances accuracy in the bulk of the integral against the risk of tail artefacts at higher orders. The grid discretisation of F_A (Appendix D), which converges cleanly as $O(h^4)$, is the binding source of numerical error.

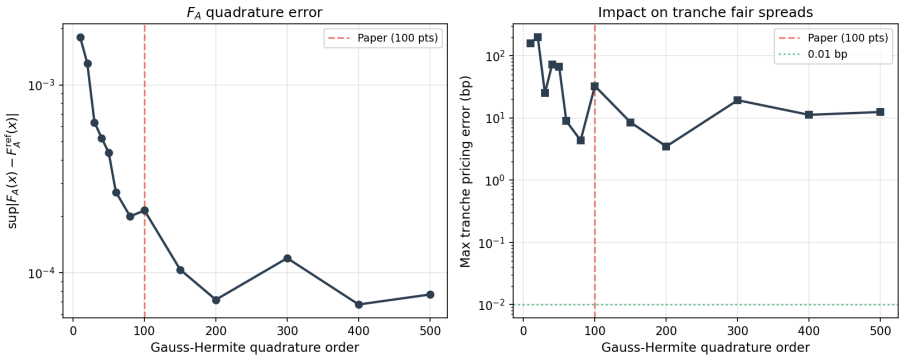

FIGURE 7. Gauss–Hermite quadrature convergence for the calibrated ANT copula (August 30, 2005). Left: supremum F_A error vs. trapezoidal reference. Right: maximum tranche fair spread error. Convergence is algebraic and non-monotone, reflecting the non-smoothness of h_M^{-1} in the integrand. Vertical red line marks the 100-point quadrature used in this paper.