LECTURE 2

- **LECTURE 2**
- The swap market model;
- Change of Numeraire: rigorous derivation of Black's market formula for swaptions;
- Incompatibility between LIBOR and SWAP models;
- Parameterizing the LIBOR model: Instantaneous volatilities:
- Diagnostics after calibration: Term structure of caplet volatilitites and Terminal correlations;
- Instantaneous correlations: Some full rank parameterizations.
- Instantaneous correlations: Reduced rank parameterizations.
- Monte Carlo pricing with the LIBOR model
- An approximated swaption formula linking the LIBOR model to the swaption market restoring mutual compatibility in practice
- Derivation of the formula for terminal correlation;

The Swap Market Model

Similarly to caplets, Black's formula for **swaptions** becomes rigorous by taking as numeraire the annuity, or the PVPBP:

$$U = C_{\alpha,\beta}(t) = \sum_{i=\alpha+1}^{\beta} \tau_i P(t, T_i), \quad Q^U = Q^{\alpha,\beta}$$

By FACT ONE the forward swap rate $S_{\alpha,\beta}$ is then a martingale under $Q^{\alpha,\beta}$:

$$S_{\alpha,\beta}(t) = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{i=\alpha+1}^{\beta} \tau_i P(t,T_i)} = \frac{P(t,T_\alpha) - P(t,T_\beta)}{C_{\alpha,\beta}(t)}$$

Take the usual martingale (zero drift) lognormal geometric brownian motion

$$d\ S_{\alpha,\beta}(t) = \sigma^{(\alpha,\beta)}(t) S_{\alpha,\beta}(t)\ dW_t^{\alpha,\beta}, \quad Q^{\alpha,\beta} \quad \textbf{(LSM)},$$

The Swap Market Model

BY FACT TWO on the change of numeraire

$$E^{\boxed{B}}\left((S_{\alpha,\beta}(T_\alpha)-K)^+\, C_{\alpha,\beta}(T_\alpha)\frac{\boxed{B(0)}}{\boxed{B(T_\alpha)}}\right) =$$

$$= \ E^{\boxed{\alpha,\beta}}\left[(S_{\alpha,\beta}(T_\alpha)-K)^+ C_{\alpha,\beta}(T_\alpha)\frac{\boxed{C_{\alpha,\beta}(0)}}{\boxed{C_{\alpha,\beta}(T_\alpha)}}\right]$$

$$= C_{\alpha,\beta}(0)\ E^{\alpha,\beta}\left[(S_{\alpha,\beta}(T_\alpha)-K)^+\right]$$

$$= C_{\alpha,\beta}(0)\,[S_{\alpha,\beta}(0)\Phi\,(d_1)-K\Phi(d_2)],\ \ d_{1,2} = \frac{\ln\frac{S_{\alpha,\beta}(0)}{K}\pm\frac{1}{2}T_\alpha v^2_{\alpha,\beta}(T_\alpha)}{\sqrt{T_\alpha}v_{\alpha,\beta}(T_\alpha)}$$

$$= C_{\alpha,\beta}(0)\ \mathsf{B\&S}(S_{\alpha,\beta}(0),K,\sqrt{T_\alpha}v_{\alpha,\beta}(T_\alpha)),$$

$$v^2_{\alpha,\beta}(T) = \frac{1}{T}\int_0^T (\sigma^{(\alpha,\beta)}(t))^2\ dt\ .$$

Theoretical incompatibility LSM / LFM

Recall **LFM**: $dF_i(t) = \sigma_i(t)F_i(t)dZ_i(t), \quad Q^i,$

$$\textbf{LSM}: d\ S_{\alpha,\beta}(t) = \sigma^{(\alpha,\beta)}(t)S_{\alpha,\beta}(t)\ dW_t, \quad Q^{\alpha,\beta}\ . \qquad (1)$$

Precisely: Can each F_i be lognormal under Q^i **and** $S_{\alpha,\beta}$ be lognormal under $Q^{\alpha,\beta}$, given that

$$S_{\alpha,\beta}(t) = \frac{1 - \prod_{j=\alpha+1}^{\beta} \frac{1}{1+\tau_j\ F_j(t)}}{\sum_{i=\alpha+1}^{\beta} \tau_i \prod_{j=\alpha+1}^{i} \frac{1}{1+\tau_j\ F_j(t)}} \quad ? \qquad (2)$$

Check distributions of $S_{\alpha,\beta}$ under $Q^{\alpha,\beta}$ for both LFM and LSM. Derive the LFM model under the LSM numeraire $Q^{\alpha,\beta}$:

$$dF_k(t) = \sigma_k(t)F_k(t)\left(\mu_k^{\alpha,\beta}(t)dt + dZ_k^{\alpha,\beta}(t)\right), \qquad (3)$$

$$\mu_k^{\alpha,\beta} = \sum_{j=\alpha+1}^{\beta} (2_{(j\le k)} - 1)\tau_j \frac{P(t,T_j)}{C_{\alpha,\beta}(t)} \sum_{i=\min(k+1,j+1)}^{\max(k,j)} \frac{\tau_i \rho_{k,i}\sigma_i F_i}{1+\tau_i F_i}.$$

When computing the swaption price as the $Q^{\alpha,\beta}$ expectation

$$C_{\alpha,\beta}(0)E^{\alpha,\beta}(S_{\alpha,\beta}(T_\alpha) - K)^+$$

we can use either LFM (2,3) or LSM (1).

In general, $S_{\alpha,\beta}$ coming from LSM (1) is LOGNORMAL, whereas $S_{\alpha,\beta}$ coming from LFM (2,3) is NOT. But in practice...

LFM instantaneous covariance structures

LFM is natural for caps and LSM is natural for swaptions. **Choose.** We choose LFM and adapt it to price swaptions.

Recall: Under numeraire $P(\cdot, T_i) \neq P(\cdot, T_k)$:

$$dF_k(t) = \mu_k^i(t)\; F_k(t)\; dt + \boxed{\sigma_k(t)} F_k(t)\, dZ_k, \quad dZ\; dZ' = \boxed{\rho}\; dt$$

Model specification: Choice of $\sigma_k(t)$ and of ρ.

- General Piecewise constant **(GPC)** vols, $\sigma_k(t) = \sigma_{k,\beta(t)}$, $T_{\beta(t)-2} < t \leq T_{\beta(t)-1}$.

Inst. Vols	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	$\ldots$	$(T_{M-2}, T_{M-1}]$
Fwd: $F_1(t)$	$\sigma_{1,1}$	Expired	Expired	$\ldots$	Expired
$F_2(t)$	$\sigma_{2,1}$	$\sigma_{2,2}$	Expired	$\ldots$	Expired
$\vdots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
$F_M(t)$	$\sigma_{M,1}$	$\sigma_{M,2}$	$\sigma_{M,3}$	$\ldots$	$\sigma_{M,M}$

Separable Piecewise const **(SPC)**, $\sigma_k(t) = \Phi_k \psi_{k-(\beta(t)-1)}$

- Parametric Linear-Exponential **(LE)** vols

$$\sigma_i(t) = \Phi_i\, \psi(T_{i-1} - t; a, b, c, d)$$
$$:= \Phi_i \left([a(T_{i-1} - t) + d] e^{-b(T_{i-1}-t)} + c \right).$$

Caplet volatilities

Recall that under numeraire $P(\cdot, T_i)$:

$$dF_i(t) = \sigma_i(t) F_i(t)\, dZ_i, \quad dZ dZ' = \rho\, dt$$

Caplet: Strike rate K, Reset T_{i-1}, Payment T_i:

Payoff: $\tau_i (F_i(T_{i-1}) - K)^+$ at T_i.

"Call option" on F_i, $F_i \sim$ lognormal under Q^i

$\Rightarrow$ Black's formula, with Black vol. parameter

$$v^2_{T_{i-1}-\text{caplet}} := \frac{1}{T_{i-1}} \int_0^{T_{i-1}} \sigma_i(t)^2 dt.$$

$v_{T_{i-1}-\text{caplet}}$ is T_{i-1}-*caplet volatility*

Only the σ's have impact on caplet (and cap) prices, the ρ's having no influence.

Caplet volatilities (cont'd)

$$dF_i(t) = \sigma_i(t) F_i(t)\, dZ_i, \qquad v^2_{T_{i-1}-\text{caplet}} := \frac{1}{T_{i-1}} \int_0^{T_{i-1}} \sigma_i(t)^2 dt.$$

Under GPC vols, $\sigma_k(t) = \sigma_{k,\beta(t)}$

$$v^2_{T_{i-1}-\text{caplet}} = \frac{1}{T_{i-1}} \sum_{j=1}^{i} (T_{j-1} - T_{j-2})\ \sigma^2_{i,j}$$

Under LE vols, $\sigma_i(t) = \Phi_i\, \psi(T_{i-1} - t; a, b, c, d)$,

$$T_{i-1} v^2_{T_{i-1}-\text{caplet}} = \Phi_i^2\ I^2(T_{i-1}; a, b, c, d)$$

$$:= \Phi_i^2 \int_0^{T_{i-1}} \left([a(T_{i-1} - t) + d] e^{-b(T_{i-1}-t)} + c \right)^2\ dt \quad .$$

Term Structure of Caplet Volatilities

The term structure of volatility at time T_j is a graph of expiry times T_{h-1} against average volatilities $V(T_j, T_{h-1})$ of the related forward rates $F_h(t)$ up to that expiry time itself, i.e. for $t \in (T_j, T_{h-1})$.

Formally, at time $t = T_j$, graph of points

$$\{(T_{j+1}, V(T_j, T_{j+1})), (T_{j+2}, V(T_j, T_{j+2})), \ldots, (T_{M-1}, V(T_j, T_{M-1}))\}$$

$$V^2(T_j, T_{h-1}) = \frac{1}{T_{h-1} - T_j} \int_{T_j}^{T_{h-1}} \sigma_h^2(t)dt, \quad h > j+1.$$

The term structure of vols at time 0 is given simply by caplets vols plotted against their expiries.

Different assumptions on the behaviour of instantaneous volatilities (SPC, LE, etc.) imply different evolutions for the term structure of volatilities in time as $t = T_0$, $t = T_1$, $t = T_2$...

Cap calibration: Some possible choices

We implemented a version with:

- Semi-annual tenors, $T_i - T_{i-1} = 6m$.
- Instantaneous correlation estimated historically, first fitted on the full rank parametric form in ρ_∞, α:

$$\rho_\infty + (1 - \rho_\infty)\exp(-\alpha|i - j|)$$

and then possibly fitted to a reduced rank correlation (no impact on caps but need for ratchets etc., more on this later)

- Vol. parameterization $\sigma_k(t) = \sigma_{k,\beta(t)} := \Phi_k \psi_{k-(\beta(t)-1)}$,

Inst. Vols	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	...	$(T_{M-2}, T_{M-1}]$
Fwd : $F_1(t)$	$\Phi_1\psi_1$	Dead	Dead	...	Dead
$F_2(t)$	$\Phi_2\psi_2$	$\Phi_2\psi_1$	Dead	...	Dead
$\vdots$	...	...	...	...	...
$F_M(t)$	$\Phi_M\psi_M$	$\Phi_M\psi_{M-1}$	$\Phi_M\psi_{M-2}$	...	$\Phi_M\psi_1$

Note: $\Phi = 1$ (use only ψ) leads to "stationary vol term structure" as in the top figure, next page;
$\psi = 1$ (use only Φ) leads to constant volatilities and is the easiest calibration possible, since then $\Phi_i = v_{T_{i-1}-\text{caplet}}$, but leads also to bad term-structure evolution, middle figure next page.

Cap calibration: Some possible choices

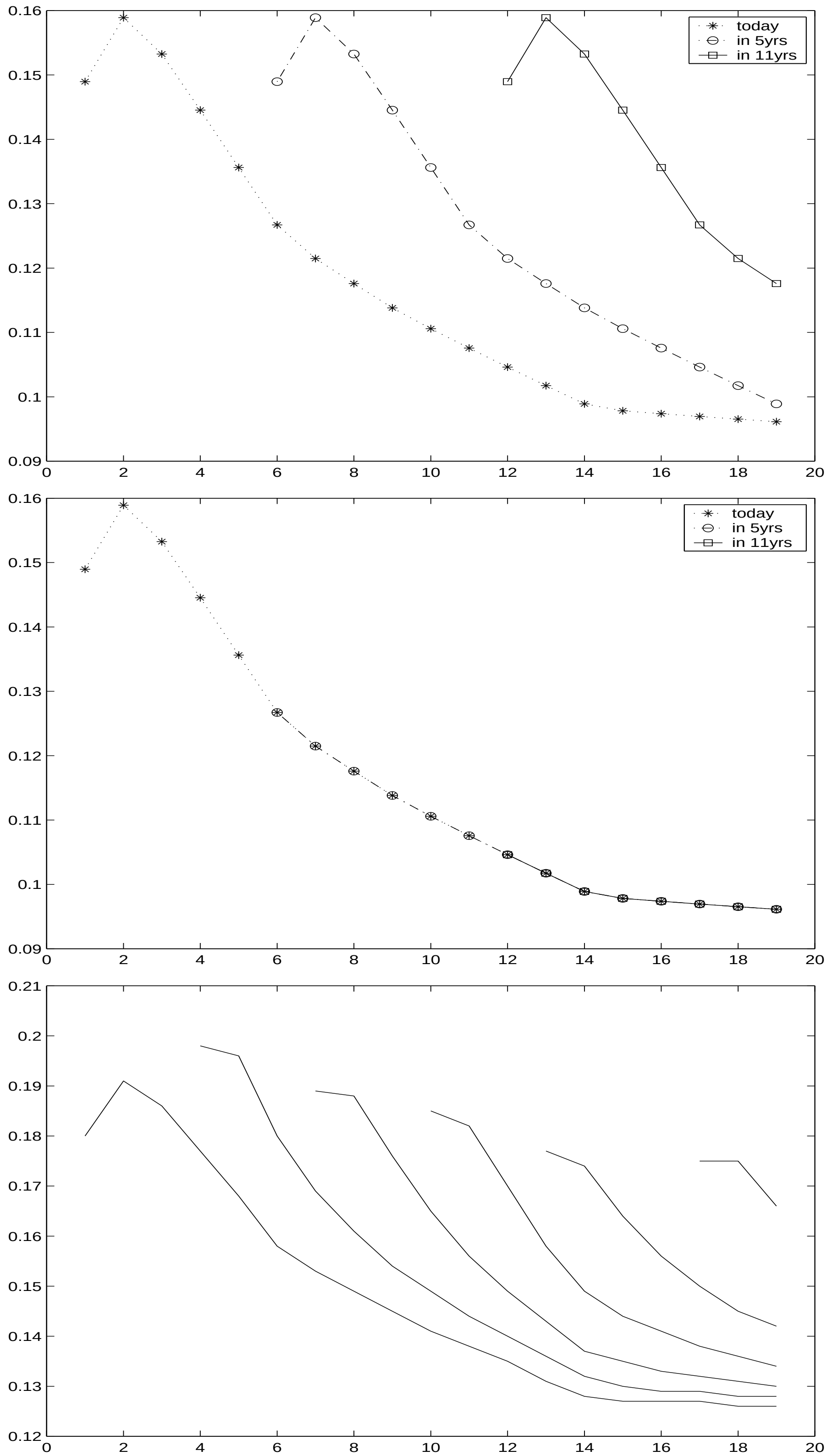
Vol. parameterization: HOMOGENEOUS IN THE TIME-TO-EXPIRY (constancy along the DIAGONALS of the "ziggurat"): $\sigma_k(t) = \psi_{k-(\beta(t)-1)}$, and in particular $\sigma_k(T_j-) = \psi_{k-j}$;

Inst. Vols	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	...	$(T_{M-2}, T_{M-1}]$
Fwd : $F_1(t)$	ψ_1	Dead	Dead	...	Dead
$F_2(t)$	ψ_2	ψ_1	Dead	...	Dead
$\vdots$	...	...	...	...	...
$F_M(t)$	ψ_M	ψ_{M-1}	ψ_{M-2}	...	ψ_1

Vol. parameterization: HOMOGENEOUS IN TIME (constancy along the ROWS of the "ziggurat"): $\sigma_k(t) = \Phi_k$

Inst. Vols	$t \in (0, T_0]$	$(T_0, T_1]$	$(T_1, T_2]$	...	$(T_{M-2}, T_{M-1}]$
Fwd : $F_1(t)$	Φ_1	Dead	Dead	...	Dead
$F_2(t)$	Φ_2	Φ_2	Dead	...	Dead
$\vdots$	...	...	...	...	...
$F_M(t)$	Φ_M	Φ_M	Φ_M	...	Φ_M

Let's see the evolution of the term structure of volatilities in the three cases: $\Phi = 1$ (homogeneous in time-to-expiry), $\psi = 1$ (homogeneous in time), and intermediate (neither Φ nor ψ set to one).


today
in 5yrs
in 11yrs
today
in 5yrs
in 11yrs

Terminal and Instantaneous correlation

Swaptions depend on **terminal** correlations among fwd rates.

E.g., the swaption whose underlying is $S_{1,3}$ depends on

$$\mathrm{corr}(F_2(T_1), F_3(T_1)).$$

This terminal corr. depends both on **inst.** corr. $\rho_{2,3}$

and and on the way the $T_1 - T_2$ and $T_2 - T_3$ caplet vols

$v_1 = v_{T_1-\text{caplet}}$ and $v_2 = v_{T_2-\text{caplet}}$ are decomposed in instantaneous vols $\sigma_2(t)$ and $\sigma_3(t)$ for t in $0, T_1$.

$$\mathrm{corr}(F_2(T_1), F_3(T_1)) \approx \frac{\int_0^{T_1} \sigma_2(t)\sigma_3(t)\rho_{2,3} dt}{\sqrt{\int_0^{T_1} \sigma_2^2(t)dt}\ \sqrt{\int_0^{T_1} \sigma_3^2(t)dt}} =$$

$$= \rho_{2,3} \frac{\sigma_{2,1}\sigma_{3,1} + \sigma_{2,2}\sigma_{3,2}}{v_2\ \sqrt{\sigma_{3,1}^2 + \sigma_{3,2}^2}}\ .$$

No such formula for general short-rate models

$$\mathrm{corr}(F_2(T_1), F_3(T_1)) \approx \rho_{2,3} \frac{\sigma_{2,1}\sigma_{3,1} + \sigma_{2,2}\sigma_{3,2}}{v_1\sqrt{T_1}\ \sqrt{\sigma_{3,1}^2 + \sigma_{3,2}^2}} \ .$$

Fix $\rho_{2,3} = 1$, $\tau_i = 1$ and caplet vols:

$$v_1^2 T_1 = \sigma_{2,1}^2 + \sigma_{2,2}^2; \qquad v_2^2 T_2 = \sigma_{3,1}^2 + \sigma_{3,2}^2 + \sigma_{3,3}^2 \ .$$

Decompose v_1 and v_2 in two different ways: **First case**

$$\sigma_{2,1} = v_1\sqrt{T_1},\ \sigma_{2,2} = 0; \qquad \sigma_{3,1} = v_2\sqrt{T_2}, \sigma_{3,2} = \sigma_{3,3} = 0.$$

In this case the above fomula yields easily

$$\mathrm{corr}(F_2(T_1), F_3(T_1)) = \rho_{2,3} = 1 \ .$$

The **second case** is obtained as

$$\sigma_{2,1} = 0,\ \sigma_{2,2} = v_1\sqrt{T_1}; \qquad \sigma_{3,1} = v_2\sqrt{T_2}, \sigma_{3,2} = \sigma_{3,3} = 0.$$

In this second case the above fomula yields immediately

$$\mathrm{corr}(F_2(T_1), F_3(T_1)) = 0\rho_{2,3} = 0 \ .$$

Instantaneous correlation: Parametric forms

Swaptions depend on **terminal** correlation among forward rates (ρ's **and** σ's). How do we model ρ?

Full Rank Parametric forms for instant. correl. ρ

Schoenmakers and Coffey (2000) propose a finite sequence

$$1 = c_1 < c_2 < \ldots < c_M, \quad \frac{c_1}{c_2} < \frac{c_2}{c_3} < \ldots < \frac{c_{M-1}}{c_M},$$

and they set ("F" stands for "Full" (Rank))

$$\rho^F(c)_{i,j} := c_i/c_j, \quad i \leq j, \quad i, j = 1, \ldots, M.$$

Notice that the correlation between changes in adjacent rates is $\rho^F_{i+1,i} = c_i/c_{i+1}$. The above requirements on c's translate into the requirement that

the sub-diagonal of the resulting correlation matrix $\rho^F(c)$ be increasing when moving from NW to SE.

This bears the interpretation that when we move along the yield curve, the larger the tenor, the more correlated changes in adjacent forward rates become. This corresponds to the experienced fact that the forward curve tends to flatten and to move in a more "correlated" way for large maturities than for small ones. This holds also for lower levels below the diagonal.

Instantaneous correlation: Parametric forms

Full rank: S& C parametric form (cont'd).

The number of parameters needed in this formulation is M, versus the $M(M-1)/2$ number of entries in the general correlation matrix. One can prove that $\rho^F(c)$ is always a viable correlation matrix if defined as above (symmetric, positive semidefinite and with ones in the diagonal).

Schoenmakers and Coffey (2000) observe also that this parameterization can be always characterized in terms of a finite sequence of non-negative numbers $\Delta_2, \ldots, \Delta_M$:

$$c_i = \exp\left[\sum_{j=2}^{i} j\Delta_j + \sum_{j=i+1}^{M} (i-1)\Delta_j\right].$$

Some particular cases in this class of parameterizations that Schoenmakers and Coffey (2000) consider to be promising can be formulated through suitable changes of variables as follows. The first is the case where all Δ's are zero except the last two: by a change of variable one has

Instantaneous correlation: Parametric forms

Stable, full rank, two-parameters, "increasing along sub-diagonals" parameterization for instantaneous correlation:

$$\rho_{i,j} = \exp\left[-\frac{|i-j|}{M-1}\left(-\ln\rho_\infty + \eta\frac{M-1-i-j}{M-2}\right)\right].$$

Stability here is meant to point out that relatively small movements in the c-parameters connected to this form cause relatively small changes in ρ_∞ and η.

Notice that $\rho_\infty = \rho_{1,M}$ is the correlation between the farthest forward rates in the family considered, whereas η is related to the first non-zero Δ, i.e. $\eta = \Delta_{M-1}(M-1)(M-2)/2$.

A 3-parameters form is obtained with Δ_i's following a straight line (two parameters) for $i = 2, 3, \ldots, M-1$ and set to a third parameter for $i = M$.
Stable, full rank, 3-parameters, "increasing along sub-diagonals" parameterization S&C3:

$$\rho_{i,j} = \exp\Big[-|i-j|\Big(\beta - \frac{\alpha_2}{6M-18}\left(i^2+j^2+ij-6i-6j-3M^2+15M-7\right) + \frac{\alpha_1}{6M-18}\left(i^2+j^2+ij-3Mi-3Mj+3i+3j+3M^2-6M+2\right)\Big)\Big]. \quad (4)$$

where the parameters should be constrained to be non-negative, if one wants to be sure all the typical desirable properties are indeed present.

Instantaneous correlation: Parametric forms

Full rank: S& C parametric form (cont'd).

In order to get parameter stability, Schoenmakers and Coffey introduce a change of variables, thus obtaining a laborious expression generalizing the earlier two-parameters one. The calibration experiments pointed out, however, that the parameter associated with the final point Δ_{M-1} of our straight line in the Δ's is practically always close to zero. Setting thus $\Delta_{M-1} = 0$ and maintaining the other characteristics of the last parameterization leads to the following

Improved, stable, full rank, two-parameters, "increasing along sub-diagonals" parameterization for instantaneous correlations (S&C2):

$$\rho_{i,j} = \exp\Bigg[- \frac{|i-j|}{M-1}\Bigg(-\ln\rho_\infty + \eta \frac{i^2+j^2+ij-3Mi-3Mj+3i+3j+2M^2-M-4}{(M-2)(M-3)}\Bigg)\Bigg]. \tag{5}$$

As before, $\rho_\infty = \rho_{1,M}$, whereas η is related to the steepness of the straight line in the Δ's.

Instantaneous correlation: Parametric forms

Full rank: Classic exp and Rebonato parametric forms.

Classical, two-parameters, exponentially decreasing parameterization

$$\rho_{i,j} = \rho_\infty + (1 - \rho_\infty) \exp[-\beta |i - j|], \quad \beta \geq 0.$$

where now ρ_∞ is only asymptotically representing the correlation between the farthest rates in the family.
Schoenmakers and Coffey (2000) point out that Rebonato's (1999c,d) full-rank parameterization, consisting in the following perturbation of the classical structure:

Rebonato's three parameters full rank parameterization

$$\rho_{i,j} = \rho_\infty + (1 - \rho_\infty) \exp[-|i - j| (\beta - \alpha (\max(i, j) - 1))], \tag{6}$$

has still the desirable property of being increasing along sub-diagonals. However, the domain of positivity for the resulting matrix is not specified "off-line" in terms of $\alpha, \beta, \rho_\infty$.

Instantaneous correlation: Reducing the rank

Instant. correl: Approximate ρ ($M \times M$, Rank M) with a n-rank $\rho^B = B \times B'$, with B an $M \times n$ matrix, $n << M$.

$$dZ\ dZ' = \rho\ dt \quad \longrightarrow \quad B\ dW(B\ dW)' = BB'dt\ .$$

$\rho^B = B \times B'$, with B an $M \times n$ matrix, $n << M$.

Eigenvalues zeroing and rescaling.

We know that, being ρ a positive definite symmetric matrix, it can be written as

$$\rho = PHP',$$

where P is a real orthogonal matrix, $P'P = PP' = I_M$, and H is a diagonal matrix of the positive eigenvalues of ρ.
The columns of P are the eigenvectors of ρ, associated to the eigenvalues located in the corresponding position in H.
Let Λ be the diagonal matrix whose entries are the square roots of the corresponding entries of H, so that if we set $A := P\Lambda$ we have both

$$AA' = \rho, \quad A'A = H\ .$$

Instantaneous correlation: Reducing the rank. Eigenvalues zeroing and rescaling (continued)

$$\rho = PHP', \ \text{“}\Lambda = \sqrt{H}\text{”}, \ A := P\Lambda, \ AA' = \rho, \ A'A = H.$$

We can try and mimic the decomposition $\rho = AA'$ by means of a suitable n-rank $M \times n$ matrix B such that BB' is an n-rank correlation matrix, with typically $n << M$.

Consider the diagonal matrix $\bar{\Lambda}^{(n)}$ defined as the matrix Λ with the $M - n$ smallest diagonal terms set to zero.

Define then $\bar{B}^{(n)} := P\bar{\Lambda}^{(n)}$, and the related candidate correlation matrix $\bar{\rho}^{(n)} := \bar{B}^{(n)}(\bar{B}^{(n)})'$.
We can also equivalently define $\Lambda^{(n)}$ as the $n \times n$ diagonal matrix obtained from Λ by taking away (instead of zeroing) the $M - n$ smallest diagonal elements and shrinking the matrix correspondingly. Analogously, we can define the $M \times n$ matrix $P^{(n)}$ as the matrix P from which we take away the columns corresponding to the diagonal elements we took away from Λ. The result does not change, in that if we define the $M \times n$ matrix $B^{(n)} = P^{(n)}\Lambda^{(n)}$ we have

$$\bar{\rho}^{(n)} = \bar{B}^{(n)}(\bar{B}^{(n)})' = B^{(n)}(B^{(n)})'.$$

We keep the $B^{(n)}$ formulation.

Instantaneous correlation: Reducing the rank. Eigenvalues zeroing and rescaling (continued)

$$\bar{\rho}^{(n)} = B^{(n)}(B^{(n)})', \quad B^{(n)} = P^{(n)}\Lambda^{(n)}.$$

Now the problem is that, in general, while $\bar{\rho}^{(n)}$ is positive semidefinite, it does not feature ones in the diagonal. Throwing away some eigenvalues from Λ has altered the diagonal. The solution is to interpret $\bar{\rho}^{(n)}$ as a *covariance* matrix, and to derive the correlation matrix associated with it. We can do this immediately by defining

$$\rho_{i,j}^{(n)} := \bar{\rho}_{i,j}^{(n)}/(\sqrt{\bar{\rho}_{i,i}^{(n)}\ \bar{\rho}_{j,j}^{(n)}}).$$

Now $\rho_{i,j}^{(n)}$ is an n-rank approximation of the original matrix ρ. But how good is the approximation, and are there more precise methods to approximate a full rank correlation matrix with a n-rank matrix? Can we find, in a sense, the **best** reduced rank correlation matrix approximating a given full rank one?

Instantaneous correlation: Reducing the rank. Angles parameterization and optimization

An angles parametric form for B. Rebonato:

$$
\begin{aligned}
b_{i,1} &= \cos\theta_{i,1} \\
b_{i,k} &= \cos\theta_{i,k}\sin\theta_{i,1}\cdots\sin\theta_{i,k-1}, \quad 1<k<n, \\
b_{i,n} &= \sin\theta_{i,1}\cdots\sin\theta_{i,n-1}, \quad \text{for} \;\; i=1,2,\ldots,M.
\end{aligned}
$$

Angles are redundant: one can assume with no loss of generality that $\theta_{i,k}=0$ for $i \le k$ ("trapezoidal" angles matrix)

For $\quad n=2, \quad \rho^B_{i,j} = b_{i,1}b_{j,1} + b_{i,2}b_{j,2} = \cos(\theta_i - \theta_j).$

(redendancy: can assume $\theta_1 = 0$ with no loss of generality.) This structure consists of M parameters $\theta_1,\ldots,\theta_M$ obtained either by forcing the LFM model to recover market swaptions prices (market implied data), or through historical estimation (time-series/econometrics). More on this later.
Given full rank ρ^F, can find optimal θ by minimizing numerically

$$
\theta^* = \text{argmin}_\theta \left(\sum_{i,j=1}^{M} (\rho^F_{i,j} - \rho_{i,j}(\theta))^2 \right).
$$

Example: full rank ρ

1	0.9756	0.9524	0.9304	0.9094	0.8894	0.8704	0.8523	0.8352	0.8188
0.9756	1	0.9756	0.9524	0.9304	0.9094	0.8894	0.8704	0.8523	0.8352
0.9524	0.9756	1	0.9756	0.9524	0.9304	0.9094	0.8894	0.8704	0.8523
0.9304	0.9524	0.9756	1	0.9756	0.9524	0.9304	0.9094	0.8894	0.8704
0.9094	0.9304	0.9524	0.9756	1	0.9756	0.9524	0.9304	0.9094	0.8894
0.8894	0.9094	0.9304	0.9524	0.9756	1	0.9756	0.9524	0.9304	0.9094
0.8704	0.8894	0.9094	0.9304	0.9524	0.9756	1	0.9756	0.9524	0.9304
0.8523	0.8704	0.8894	0.9094	0.9304	0.9524	0.9756	1	0.9756	0.9524
0.8352	0.8523	0.8704	0.8894	0.9094	0.9304	0.9524	0.9756	1	0.9756
0.8188	0.8352	0.8523	0.8704	0.8894	0.9094	0.9304	0.9524	0.9756	1

Rank-2 optimal approximation:

$$\theta^{*(2)} = [1.2367 \;\; 1.2812 \;\; 1.3319 \;\; 1.3961 \;\; 1.4947 \;\; 1.6469 \;\; 1.7455 \;\; 1.8097 \;\; 1.8604 \;\; 1.9049].$$

The resulting optimal rank-2 matrix $\rho(\theta^{*(2)})$ is

1	0.999	0.9955	0.9873	0.9669	0.917	0.8733	0.8403	0.8117	0.7849
0.999	1	0.9987	0.9934	0.9773	0.9339	0.8941	0.8636	0.8369	0.8117
0.9955	0.9987	1	0.9979	0.9868	0.9508	0.9157	0.888	0.8636	0.8403
0.9873	0.9934	0.9979	1	0.9951	0.9687	0.9396	0.9157	0.8941	0.8733
0.9669	0.9773	0.9868	0.9951	1	0.9885	0.9687	0.9508	0.9339	0.917
0.917	0.9339	0.9508	0.9687	0.9885	1	0.9951	0.9868	0.9773	0.9669
0.8733	0.8941	0.9157	0.9396	0.9687	0.9951	1	0.9979	0.9934	0.9873
0.8403	0.8636	0.888	0.9157	0.9508	0.9868	0.9979	1	0.9987	0.9955
0.8117	0.8369	0.8636	0.8941	0.9339	0.9773	0.9934	0.9987	1	0.999
0.7849	0.8117	0.8403	0.8733	0.917	0.9669	0.9873	0.9955	0.999	1

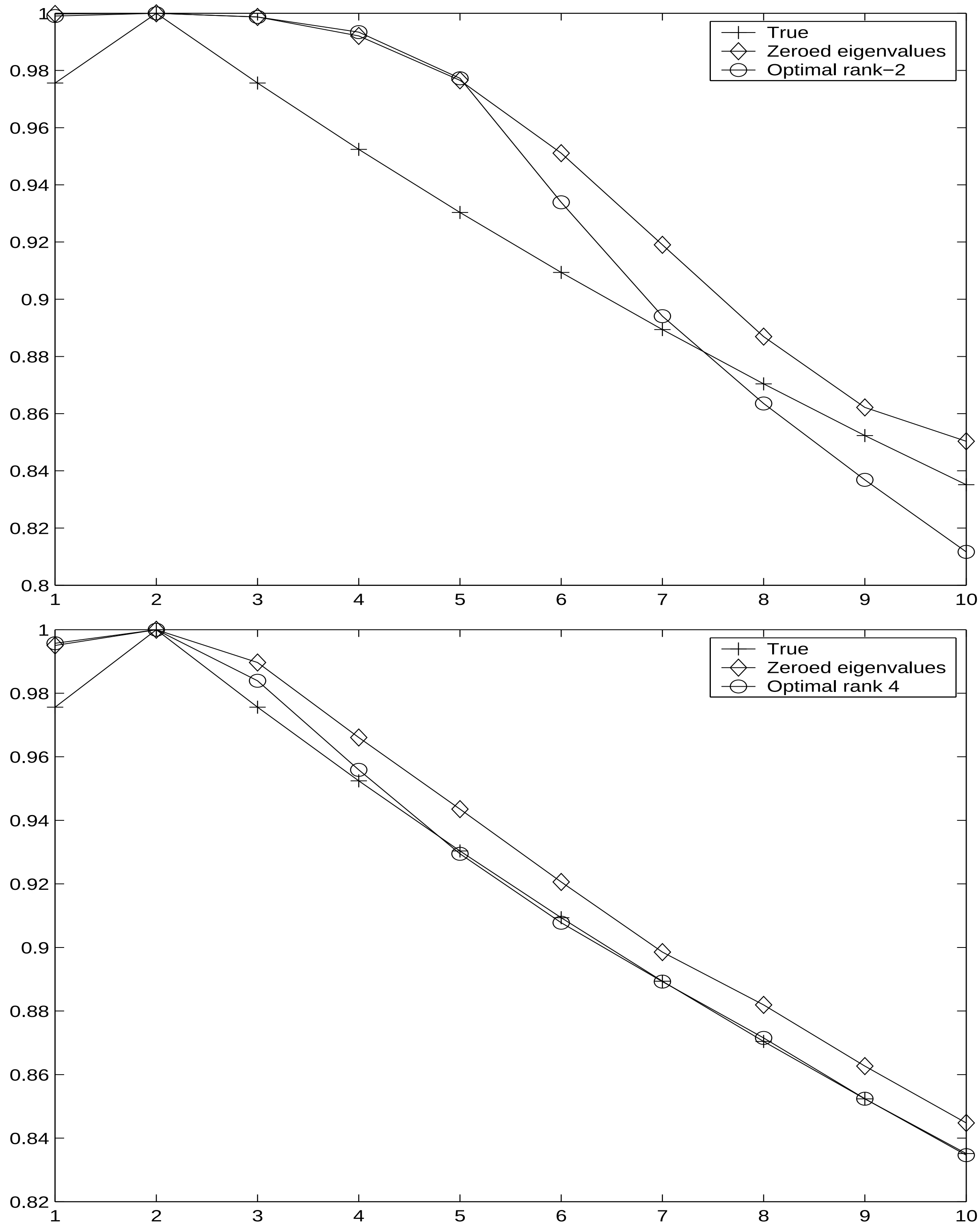


Figure 2: Problems of low rank correlation: sigmoid shape

Another example: Consider the rapidly decreasing 10×10 full-rank $\hat{\rho}_{i,j} = \exp[-|i-j|]$.
rank-4 approximation: the zeroed-eigenvalues procedure yields a matrix $\rho^{(4)}$ given by

1	0.9474	0.5343	-0.0116	-0.1967	-0.0427	0.1425	0.1378	-0.042	-0.1511
0.9474	1	0.775	0.2884	0.0164	-0.03	0.0316	0.0538	0	-0.042
0.5343	0.775	1	0.8137	0.4993	0.0979	-0.1229	-0.1035	0.0538	0.1378
-0.0116	0.2884	0.8137	1	0.8583	0.3725	-0.0336	-0.1229	0.0316	0.1425
-0.1967	0.0164	0.4993	0.8583	1	0.7658	0.3725	0.0979	-0.03	-0.0427
-0.0427	-0.03	0.0979	0.3725	0.7658	1	0.8583	0.4993	0.0164	-0.1967
0.1425	0.0316	-0.1229	-0.0336	0.3725	0.8583	1	0.8137	0.2884	-0.0116
0.1378	0.0538	-0.1035	-0.1229	0.0979	0.4993	0.8137	1	0.775	0.5343
-0.042	0	0.0538	0.0316	-0.03	0.0164	0.2884	0.775	1	0.9474
-0.1511	-0.042	0.1378	0.1425	-0.0427	-0.1967	-0.0116	0.5343	0.9474	1

optimal angle-parameterized rank-4 matrix $\rho(\theta^{*(4)})$:

1	0.9399	0.4826	-0.0863	-0.2715	-0.0437	0.1861	0.1808	-0.077	-0.2189
0.9399	1	0.7515	0.234	-0.0587	-0.0572	0.0496	0.0843	-0.0135	-0.077
0.4826	0.7515	1	0.7935	0.4329	0.015	-0.1745	-0.1195	0.0843	0.1808
-0.0863	0.234	0.7935	1	0.8432	0.3222	-0.0872	-0.1745	0.0496	0.1861
-0.2715	-0.0587	0.4329	0.8432	1	0.7421	0.3222	0.015	-0.0572	-0.0437
-0.0437	-0.0572	0.015	0.3222	0.7421	1	0.8432	0.4329	-0.0587	-0.2715
0.1861	0.0496	-0.1745	-0.0872	0.3222	0.8432	1	0.7935	0.234	-0.0863
0.1808	0.0843	-0.1195	-0.1745	0.015	0.4329	0.7935	1	0.7515	0.4826
-0.077	-0.0135	0.0843	0.0496	-0.0572	-0.0587	0.234	0.7515	1	0.9399
-0.2189	-0.077	0.1808	0.1861	-0.0437	-0.2715	-0.0863	0.4826	0.9399	1

again 10×10 full-rank $\hat{\rho}_{i,j} = \exp[-|i-j|]$.
If we resort to a rank-7 approximation, the zeroed-eigenvalues approach yields the following matrix $\rho^{(7)}$:

1	0.5481	0.0465	0.0944	0.0507	-0.0493	0.034	0.0169	-0.0441	0.0284
0.5481	1	0.6737	0.0647	0.0312	0.112	-0.0477	-0.0162	0.0691	-0.0441
0.0465	0.6737	1	0.579	0.1227	0.0353	0.0562	0.0012	-0.0162	0.0169
0.0944	0.0647	0.579	1	0.5822	0.0674	0.0806	0.0562	-0.0477	0.034
0.0507	0.0312	0.1227	0.5822	1	0.6472	0.0674	0.0353	0.112	-0.0493
-0.0493	0.112	0.0353	0.0674	0.6472	1	0.5822	0.1227	0.0312	0.0507
0.034	-0.0477	0.0562	0.0806	0.0674	0.5822	1	0.579	0.0647	0.0944
0.0169	-0.0162	0.0012	0.0562	0.0353	0.1227	0.579	1	0.6737	0.0465
-0.0441	0.0691	-0.0162	-0.0477	0.112	0.0312	0.0647	0.6737	1	0.5481
0.0284	-0.0441	0.0169	0.034	-0.0493	0.0507	0.0944	0.0465	0.5481	1

Optimization on an angle-parameterized rank-7 matrix yields the following output matrix $\rho(\theta^{*(7)})$:

1	0.5592	-0.0177	0.1085	0.0602	-0.0795	0.0589	0.018	-0.0734	0.0667
0.5592	1	0.5992	0.0202	0.0277	0.1123	-0.0652	-0.008	0.0797	-0.0734
-0.0177	0.5992	1	0.5464	0.0618	0.0401	0.0561	-0.012	-0.008	0.018
0.1085	0.0202	0.5464	1	0.5556	0.018	0.0834	0.0561	-0.0652	0.0589
0.0602	0.0277	0.0618	0.5556	1	0.5819	0.018	0.0401	0.1123	-0.0795
-0.0795	0.1123	0.0401	0.018	0.5819	1	0.5556	0.0618	0.0277	0.0602
0.0589	-0.0652	0.0561	0.0834	0.018	0.5556	1	0.5464	0.0202	0.1085
0.018	-0.008	-0.012	0.0561	0.0401	0.0618	0.5464	1	0.5992	-0.0177
-0.0734	0.0797	-0.008	-0.0652	0.1123	0.0277	0.0202	0.5992	1	0.5592
0.0667	-0.0734	0.018	0.0589	-0.0795	0.0602	0.1085	-0.0177	0.5592	1

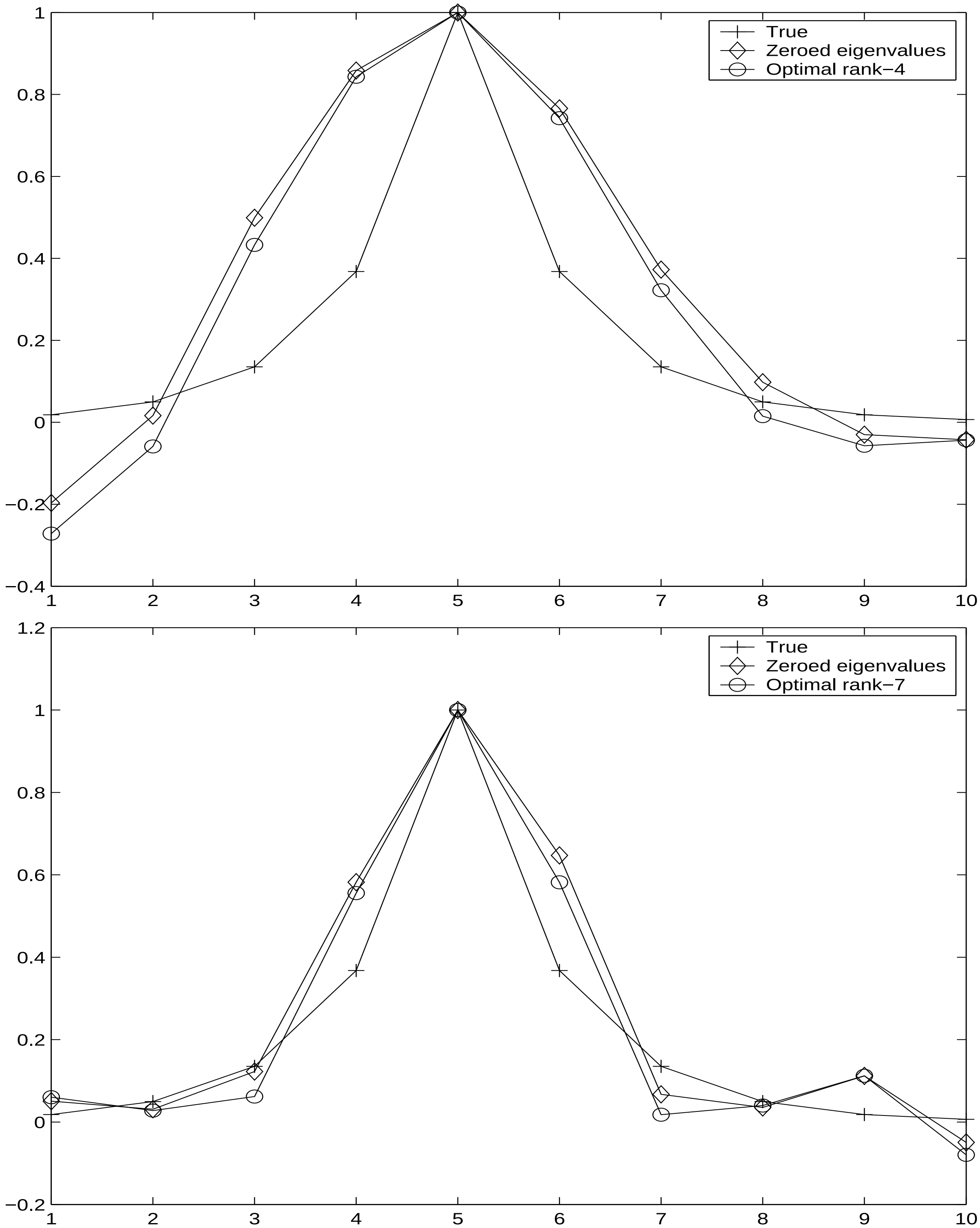


Figure 3: Problems of low rank correlation: sigmoid shape

Monte Carlo pricing swaptions with LFM

$$E^{\boxed{B}}\left(\frac{\boxed{B(0)}}{\boxed{B(T_\alpha)}}(S_{\alpha,\beta}(T_\alpha)-K)^+\sum_{i=\alpha+1}^{\beta}\tau_i P(T_\alpha,T_i)\right)=$$

$$=E^{\boxed{\alpha}}\left[\frac{\boxed{P(0,T_\alpha)}}{\boxed{P(T_\alpha,T_\alpha)}}(S_{\alpha,\beta}(T_\alpha)-K)^+\sum_{i=\alpha+1}^{\beta}\tau_i P(T_\alpha,T_i)\right].$$

$$=P(0,T_\alpha)E^{\alpha}\left[(S_{\alpha,\beta}(T_\alpha)-K)^+\sum_{i=\alpha+1}^{\beta}\tau_i P(T_\alpha,T_i)\right].$$

Since $S_{\alpha,\beta}(T_\alpha)=\dfrac{1-\prod_{j=\alpha+1}^{\beta}\frac{1}{1+\tau_j\ F_j(T_\alpha)}}{\sum_{i=\alpha+1}^{\beta}\tau_i\prod_{j=\alpha+1}^{i}\frac{1}{1+\tau_j\ F_j(T_\alpha)}}$

the above expectation depends on the *joint* distrib. under Q^α of

$$F_{\alpha+1}(T_\alpha),F_{\alpha+2}(T_\alpha),\ldots,F_\beta(T_\alpha)$$

Recall the dynamics of forward rates under Q^α:

$$dF_k(t)=\sigma_k(t)F_k(t)\sum_{j=\alpha+1}^{k}\frac{\rho_{k,j}\ \tau_j\ \sigma_j\ F_j}{1+\tau_j F_j(t)}\,dt+\sigma_k(t)F_k(t)dZ_k,$$

$$E^B\left(D(0,T_\alpha)\ (S_{\alpha,\beta}(T_\alpha)-K)^+\ \sum_{i=\alpha+1}^{\beta}\tau_i P(T_\alpha,T_i)\right)=$$

$$=P(0,T_\alpha)E^\alpha\left[(S_{\alpha,\beta}(T_\alpha)-K)^+\ \sum_{i=\alpha+1}^{\beta}\tau_i P(T_\alpha,T_i)\right].$$

Since $\quad S_{\alpha,\beta}(T_\alpha)=\dfrac{1-\prod_{j=\alpha+1}^{\beta}\frac{1}{1+\tau_j\ F_j(T_\alpha)}}{\sum_{i=\alpha+1}^{\beta}\tau_i\prod_{j=\alpha+1}^{i}\frac{1}{1+\tau_j\ F_j(T_\alpha)}}$

Milstein scheme for $\ln F$:

$$\ln F_k^{\Delta t}(t+\Delta t)=\ln F_k^{\Delta t}(t)+\sigma_k(t)\ \sum_{j=\alpha+1}^{k}\frac{\rho_{k,j}\ \tau_j\ \sigma_j(t)\ F_j^{\Delta t}}{1+\tau_j F_j^{\Delta t}}\,\Delta t+$$

$$-\frac{\sigma_k^2(t)}{2}\,\Delta t+\sigma_k(t)(Z_k(t+\Delta t)-Z_k(t))$$

leads to an approximation such that there exists a δ_0 with

$E^\alpha\{|\ln F_k^{\Delta t}(T_\alpha)-\ln F_k(T_\alpha)|\}\leq C(T_\alpha)(\Delta t)^1$ for all $\Delta t\leq\delta_0$

where $C(T_\alpha)>0$ is a constant (strong convergence of order 1).

$(Z_k(t+\Delta t)-Z_k(t))$ is GAUSSIAN and KNOWN, easy to simulate.

Monte Carlo pricing with LFM

A refined variance for simulating the shocks: Notice that in integrating **exactly** the dF equation between t and $t + \Delta t$, the resulting Brownian-motion part, in vector notation, is

$$\Delta\zeta_t := \int_t^{t+\Delta t} \underline{\sigma}(s) dZ(s) \sim \mathcal{N}(0, \mathsf{COV}_t)$$

(here the product of vectors acts component by component), where the matrix COV_t is given by

$$(\mathsf{COV}_t)_{h,k} = \int_t^{t+\Delta t} \rho_{h,k} \sigma_h(s) \sigma_k(s)\, ds.$$

Therefore, in principle we have no need to approximate this term by

$$\underline{\sigma}(t)(Z(t+\Delta t) - Z(t)) \sim \mathcal{N}(0,\ \Delta t\, \underline{\sigma}(t)\, \rho\, \underline{\sigma}(t)')$$

as is done in the classical general MC scheme given earlier. Indeed, we may consider a more refined scheme where the following substitution occurs:

$$\underline{\sigma}(t)(Z(t+\Delta t) - Z(t)) \longrightarrow \Delta\zeta_t.$$

The new shocks vector $\Delta\zeta_t$ can be simulated easily through its Gaussian distribution given above.

Monte Carlo pricing with LFM: Standard error

Assume we need to value a payoff $\Pi(T)$ depending on the realization of different forward LIBOR rates

$$F(t) = [F_{\alpha+1}(t), \dots, F_\beta(t)]'$$

in a time interval $t \in [0, T]$, where typically $T \leq T_\alpha$.

We have seen a particular case of $\Pi(T) = \Pi(T_\alpha)$ as the swaption payoff. The earlier simulation scheme for the rates entering the payoff provides us with the F's needed to form scenarios on $\Pi(T)$. Denote by a superscript the scenario (or path) under which a quantity is considered, $n_p = \#$ paths.

The Monte Carlo price of our payoff is computed, based on the simulated paths, as $E[D(0,T)\Pi(T)] =$

$$= P(0,T)E^T(\Pi(T)) = P(0,T)\sum_{j=1}^{n_p} \Pi^j(T)/n_p,$$

where the forward rates F^j entering $\Pi^j(T)$ have been simulated under the T-forward measure. We omit the T-argument in $\Pi(T), E^T$ and Std^T to contain notation: all distributions, expectations and statistics are under the T-forward measure. However, the reasoning is general and holds under any other measure.

Monte Carlo pricing with LFM: standard error

We wish to have an estimate of the error me have when estimating the true expectation $E(\Pi)$ by its Monte Carlo estimate $\sum_{j=1}^{n_p} \Pi^j / n_p$. To do so, the classic reasoning is as follows.

Let us view $(\Pi^j)_j$ as a sequence of independent identically distributed (iid) random variables, distributed as Π. By the central limit theorem, we know that under suitable assumptions one has

$$\frac{\sum_{j=1}^{n_p} (\Pi^j - E(\Pi))}{\sqrt{n_p}\, \mathsf{Std}(\Pi)} \to \mathcal{N}(0,1),$$

in law, as $n_p \to \infty$, from which we have that we may write, approximately and for large n_p:

$$\frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} - E(\Pi) \sim \frac{\mathsf{Std}(\Pi)}{\sqrt{n_p}}\, \mathcal{N}(0,1).$$

It follows that

$$Q^T \left\{ \left| \frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} - E(\Pi) \right| < \epsilon \right\} = Q^T \left\{ |\mathcal{N}(0,1)| < \epsilon \, \frac{\sqrt{n_p}}{\mathsf{Std}(\Pi)} \right\}$$

$$= 2\Phi\left(\epsilon \, \frac{\sqrt{n_p}}{\mathsf{Std}(\Pi)} \right) - 1,$$

where as usual Φ denotes the cumulative distribution function of the standard Gaussian random variable.

Monte Carlo pricing with LFM: standard error

$$Q^T \left\{ \left| \frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} - E(\Pi) \right| < \epsilon \right\} = 2\Phi\left(\epsilon \, \frac{\sqrt{n_p}}{\mathsf{Std}(\Pi)}\right) - 1,$$

The above equation gives the probability that our Monte Carlo estimate $\sum_{j=1}^{n_p} \Pi^j / n_p$ is not farther than ϵ from the true expectation $E(\Pi)$ we wish to estimate. Typically, one sets a desired value for this probability, say 0.98, and derives ϵ by solving

$$2\Phi\left(\epsilon \, \frac{\sqrt{n_p}}{\mathsf{Std}(\Pi)}\right) - 1 = 0.98.$$

For example, since we know from the Φ tables that

$$2\Phi(z) - 1 = 0.98 \iff \Phi(z) = 0.99 \iff z \approx 2.33,$$

we have that

$$\epsilon = 2.33 \, \frac{\mathsf{Std}(\Pi)}{\sqrt{n_p}}.$$

Monte Carlo pricing with LFM: standard error

The true value of $E(\Pi)$ is thus inside the "window"

$$\left[\frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} - 2.33\,\frac{\mathsf{Std}(\Pi)}{\sqrt{n_p}}, \quad \frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} + 2.33\,\frac{\mathsf{Std}(\Pi)}{\sqrt{n_p}}\right]$$

with a 98% probability. This is called a 98% confidence interval for $E(\Pi)$. Other typical confidence levels are given in Table 1.

$2\Phi(z) - 1$	$z \approx$
99%	2.58
98%	2.33
95.45%	2
95%	1.96
90%	1.65
68.27%	1

Table 1: Confidence levels

We can see that, ceteris paribus, as n_p increases, the window shrinks as $1/\sqrt{n_p}$, which is worse than $1/n_p$. If we need to reduce the window size to one tenth, we have to increase the number of scenarios by a factor 100. Sometimes, to reach a chosen accuracy (a small enough window), we need to take a huge number of scenarios n_p. When this is too time-consuming, there are "variance-reduction" techniques that may be used to reduce the above window size.

Monte Carlo pricing with LFM: standard error

A more fundamental problem with the above window is that the true standard deviation $\text{Std}(\Pi)$ of the payoff is usually unknown. This is typically replaced by the known sample standard deviation obtained by the simulated paths,

$$(\widehat{\text{Std}}(\Pi; n_p))^2 := \sum_{j=1}^{n_p} (\Pi^j)^2 / n_p - (\sum_{j=1}^{n_p} \Pi^j / n_p)^2$$

and the actual 98% Monte Carlo window we compute is

$$\left[\frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} - 2.33 \, \frac{\widehat{\text{Std}}(\Pi; n_p)}{\sqrt{n_p}}, \quad \frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} + 2.33 \, \frac{\widehat{\text{Std}}(\Pi; n_p)}{\sqrt{n_p}} \right]. \tag{7}$$

To obtain a 95% (narrower) window it is enough to replace 2.33 by 1.96, and to obtain a (still narrower) 90% window it is enough to replace 2.33 by 1.65. All other sizes may be derived by the Φ tables.

Monte Carlo pricing with LFM: Control Variate

We know that in some cases, to obtain a 98% window whose (half-) width $2.33\ \widehat{\mathsf{Std}}(\Pi; n_p)/\sqrt{n_p}$ is small enough, we are forced to take a huge number of paths n_p. This can be a problem for computational time. A way to reduce the impact of this problem is, for a given n_p that we deem to be large enough, to find alternatives that reduce the variance $(\widehat{\mathsf{Std}}(\Pi; n_p))^2$, thus narrowing the above window without increasing n_p.

One of the most effective methods to do this is the control variate technique.

We begin by selecting an alternative payoff Π^{an} which we know how to evaluate analytically, in that

$$E(\Pi^{\mathsf{an}}) = \pi^{\mathsf{an}}$$

is known. When we simulate our original payoff Π we now simulate also the analytical payoff Π^{an} as a function of the same scenarios for the underlying variables F. We define a new control-variate estimator for $E\Pi$ as

$$\widehat{\Pi}_c(\gamma; n_p) := \frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} + \gamma \left(\frac{\sum_{j=1}^{n_p} \Pi^{\mathsf{an},j}}{n_p} - \pi^{\mathsf{an}} \right),$$

Monte Carlo pricing with LFM: Control Variate

$$\widehat{\Pi}_c(\gamma; n_p) := \frac{\sum_{j=1}^{n_p} \Pi^j}{n_p} + \gamma \left(\frac{\sum_{j=1}^{n_p} \Pi^{\mathsf{an},j}}{n_p} - \pi^{\mathsf{an}} \right),$$

with γ a constant to be determined. When viewing Π^j as iid copies of Π and $\Pi^{\mathsf{an},j}$ as iid copies of Π^{an}, the above estimator remains unbiased, since we are subtracting the true known mean π^{an} from the correction term in γ. So, once we have found that the estimator has not been biased by our correction, we may wonder whether our correction can be used to lower the variance.

Consider the random variable

$$\Pi_c(\gamma) := \Pi + \gamma(\Pi^{\mathsf{an}} - \pi^{\mathsf{an}})$$

whose expectation is the $E(\Pi)$ we are estimating, and compute

$$\mathsf{Var}(\Pi_c(\gamma)) = \mathsf{Var}(\Pi) + \gamma^2 \mathsf{Var}(\Pi^{\mathsf{an}}) + 2\gamma \mathsf{Corr}(\Pi, \Pi^{\mathsf{an}}) \mathsf{Std}(\Pi) \mathsf{Std}(\Pi^{\mathsf{an}}),$$

We may minimize this function of γ by differentiating and setting the first derivative to zero.

Monte Carlo pricing with LFM: Control Variate

We obtain easily that the variance is minimized by the following value of γ: $\gamma^* := -\text{Corr}(\Pi, \Pi^{\text{an}})\text{Std}(\Pi)\ /\text{Std}(\Pi^{\text{an}})$. By plugging $\gamma = \gamma^*$ into the above expression, we obtain easily

$$\text{Var}(\Pi_c(\gamma^*)) = \text{Var}(\Pi)(1 - \text{Corr}(\Pi, \Pi^{\text{an}})^2),$$

from which we see that $\Pi_c(\gamma^*)$ has a smaller variance than our original Π, the smaller this variance the larger (in absolute value) the correlation between Π and Π^{an}. Accordingly, when moving to simulated quantities, we set

$$\widehat{\text{Std}}(\Pi_c(\gamma^*); n_p) = \widehat{\text{Std}}(\Pi; n_p)(1 - \widehat{\text{Corr}}(\Pi, \Pi^{\text{an}}; n_p)^2)^{1/2},$$

where $\widehat{\text{Corr}}(\Pi, \Pi^{\text{an}}; n_p)$ is the sample correlation

$$\widehat{\text{Corr}}(\Pi, \Pi^{\text{an}}; n_p) = \frac{\widehat{\text{Cov}}(\Pi, \Pi^{\text{an}}; n_p)}{\widehat{\text{Std}}(\Pi; n_p)\ \widehat{\text{Std}}(\Pi^{\text{an}}; n_p)}$$

and the sample covariance is

$$\widehat{\text{Cov}}(\Pi, \Pi^{\text{an}}; n_p) = \sum_{j=1}^{n_p} \Pi^j \Pi^{\text{an},j}/n_p - (\sum_{j=1}^{n_p} \Pi^j)(\sum_{j=1}^{n_p} \Pi^{\text{an},j})/(n_p^2)$$

and

$$(\widehat{\text{Std}}(\Pi^{\text{an}}; n_p))^2 := \sum_{j=1}^{n_p} (\Pi^{\text{an},j})^2/n_p - (\sum_{j=1}^{n_p} \Pi^{\text{an},j}/n_p)^2.$$

Monte Carlo pricing with LFM: Control Variate

One may include the correction factor $n_p/(n_p-1)$ to correct for the bias of the variance estimator, although the correction is irrelevant for large n_p.

We see from

$$\widehat{\mathrm{Std}}(\Pi_c(\gamma^*);n_p) = \widehat{\mathrm{Std}}(\Pi;n_p)(1-\widehat{\mathrm{Corr}}(\Pi,\Pi^{\mathrm{an}};n_p)^2)^{1/2},$$

that for the variance reduction to be relevant, we need to choose the analytical payoff Π^{an} to be as (positively or negatively) correlated as possible with the original payoff Π. Notice that in the limit case of correlation equal to one the variance shrinks to zero.

The window for our control-variate Monte Carlo estimate $\widehat{\Pi}_c(\gamma;n_p)$ of $E(\Pi)$ is now:

$$\left[\widehat{\Pi}_c(\gamma;n_p) - 2.33\,\frac{\widehat{\mathrm{Std}}(\Pi_c(\gamma^*);n_p)}{\sqrt{n_p}},\quad \widehat{\Pi}_c(\gamma;n_p) + 2.33\,\frac{\widehat{\mathrm{Std}}(\Pi_c(\gamma^*);n_p)}{\sqrt{n_p}}\right],$$

This window is narrower than the corresponding simple Monte Carlo one by a factor $(1-\widehat{\mathrm{Corr}}(\Pi,\Pi^{\mathrm{an}};n_p)^2)^{1/2}$.

Monte Carlo pricing with LFM: Control Variate

We may wonder about a good possible Π^{an}. We may select as Π^{an} the simplest payoff depending on the underlying rates

$$F(t) = [F_{\alpha+1}(t), \ldots, F_\beta(t)]'.$$

This is given by the Forward Rate Agreement (FRA) contract seen earlier. We consider the sum of at-the-money FRA payoffs, each on a single forward rate included in our family.

In other terms, if we are simulating under the T_j forward measure a payoff paying at T_α, with , the payoff we consider is

$$\Pi^{\mathrm{an}}(T_\alpha) = \sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i)(F_i(T_\alpha) - F_i(0))/P(T_\alpha, T_j)$$

whose expected value under the Q^j measure is easily seen to be 0 by remembering that quantities featuring $P(\cdot, T_j)$ as denominator are martingales. Thus in our case $\pi^{\mathrm{an}} = 0$ and we may use the related control-variate estimator. Somehow surprisingly, this simple correction has allowed us to reduce the number of paths of up to a factor 10 in several cases, including for example Monte Carlo evaluation of ratchet caps.

Analytical swaption prices with LFM

Approximated method to compute swaption prices with the LFM LIBOR MODEL without resorting to Monte Carlo simulation.

This method is rather simple and its quality has been tested in Brace, Dun, and Barton (1999) and by ourselves.

Recall the SWAP MODEL LSM leading to Black's formula for swaptions:

$$d\ S_{\alpha,\beta}(t) = \sigma^{(\alpha,\beta)}(t) S_{\alpha,\beta}(t)\ dW_t^{\alpha,\beta}, \quad Q^{\alpha,\beta}\ .$$

A crucial role is played by the Black swap volatility component

$$\int_0^{T\alpha} \sigma^2_{\alpha,\beta}(t) dt = \int_0^{T\alpha} \sigma_{\alpha,\beta}(t) dW^{\alpha,\beta}(t) \sigma_{\alpha,\beta}(t) dW^{\alpha,\beta}(t)$$

$$= \int_0^{T\alpha} (d \ln S_{\alpha,\beta}(t))(d \ln S_{\alpha,\beta}(t))$$

We compute an analogous approximated quantity in the LFM.

$$\begin{aligned} S_{\alpha,\beta}(t) &= \sum_{i=\alpha+1}^{\beta} w_i(t)\ F_i(t), \\ w_i(t) &= w_i(F_{\alpha+1}(t), F_{\alpha+2}(t), \ldots, F_\beta(t)) = \\ &= \frac{\tau_i \prod_{j=\alpha+1}^{i} \frac{1}{1+\tau_j F_j(t)}}{\sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \frac{1}{1+\tau_j F_j(t)}}\ . \end{aligned}$$

Freeze the w's at time 0:

$$S_{\alpha,\beta}(t) = \sum_{i=\alpha+1}^{\beta} w_i(t)\ F_i(t) \approx \sum_{i=\alpha+1}^{\beta} w_i(0)\ F_i(t)\ .$$

(variability of the w's is much smaller than variability of F's)

$$dS_{\alpha,\beta} \approx \sum_{i=\alpha+1}^{\beta} w_i(0)\, dF_i = (\ldots)dt + \sum_{i=\alpha+1}^{\beta} w_i(0)\sigma_i(t)F_i(t)dZ_i(t)\ ,$$

under any of the forward adjusted measures. Compute

$$dS_{\alpha,\beta}(t)dS_{\alpha,\beta}(t) \approx \sum_{i,j=\alpha+1}^{\beta} w_i(0)\sigma_i(t)F_i(t)dZ_i w_j(0)F_j(t)\sigma_j(t)\, dZ_j\ =$$

$$= \sum_{i,j=\alpha+1}^{\beta} w_i(0)w_j(0)F_i(t)F_j(t)\rho_{i,j}\sigma_i(t)\sigma_j(t)\ dt\ .$$

The percentage quadratic covariation is

$$(d\ln S_{\alpha,\beta}(t))(d\ln S_{\alpha,\beta}(t)) = \frac{dS_{\alpha,\beta}(t)}{S_{\alpha,\beta}(t)}\frac{dS_{\alpha,\beta}(t)}{S_{\alpha,\beta}(t)} =$$

$$\approx \frac{\sum_{i,j=\alpha+1}^{\beta} w_i(0)w_j(0)F_i(t)F_j(t)\rho_{i,j}\sigma_i(t)\sigma_j(t)}{S_{\alpha,\beta}(t)^2}\ dt\ .$$

Introduce a further approx by freezing again all F's (as was done earlier for the w's) to time zero: $(d\ln S_{\alpha,\beta})(d\ln S_{\alpha,\beta}) \approx$

$$\approx \sum_{i,j=\alpha+1}^{\beta} \frac{w_i(0)w_j(0)F_i(0)F_j(0)\rho_{i,j}}{S_{\alpha,\beta}(0)^2}\,\sigma_i(t)\sigma_j(t)\;dt\ .$$

Now compute the time-averaged percentage variance of S as

(Rebonato's Formula)

$$(v_{\alpha,\beta}^{\mathsf{LFM}})^2 = \frac{1}{T_\alpha}\int_0^{T_\alpha}(d\ln S_{\alpha,\beta}(t))(d\ln S_{\alpha,\beta}(t))$$

$$= \sum_{i,j=\alpha+1}^{\beta} \frac{w_i(0)w_j(0)F_i(0)F_j(0)\rho_{i,j}}{T_\alpha\ S_{\alpha,\beta}(0)^2}\ \int_0^{T_\alpha}\sigma_i(t)\sigma_j(t)\;dt\ .$$

$v_{\alpha,\beta}^{\mathsf{LFM}}$ can be used as a proxy for the Black volatility $v_{\alpha,\beta}(T_\alpha)$.

Use Black's formula for swaptions with volatility $v_{\alpha,\beta}^{\mathsf{LFM}}$ to price swaptions **analytically** with the LFM.

Analytical swaption prices with LFM

It turns out that the approximation is not at all bad, as pointed out by Brace, Dun and Barton (1999) and by ourselves.

A slightly more sophisticated version of this procedure has been pointed out for example by Hull and White (1999).

This pricing formula is ALGEBRAIC and very quick (compare with short-rate models)

H–W refine this formula by differentiating $S_{\alpha,\beta}(t)$ without immediately freezing the w. Same accuracy in practice.

Analytical terminal correlation

By similar arguments (freezing the drift and collapsing all measures) we may find a formula for terminal correlation.

Corr$(F_i(T_\alpha), F_j(T_\alpha))$ should be computed with MC simulation and depends on the chosen numeraire

Useful to have a first idea on the stability of the model correlation at future times.

Traders need to check this quickly, no time for MC

In Brigo and Mercurio (2001), we obtain easily

$$\frac{\exp\left(\int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)\rho_{i,j}dt\right) - 1}{\sqrt{\exp\left(\int_0^{T_\alpha} \sigma_i^2(t)dt\right) - 1}\ \sqrt{\exp\left(\int_0^{T_\alpha} \sigma_j^2(t)dt\right) - 1}}$$

$$\approx \rho_{i,j}\frac{\int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)\ dt}{\sqrt{\int_0^{T_\alpha} \sigma_i^2(t)dt}\ \sqrt{\int_0^{T_\alpha} \sigma_j^2(t)dt}}\ ,$$

the second approximation as from Rebonato (1999). Schwartz's inequality: *terminal correlations are always smaller, in absolute value, than instantaneous correlations*.