PCA: invariant risk metrics and representation of residuals for bond returns

K. Musaelian[1], S. Nagarajan *[2], and D. Villani[2,3]

[1]HiQu Capital, 123 East 71st Street, New York, NY 10021
[2]Tudor Investment Corporation, 51 Astor Place, New York, NY 10003
[3]Department of Operations Research & Financial Engineering and Bendheim Center for Finance, Princeton University, Princeton, NJ 08544

September 15, 2016

We show how to build robust risk metrics for bond returns based on a global structure in the form of principal components and a novel quasi-local representation for the residuals.

1 Introduction

The highly correlated nature of bond returns creates a particular challenge for portfolio management of interest rate products. Unlike stock portfolios where large idiosyncratic components of single name returns limit the amount of concentration and leverage, interest rate strategies rely on large and varying amounts of leverage and have limited avenues for diversification. This makes it especially valuable to create a parsimonious representation of risk that reflects the correlation structure of returns of various maturities and takes into account risks associated with the possibility of a breakdown of the correlation structure. The former was addressed by R. Litterman and J. Scheinkman [1] in their seminal paper. In this approach, dimensional reduction is accomplished by representing risk with two or three principal components (PCs), and since these explain most of the variations in the curve, the space orthogonal to these components is ignored. Since their work, rates trading has expanded in scope and increased in complexity; risk taking now includes many strategies that are not exposed to the first few PCs. It is indeed very easy to build a large portfolio and zero out exposures to the first few principal components, but that does not mean the portfolio is riskless. In fact, the issue is that all the higher order principal components have smaller eigenvalues; capping these is a challenge given their non-intuitive nature and the statistical noise associated with their small magnitude. In order to effectively risk manage a portfolio of such strategies, the following four drawbacks of the traditional principal component approach have to be addressed: a) it requires the specification of multiple risk limits (one for each PC chosen), b) misses convexity or includes it as a separate limit, c) does not preclude portfolios with enormous notional and leverage, and d) does not properly address tail risks due to market stress. Limiting gross notional exposure is a crude way of dealing with some of these problems, but it suffers from arbitrariness and inconsistency. In this paper, we propose a framework that combines a principal component based global risk measure with a novel ***flylet*** representation for the residual risk. In deriving this framework, we address each of the aforementioned drawbacks individually. First, we stress that there is no need to impose independent risk limits on the first few PCs (we choose the first two PCs). We derive a measure that encapsulates risks of all included principal components, thus reducing dimensionality. Furthermore, the measure is invariant under rotation, and allows easy risk aggregation across different curves (*e.g.*, different currencies in a global fixed income portfolio) without adjustments for the actual orientation of PC_1 and PC_2. Second, we expand the framework to compute an invariant measure that includes convexity. Third, we provide a sensible structure to the space not spanned by the first few principal components. Using quasi-local objects we call ***flylets*** (similar in spirit to wavelets [2])

*Corresponding Author: nagsanthanam@gmail.com

we build a complete and intuitive basis in the space orthogonal to PC_1 and PC_2. An added benefit of ***flylets*** is that they are traded in the market and their volatilities are of comparable magnitude; hence it is easier to define reasonable exposure limits. Lastly, we generalize the method to non-normal distributions, showing how the methodology can be extended to account for tail risk.

2 Invariant Risk Metrics

The use of Principal Component Analysis (PCA) for describing common factors affecting bond returns was put forward in [1]. Here, we consider the same approach for a curve $\mathscr{C}$ consisting of n points, and start by representing the vector daily yield changes Δy^d (where d identifies the day over which the change is computed) using the basis formed by the principal components of $\mathscr{C}$ as:

$$\Delta y^d = \sum_{j=1}^{n} \alpha_j^d PC_j \tag{1}$$

We want to represent the impact of changes in $\mathscr{C}$ on a portfolio Π of interest rate products. In the following sections, we build the framework one step at a time, starting with the risk metric for the net delta exposure of Π to the principal components in section 2.1, adding convexity terms in section 2.2, accounting for exposure in the subspace orthogonal to PC_1 and PC_2 in section 2.3, and finally, in section 2.4 we extend the analysis to fat-tailed distributions.

2.1 Invariant measure - Delta

In this section, we compute an invariant metric that depends solely on the delta exposure of Π, *i.e.*, the first derivative of Π with respect to the principal components. We consider the n-vector of yield changes in Eq. (1), and collect the de-trended changes over the m-days in the $m \times n$ matrix Y, where each row is a vector of yield changes over a single day. Principal component analysis is associated with the singular value decomposition $Y = U\Sigma V^T$. Since $m > n$, we start with the reduced form $Y = \hat{U}\hat{\Sigma}V^T$ where $\hat{U}(m \times n)$ and $V(n \times n)$ are unitary matrices whose columns are the left and right singular vectors of Y respectively, and $\hat{\Sigma}$ is the $n \times n$ diagonal matrix containing the singular values of Y. The covariance matrix $C = Y^TY$ of the n points on curve $\mathscr{C}$ then becomes $C = V\hat{\Sigma}^2V^T$, and the principal components are the columns of $\frac{1}{\sqrt{m}}\hat{\Sigma}V^T$.

Defining the m-vector ΔP consisting of the daily changes in the portfolio value due to de-trended yield changes, we can write

$$\begin{aligned} \Delta P &= \sqrt{m}\,\hat{U}J \qquad\qquad \text{where } J_j = \frac{\partial \Pi}{\partial PC_j} \\ \implies \frac{\Delta P^T \Delta P}{m} &= J^T\hat{U}^T\hat{U}J = \|J\|_2^2 \end{aligned} \tag{2}$$

Therefore, the realized *in-sample* variance of the portfolio played back over the period of PCA is $\|J\|_2^2$, with $\|\cdot\|_2$ being the rotationally invariant L_2-norm. The main drawback of this derivation, in addition to the fact that it holds only in-sample, is that the result does not work well when applied to portfolios that include convex instruments.

2.2 Invariant measure - Delta and Gamma

In this section, we use a different approach to extend the invariant measure to include the second derivatives of the portfolio with respect to the principal components. The change in the value of the portfolio Π with respect to movement of the curve $\mathscr{C}$ over a single day (for the rest of this paper, we drop the day identifier d) is

$$\Delta\Pi = J^T\alpha + \frac{1}{2}\alpha^T H\alpha \tag{3}$$

where $H_{ij} = \frac{\partial^2 \Pi}{\partial PC_i \partial PC_j}$. If we assume that the PC coefficients α_i are i.i.d. and normally distributed ($\sim N(0,1)$), it can be shown that the variance of $\Delta\Pi$ is

$$\mathbb{E}\big(\Delta\Pi^2\big) = \|J\|_2^2 + \frac{1}{2}\|H\|_F^2 + \frac{1}{4}\left[Tr(H)\right]^2 = \|J\|_2^2 + \frac{1}{2}\|H\|_F^2 + \frac{1}{4}\|H\|_*^2 \tag{4}$$

where $\|\cdot\|_F$ and $\|\cdot\|_*$ are the Frobenius and nuclear norms respectively. In this paper, we choose the first two principal components, and Eq. (4) reduces to

$$\begin{aligned}\mathbb{E}\big(\Delta\Pi^2\big) \;=\; & \left(\frac{\partial \Pi}{\partial PC_1}\right)^2 + \left(\frac{\partial \Pi}{\partial PC_2}\right)^2 + \frac{1}{2}\left[\left(\frac{\partial^2 \Pi}{\partial PC_1^2}\right)^2 + \left(\frac{\partial^2 \Pi}{\partial PC_2^2}\right)^2 + \left(\frac{\partial^2 \Pi}{\partial PC_1 \partial PC_2}\right)^2\right] \\ & + \frac{1}{4}\left[\frac{\partial^2 \Pi}{\partial PC_1^2} + \frac{\partial^2 \Pi}{\partial PC_2^2}\right]^2 \end{aligned} \tag{5}$$

For example, the i^{th} element of J is computed as the change in the value of the portfolio Π when the underlying rate curve moves by PC_i (and the second derivatives are computed similarly). Note also that the formula (2) is $O(N)$ invariant given definitions of L_2 norm of a vector, and Frobenius & nuclear norms of a matrix. In this sense, capping (4) is an elegant and invariant way to cap both linear and convexity terms (especially important for portfolios with options) at once without having to choose an orientation in the space spanned by PCs.

Note that while we use standard PCA in section 2.1, the results in this section hold for any subspace as long as the coefficients are normally distributed. We use PCs for notational convenience and intuitive clarity; the actual methodology used to identify the subspace is not relevant to the developments herein. The only requirement is that we have a good low rank representation of the dynamics of curve $\mathscr{C}$, and more advanced techniques can help in this regard.

2.3 *Flylet* representation of residuals

The result in Eq. (5) provides a metric for the risk in the 2-dimensional sub-space spanned by PC_1 and PC_2. To represent the remaining risk in the space orthogonal to PC_1 and PC_2, we span this $n-2$ dimensional sub-space with a basis of ***flylets*** (Fig. 1). The i^{th} ***flylet*** is denoted by

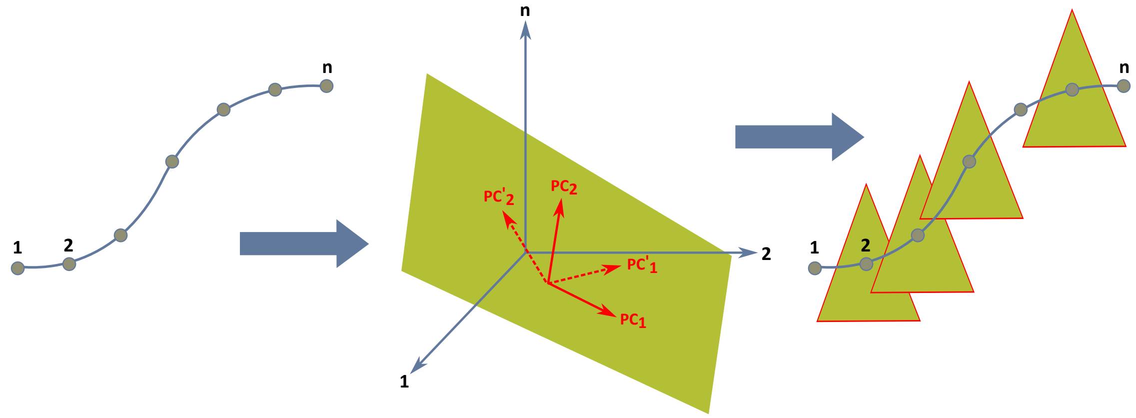
$$f_i = \big(0_{(i-2)}, a_i, b_i, c_i, 0_{(n-i-1)}\big) \text{ for } i = 2, \cdots, n-1 \tag{6}$$

where the vector f_i denotes the weights of the flylet centered at the i^{th} node, and $0_{(k)}$ is a vector of k zeros. The weights are computed such that

$$f_i \cdot PC_1 \;=\; f_i \cdot PC_2 = 0 \tag{7}$$

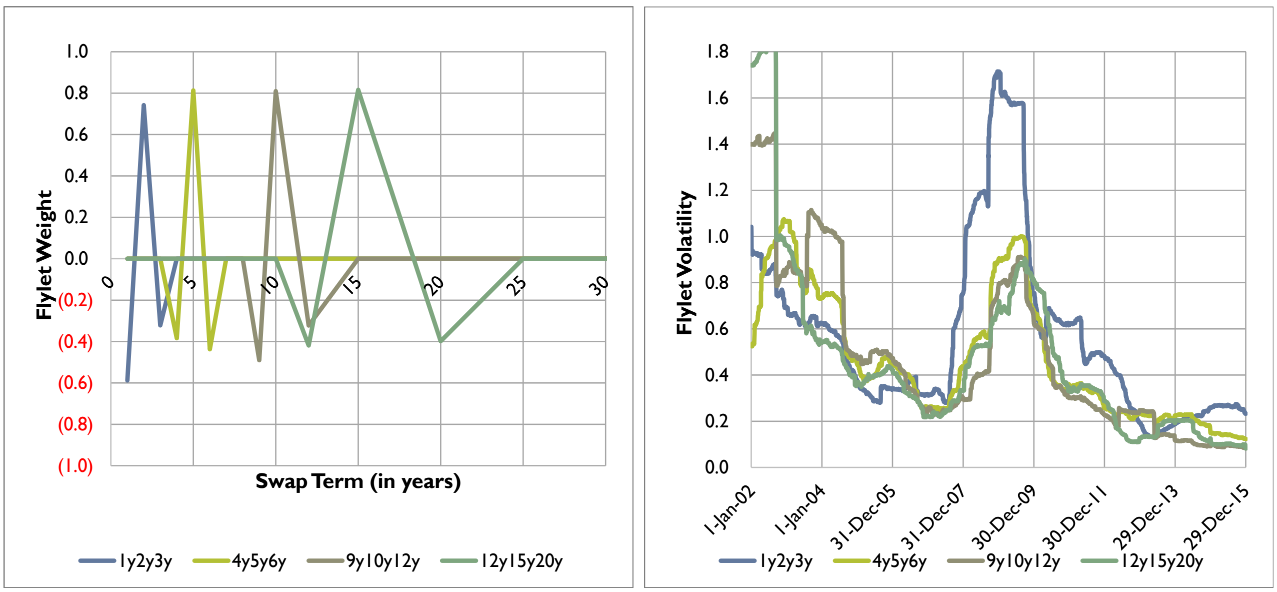
$$\|f_i\|_2^2 \;=\; 1 \tag{8}$$

By definition, the ***flylets*** span the space orthogonal to PC_1 and PC_2. These are not the standard flies with weights $1 \times 2 \times 1$ (see Fig. 2a for examples). The advantage of using ***flylets*** in conjunction with the invariant measure of the previous section is that it combines a global and a local measure. The ***flylets*** represent the local curvature of the yield curve, and have the additional advantage that all ***flylet*** volatilities are of the same order, as shown in Fig. 2b. Defining the vector of exposures F such that $F_i = \partial\Pi/\partial f_i$ is the sensitivity of the portfolio to the normalized ***flylet*** f_i, a generic metric for the net ***flylet*** exposure of portfolio Π may be expressed as $\sigma_F^2 \|F\|_2^2$, where σ_F is a representative ***flylet*** volatility (see Fig. 2b).



Curve $\mathscr{C}$ $\quad$ PC_1-PC_2 subspace $\quad$ ***Flylets***

Figure 1: The ***flylets*** form a quasi-local complete basis for the space orthogonal to that spanned by PC_1 and PC_2



Data Source: Bloomberg

(a) Weights of normalized ***flylets***

Data Source: Bloomberg

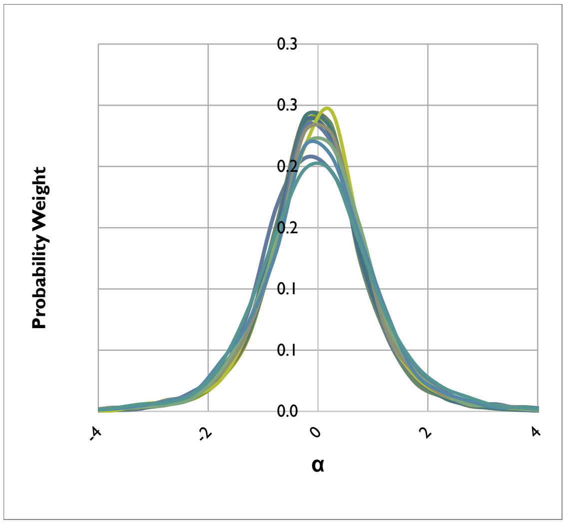
(b) Trailing 1 year volatility (in bps/day) of normalized ***flylets***

Figure 2: Characteristics of sample ***flylets*** at the short, belly, and long end of the USD curve

Buy and sell-side rates desks, and risk management groups routinely rely on local exposures in terms of partial durations, which are the portfolio sensitivities $\partial\Pi/\partial e_i$, where $e_i = \{0, \cdots, 1, \cdots, 0\}^T$ is the n-vector representing a small perturbation to the i^{th} point of curve $\mathscr{C}$. The disadvantage of using this approach is that when projected to the subspace orthogonal to PC_1 and PC_2 (applying Eq. (7) and (8) to e_i), the vector is no longer local, *i.e.*, all n-points are non-zero.

2.4 Tail risk and market stress

Equation (4) was derived with the assumption that the coefficients α_i are i.i.d. $N(0, 1)$. After a period of relative calm, PC_1 and PC_2 volatilities would drift down and the exposures (J and H in Eq. (4)) will be low, providing a false sense of comfort that the portfolio is benign. Fig. 3 shows the distributions of PC coefficients computed for daily yield changes as in Eq. (1) using 1 year trailing PCs. The coefficients are distributed symmetrically, but are generally heavy-tailed, as seen from the excess kurtosis in Fig. 3. For the current methodology to be useful, it is therefore essential to extend the framework to include the impact of fat tails.



	MEAN	STD DEV	SKEWNESS	EXCESS KURTOSIS
α_1	0.00	1.00	0.16	1.19
α_2	0.01	1.05	-0.15	6.40
α_3	-0.03	1.05	0.51	8.82
α_4	-0.04	1.12	-0.02	18.29
α_5	-0.05	1.17	-1.92	30.19
α_6	-0.01	1.08	-0.42	11.43
α_7	-0.03	1.10	-0.75	17.71
α_8	0.02	1.11	-1.11	26.65
α_9	0.02	1.11	0.24	6.98
α_{10}	0.02	1.12	-0.38	7.43
α_{11}	0.01	1.28	5.10	140.90
α_{12}	0.03	1.32	2.84	42.29

Data Source: Bloomberg

Data Source: Bloomberg

Figure 3: Empirical characteristics of α_i distributions

One approach is to consider the α_i distribution to be mixed normal, *i.e.*, assume that the probability of a market stress is p so that the the coefficients are distributed as $N(0, \theta)$ with probability p, and as $N(0, 1)$ with probability $1 - p$, where θ is the volatility in stressed market conditions. With this mixed distribution, the expression derived in (2) takes the form

$$\mathbb{E}\big(\Delta\Pi^2\big) = \left[1 + p\left(\theta^2 - 1\right)\right] \|J\|_2^2 + \left[1 + p\left(\theta^4 - 1\right)\right] \left\{\frac{1}{2}\|H\|_F^2 + \frac{1}{4}\left[Tr(H)\right]^2\right\} \tag{9}$$

where we have assumed that all α_i jump to the stressed state simultaneously. It is possible to derive such expressions for more generic distributions of α_i following the general approach in the appendix.

The impact of fat tails on expected variance of the portfolio due to ***flylet*** exposure is linear, and therefore similar to the linear term in Eq. (9)

$$\mathbb{E}\big(\Delta\Pi_F^2\big) = \sigma_F^2\left[1 + p\left(\theta^2 - 1\right)\right] \|F\|_2^2 \tag{10}$$

where $\Delta\Pi_F$ is the change in value of the portfolio due to changes in the ***flylet*** levels. Note that the PC and ***flylet*** risk measures should be treated separately, given the difference in the orders of magnitude of these exposures.

3 Application

To illustrate these concepts, we consider the case of the USD swap curve, denoted by $\mathscr{C}$. The value of any rates instrument π_i (swap/swaptions etc.) can then be written as $\pi_i = \pi_i(\mathscr{C}, \vec{p})$, where $\vec{p}$ are other factors (in addition to $\mathscr{C}$) to which π_i is sensitive (*e.g.*, volatility in the case of swaptions). The swap curve $\mathscr{C}$ is built every day, and principal components of the zero rates are used to compute the first order PC sensitivities as

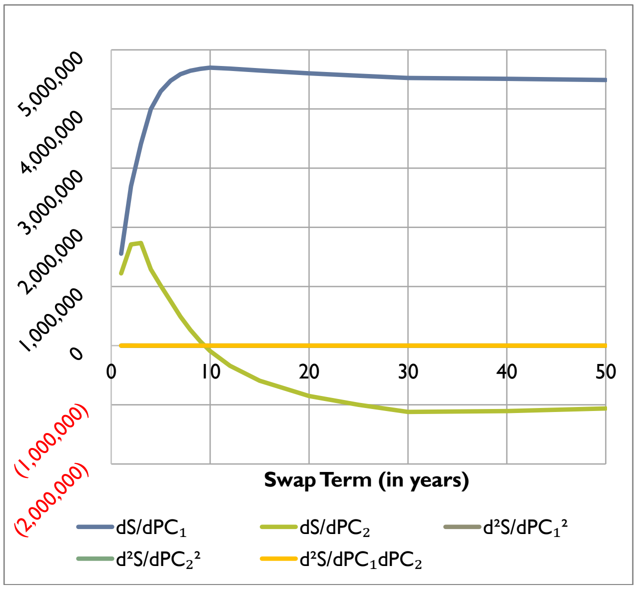
$$\frac{\partial \pi_i}{\partial PC_j} = \frac{\pi_i(\mathscr{C} + \varepsilon PC_j) - \pi_i(\mathscr{C} - \varepsilon PC_j)}{2\varepsilon}, \quad \text{where } \epsilon \ll 1 \tag{11}$$

The second derivatives (elements of H) are computed using the corresponding central difference formulae, and sensitivities of individual instruments aggregated to the portfolio level.

The examples below follow the development of the framework in the previous section. An example of the risk metric for delta products (swaps) is given in section 3.1, followed by delta and gamma (for swaptions) in section 3.2. Section 3.3 illustrates the ***flylet*** representation of residuals for the case of PC_1 and PC_2 neutral swap flies, and finally, section 3.4 extends the swap and swaption examples to include heavy tails.

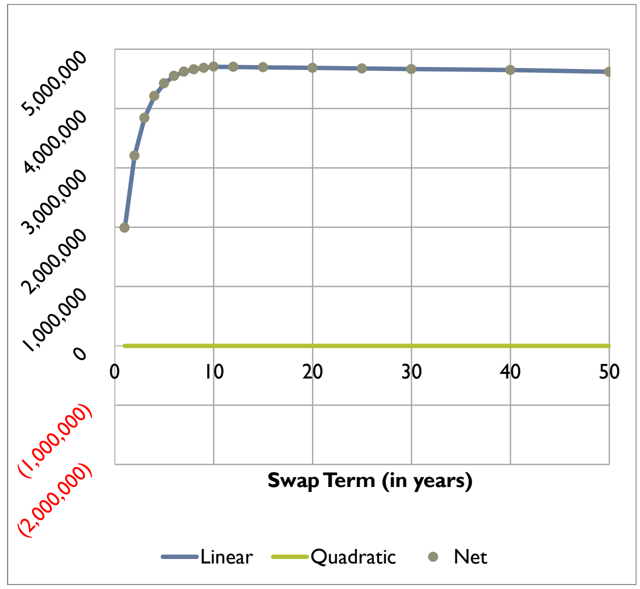
3.1 Invariant measure - Delta

We first consider the example of USD swaps of various tenors with $DV01 = 1,000,000$. Fig. 4a shows derivatives of swaps with respect to PC_1 and PC_2, *i.e.*, the elements of J and H. Swaps are close to linear (except for slight non-linearity due to dependence of DV01 on the spread), and this is borne out in this figure. If we were to risk manage a book of swaps based on these risks, we would need two risk limits, one each for PC_1 and PC_2 respectively. Fig. 4b shows linear, quadratic, and the net invariant risk measure derived in Eq. (4).



Data Source: Bloomberg

(a) Sensitivities of swaps to PC_1 and PC_2



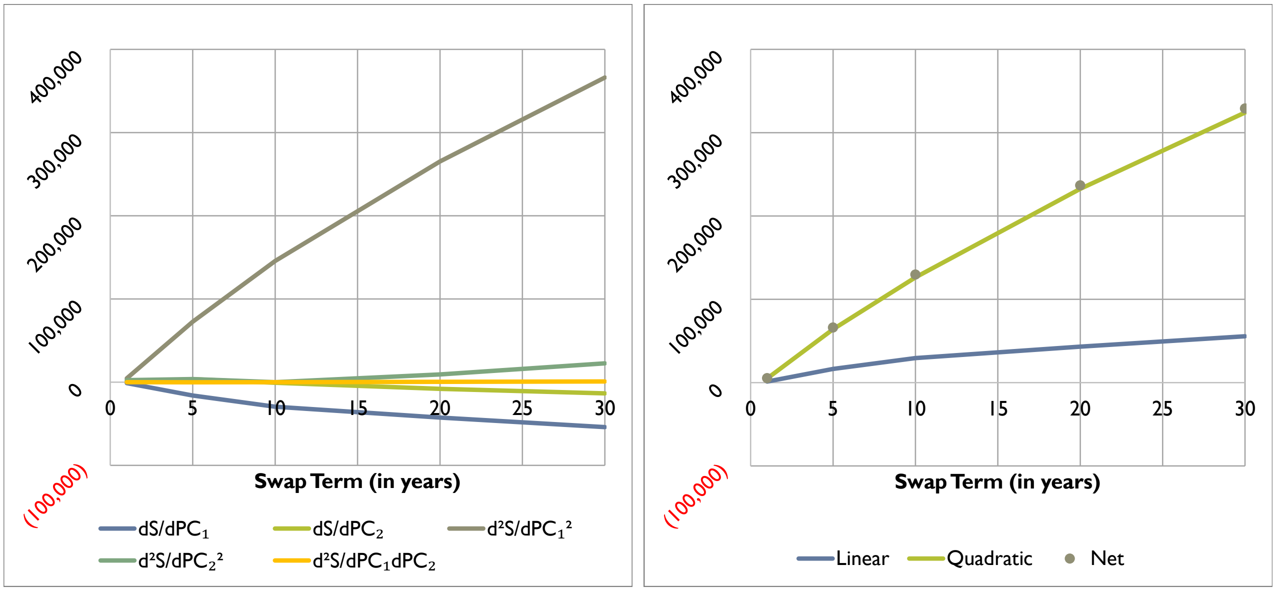
Data Source: Bloomberg

(b) $\sqrt{\mathbb{E}(\Delta\Pi^2)}$ for USD swaps as a function of term

Figure 4: PC sensitivities and combined invariant measure for outright swap positions

3.2 Invariant measure - Delta and Gamma

To illustrate how convexity is captured in this framework, we consider USD swaption straddles of 1 month expiry and terms of 1y/5y/10y/20y/30y. The notional of each leg is taken to be 200 million. In Fig. 5a, we see that swaption straddles have both linear and quadratic exposures as expected, but the Hessian terms of Eq. (4) dominate as seen in Fig. 5b. The advantage of using the invariant risk measure is now obvious: ***a complex book that is a mix of linear and convex products can use a single, consistent measure for risk management***.



Data Source: Bloomberg

(a) Sensitivities to PC_1 and PC_2

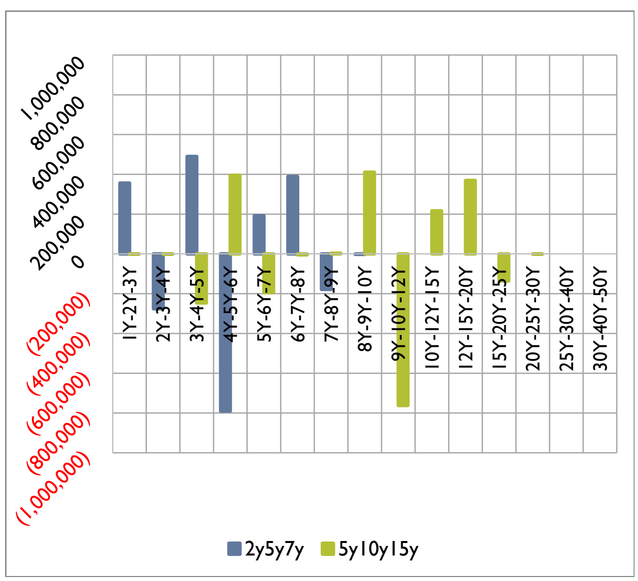
Data Source: Bloomberg

(b) $\sqrt{\mathbb{E}(\Delta\Pi^2)}$ as a function of term

Figure 5: PC sensitivities and combined invariant measure for 1 month USD swaption straddles

3.3 *Flylet* representation of residuals

Many rates trading strategies explicitly rely on PCA for risk management, and many buy and sell side desks run portfolios that are PC_1 and PC_2 neutral. The invariant measure of Eq. (4) would miss these risks completely. The example below shows two swap flies, a 2y5y7y and a 5y10y15y, constructed specifically to neutralize PC_1 and PC_2 exposures, and hence the invariant measure is negligible. However, as shown in Fig. 6, the ***flylets*** pick up exposures associated with such trades/strategies.

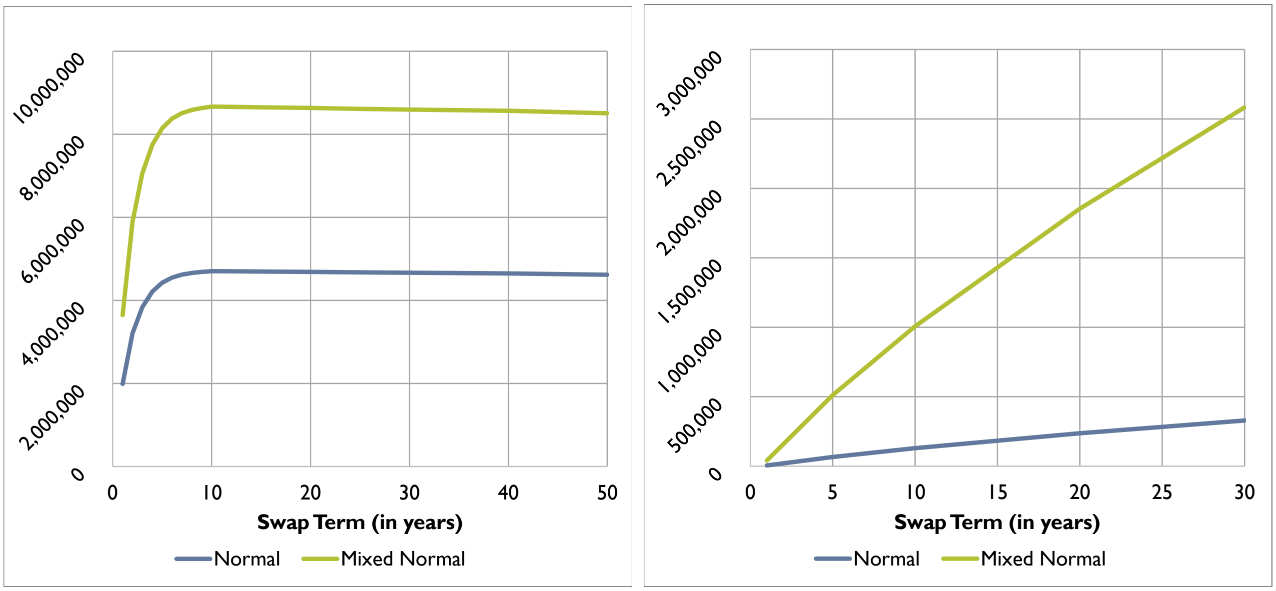


Data Source: Bloomberg

Figure 6: ***Flylet*** exposures of two USD swap flies

3.4 Tail risk and market stress

Lastly, we add the impact of including market stresses. For the purpose of illustration, we choose an extreme set of parameters; assume that there is a 90% chance of normal market conditions persisting, with a 10% chance of entering stressed market conditions where volatility increases fivefold. Using Eq. (9) we compute the net risk for both the swaps and swaption straddles of Figs. 4b and 5b, respectively. The impact of market stress doubles the net risk for swaps (Fig. 7a), while swaption straddles see an eight-fold increase (Fig. 7b). These factors are consistent with the θ^2 and θ^4 dependence of linear and quadratic terms in Eq. (9).



Data Source: Bloomberg

(a) Swaps of various tenors, DV01 = 1 million

Data Source: Bloomberg

(b) 1 month swaption straddles (200 million notional)

Figure 7: Combined invariant measure for mixed normal distributions

4 Summary

In this work, we have identified and addressed the four main drawbacks of using PCA for risk management of bond portfolios and other rates products:

1. Specification of multiple risk limits avoided through a combined risk measure,
2. Missing convexity also addressed by the combined risk measure derived in section 2.2,
3. Residual risk represented with a quasi-local ***flylet*** basis. This is especially important for portfolios with large notionals and low *PC* risk, and
4. Impact of tail risk included by extending the combined risk measure of Eq. (4) to fat tailed distributions

References

[1] R. Litterman, J. Scheinkman *Common factors affecting bond returns*, The Journal of Fixed Income, June 1991

[2] C. Gasquet, P. Witomski *Fourier Analysis and Applications: Filtering, Numerical Computation, Wavelets*, Texts in Applied Mathematics, Springer 1999

[3] S. Kullback, R. A. Leibler *On information and sufficiency*, Annals of Mathematical Statistics **22**, 1951

[4] F. Black *Interest rates as options*, Journal of Finance, **50**, 1995

[5] L. Krippner *Modifying Gaussian term structure models when interest rates are near the zero lower bound*, Reserve Bank of New Zealand, Discussion Paper DP2012/02

A Derivation of invariant measure

Following equation (3), the variance of the change in value of portfolio Π is given by

$$\mathbb{E}\big[\Delta\Pi^2\big] = \mathbb{E}\Big[\alpha^T J J^T \alpha + \frac{1}{2}\left(\alpha^T J \alpha^T H \alpha + \alpha^T H^T \alpha J^T \alpha\right) + \frac{1}{4}\left(\alpha^T H^T \alpha\right)\left(\alpha^T H \alpha\right)\Big] \tag{12}$$

In this work, we restrict ourselves to the case of symmetric distributions of α, and the distributions in Fig. 3 support this assumption. However, skewed distributions may be observed in practice, especially when rates are close to a floor [4, 5], and we can use any generic distribution in equation (12) to compute risk limits, which for distributions with zero mean, reduces to

$$\begin{aligned}
\mathbb{E}\big[\Delta\Pi^2\big] &= \sum_{i=1}^{n} J_i^2\ \mathbb{E}\big(\alpha_i^2\big) + \mathbb{E}\left[\sum_{i,j,k=1}^{n} (\alpha_i J_i)(\alpha_j H_{jk} \alpha_k)\right] + \frac{1}{4}\ \mathbb{E}\left[\sum_{i,j,k,l=1}^{n} (\alpha_i H_{ik} \alpha_k)(\alpha_j H_{jl} \alpha_l)\right] \\
&= \sum_{i=1}^{n} J_i^2\ \mathbb{E}\big(\alpha_i^2\big) + \sum_{i=1}^{n} J_i H_{ii} \mathbb{E}\left(\alpha_i^3\right) + \frac{1}{4}\sum_{\substack{i,j=1 \\ i\neq j}}^{n} \left[H_{ii}H_{jj} + 2H_{ij}^2\right]\mathbb{E}\big(\alpha_i^2\alpha_j^2\big) + \frac{1}{4}\sum_{i=1}^{n} H_{ii}^2\ \mathbb{E}\big(\alpha_i^4\big)
\end{aligned}$$

Note that we ignore 4^{th}-order terms arising from multiplication of J with the third derivative of the portfolio with respect to α. Any distribution with well defined skewness and kurtosis can be used in the above set of equations to derive combined risk limits.

Acknowledgment

The authors thank Neil Chriss and Hutchin Hill Capital, LP for discussions that contributed to the development of this paper. S.N. and D.V. also thank Albert Honein, Sean Zhou, Jonathan Kommemi, Joanne Im and Deepika Pyla for their careful reading of the manuscript, and Stephen Taylor for his feedback and suggestion to include the third order term in the last equation.