Credit risk and incomplete information: filtering and EM parameter estimation

Wolfgang Runggaldier

Dipartimento di Matematica Pura ed Applicata, Università di Padova
www.math.unipd.it/runggaldier

CREST-Sakigake Conference, Tokyo, December 2010.

Outline

I. General introduction.
- Filtering in financial market models in particular Markovian factor models;
- Filtering vs. calibration.

II. Basic facts from Credit Risk.
- Reduced-form/intensity-based models;
- Incomplete information (investor filtration);
- Linear vs. nonlinear models.

III. Pricing and parameter estimation in an “affine” credit risk model under partial information.
- Model and incomplete information setup;
- The filter methodology;
- Parameter estimation via the EM algorithm and numerical results.

- Parts I and II are based on

  R.Frey, W.Runggaldier, “Nonlinear Filtering in Models for Interest Rate and Credit Risk”. To appear in *Handbook of Nonlinear Filtering* (D.Crisan and B.Rozovski eds.), Oxford University Press.

- Part III is based on

  C.Fontana, W.J.Runggaldier, “Credit risk and incomplete information: filtering and EM parameter estimation”. International Journal of Theoretical and Applied Finance, 13 (2010) No. 1, pp. 683-715.

Filtering in financial market models

→ Filtering: when the underlying financial model is not fully known *(here: unknown factor process)*

- Filtering to price illiquid assets *(filtering under a martingale/pricing measure)*
- Filtering in risk management such as hedging, portfolio optimization,....*(filtering under the physical measure)*
- Mixed problems.

→ Here filtering mainly for pricing purposes (in credit risk models)

Pricing by Martingale Methods

- Given is a triple $(\mathcal{G}_t, N_t, Q^N)$ with

  $\mathcal{G}_t$: a filtration (global filtration)

  N_t : a "numeraire"

  Q^N : a martingale measure corresponding to N_t as numeraire.

$\longrightarrow$ If Π_t denotes the arbitrage-free price of an asset at time t, then

$$\Pi_t = N_t \, E^{Q^N} \left(\frac{\Pi_T}{N_T} \mid \mathcal{G}_t \right)$$

$\longrightarrow$ *For derivative pricing, $\Pi_T \in \mathcal{F}_T^S \subset \mathcal{G}_T$ with S an "underlying" primary asset having given dynamics.*

Factor Models

- Factor models are a convenient setup for many purposes; they are parsimonious and numerically tractable *(allow also to model dependence among different quantities)*
- Factors may represent (macro-)economic quantities that may or may not be observable *(interest rates, volatilities, value of a firm)*; they may also simply be abstract factors.

$\longrightarrow$ Model the factors as Markovian processes in $\mathcal{G}_t$ and assume them not to be directly observable.

Filtering in Markovian factor models

- Given a Markovian factor process X_t, claims as well as numeraires can in many situations be expressed as functions of X_t.

  $\longrightarrow$ Due the Markovianity of X_t one has in fact

$$\Pi_t = N_t \, E^{Q^N}\left(\frac{\Pi_T}{N_T} \mid \mathcal{G}_t\right) := \Pi(t, X_t)$$

(Π_t and N_t are functionals of future values of X_t)

Filtering in Markovian factor models

- *Let the investor filtration be* $\mathcal{Y}_t \subset \mathcal{G}_t \;\; (X_t \notin \mathcal{Y}_t)$

$\longrightarrow$ If $N_t \in \mathcal{Y}_t$ then $\hat{\Pi}_t = N_t \, E^{Q^N}\left(\frac{\Pi_T}{N_T} \mid \mathcal{Y}_t\right)$ is an arbitrage-free price in the filtration $\mathcal{Y}_t$ and one has

$$\hat{\Pi}_t = E^{Q^N}\left\{\Pi(t, X_t) \mid \mathcal{Y}_t\right\}$$

$\longrightarrow$ *Given* $\Pi(t, X_t)$*, to compute* $\hat{\Pi}_t$ *it thus suffices to have the* filter distribution $\pi(X_t \mid \mathcal{Y}_t)$.

Filtering in Markovian factor models

- It concerns thus typically a two-step procedure:

Step 1 Determine the quantities of interest under full information as instantaneous functions of the factors.

Step 2 *Derive the values under the actual market information corresponding to $\mathcal{Y}_t \subset \mathcal{G}_t$ by projecting the full information values on the subfiltration corresponding to market information.*

Filtering in Markovian factor models

- Filtering allows for a continuous updating of the filtered prices to the current investor information *("self-tuning" in engineering applications).*

- If $\Pi(t, X_t) \in \mathcal{Y}_t$ then $\hat{\Pi}_t = \Pi(t, X_t)$, i.e. the arbitrage-free filtered model is automatically calibrated to market prices.

  $\longrightarrow$ *In incomplete markets the underlying martingale measure can be determined by traditional calibration, but also (and in an adaptive way) by filtering the market price of risk.*

Filtering vs Calibration

- The model may contain parameters that need to be calibrated to market data *(even in the case when the market price of risk is filtered)*.
- Traditional calibration corresponds to an inverse problem that leads to a static *point estimation* without indication of the accuracy.

$\longrightarrow$ Filtering allows for a dynamic parameter estimation *(continuous successive updating)*.

Possibilities of calibration related to filtering

i) Combined filtering and parameter estimation;

ii) Expectation maximization *(can be naturally linked to filtering so that estimates evolve according to the filter solution; see part III)*;

iii) Maximization of the innovations likelihood *(partly dynamic)*;

iv) Others.

$\longrightarrow$ Combined filtering and parameter estimation:
Taking the *Bayesian point of view*, the parameter vector θ is considered a random variable with given prior distribution $\longrightarrow$ compute

$$\pi(X_t, \theta \mid \mathcal{Y}_t)$$

Modeling approaches

Existing credit risk models $\left\{ \right.$
- Structural models
- Reduced-from (intensity-based) models

- Structural models:

$$\begin{aligned} V_t &: \text{ asset value of the firm} \\ K_t &: \text{ default barrier} \\ \tau &= \inf\{t \geq 0 \mid V_t \leq K_t\} \end{aligned}$$

$\longrightarrow$ Default occurs at the first time when the asset value of the firm does not cover its liabilities *(predictable stopping time w.r.t. to the global filtration $\mathcal{G}_t$)*.

$\longrightarrow$ Since τ is predictable, structural models lead to unrealistic credit spreads.

Intensity-based model

A framework with a number *m* of defaultable firms

- τ_j: random time of default of firm j, $\quad j = 1, \cdots, m$;
- $H_t^j := \mathbf{1}_{\{\tau_j \leq t\}}$; $\quad H_t := (H_t^1, \cdots, H_t^m)$;
- $\mathcal{H}_t = \sigma\left(H_s, s \leq t\right)$.

Intensity-based model

- $(\mathcal{F}_t)_{0\le t\le T}$ a given *background filtration*. The underlying filtered probability space is $(\Omega, \mathcal{G}, \mathcal{G}_t, Q)$ with $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t\,,\ t \in [0, T]\,,\ \mathcal{G}_T = \mathcal{G}$ (*full information filtration*);
- Q martingale (pricing) measure, *numeraire*: $B(t) = B(0)\exp\left[\int_0^t r_s ds\right]$
- τ_i is directly modeled as a totally inaccessible stopping time w.r.to $\mathcal{G}_t$ with $(Q, \mathcal{G}_t)-$intensity $\lambda_i = (\lambda_{t,i})$ i.e. such that

$$H_t^i - \int_0^{t\wedge\tau_i} \lambda_{s,i} ds \quad \text{is a} \quad (Q, (\mathcal{G}_t))\text{-martingale.}$$

(Factors in) **Intensity-based Models**

- Assume given a common Markovian factor process $\mathbf{X}_t \in \mathbb{R}^d$ and let
$$\lambda_{t,i} = \lambda_i(\mathbf{X}_t)$$
  - $\longrightarrow$ Allows to model physical and information-induced dependence/contagion among the defaults.
  - $\longrightarrow$ A common modeling approach: Conditionally independent doubly stochastic default times

$$Q(\tau_i > t \mid \mathcal{F}_\infty^{\mathbf{X}}) = \exp(-\int_0^t \lambda_i(\mathbf{X}_s)ds), \quad t > 0$$

- The factors may include economic covariates, but also unobservable (abstract) factors: evidence is given in the recent literature (see Das et al 2007, Duffie et al 2009) that unobservable factors, driving the default intensities, are needed on top of observable co-variates to better explain clustering of defaults and large co-movements of credit spreads.

$\longrightarrow$ *Below, the entire factor process X_t will be considered as unobservable*.

Pricing under incomplete information on the factors

- Markovianity of the factors allows to follow the usual two-step procedure to price credit derivatives under incomplete information on the factors themselves:

> i) *Determine the derivative prices under the full information $\mathcal{G}_t$ as functions of $\mathbf{X}_t$;*
>
> ii) *Use stochastic filtering to "project" these prices onto the subfiltration corresponding to market information.*

$\longrightarrow$ In the case $\mathbf{X}_t \equiv \mathbf{X}$ this $\mathbf{X}$ is called "frailty" parameter (e.g. Schönbucher) and filtering of $\mathbf{X}_t$ reduces to Bayesian updating of $\mathbf{X}$.

Investor filtration (market information)

- Denote the investor filtration by $\mathcal{Y}_t \subset \mathcal{G}_t$. It is supposed to always contain the default history. It may also include (besides possible observable covariates) noisy observations of prices of credit risky assets as well as yields and spreads on default-free and defaultable bonds *(the latter are representative of more general market data)*

$\longrightarrow$ It can be shown (e.g. Jarrow-Protter, see also Guo et al) that the distinction between structural and reduced-form models is actually a distinction between full and partial observability of firm values and liabilities: τ_i *predictable w.r.to* $\mathcal{G}_t$ *becomes totally inaccessible w.r.to* $\mathcal{Y}_t$ *and it admits an* $\mathcal{Y}_t-$*intensity.*

Filtering: general

- It concerns a partially observable process (X_t, Y_t) where:

X_t : unobservable component *(known stochastic dynamics)*
Y_t : observations *(distribution of Y_t, given X_t, is known)*

$\longrightarrow$ Determine (recursively) the filter distribution $\pi(X_t \mid \mathcal{F}_t^Y)$.

$\longrightarrow$ If X_t is Gaussian and $f(Y_t \mid X_t)$ is also Gaussian *(e.g. linear Gaussian models)*, then $\pi(X_t \mid \mathcal{F}_t^Y)$ is Gaussian as well and characterized by its conditional mean and variance *(Kalman filter)*.

Linear-Gaussian models?

- Assume a linear-Gaussian model for the factors (e.g. mean reverting model) and let

$$\lambda_{t,i} = \lambda_i(\mathbf{X}_t) \quad \text{affine in } \mathbf{X}_t$$

$\longrightarrow$ It implies an affine credit risk model (e.g. Duffie-Garleanu) where bond prices are exponentially affine in $\mathbf{X}_t$.

- Consider observations that can be expressed in terms of log-bond prices.

$\longrightarrow$ *The filtering problem becomes linear-Gaussian*.

.

- Problems:
  - i) intensities may become negative *(factors though satisfy a mean reverting model and are filtered)*
  - ii) At a default of firm j the filter update is

$$\pi_{\mathbf{X}_t|\mathcal{Y}_t}(dx) = \frac{\lambda_j(x)\pi_{\mathbf{X}_t|\mathcal{Y}_{t-}}(dx)}{\int \lambda_j(x)\pi_{\mathbf{X}_t|\mathcal{Y}_{t-}}(dx)} \quad ; \quad t = \tau_j$$

→ Being $\lambda_j(x)$ affine in x, the Gaussianity of $\pi_{\mathbf{X}_t|\mathcal{Y}_{t-}}(dx)$ is destroyed *(can be preserved approximately by a Gaussian sum approximation)*.

→ In the differential of yields and credit spreads the volatility becomes a deterministic function of time *(not realistic)*.

- If however we let (see part III) $\mathbf{X}_t = \log \mathbf{Z}_t$ with $\mathbf{Z}_t$ a (multivariate) CIR process and let

$$\lambda_{t,i} = \lambda_i(\mathbf{X}_t) \quad \text{affine in } e^{\mathbf{X}_t} \text{ then}$$

$\longrightarrow$ Intensities are positive

$\longrightarrow$ at a default $\lambda_j(x)\pi_{\mathbf{X}_t|\mathcal{Y}_{t-}}(dx)$ preserves Gaussianity

- However: The filtering model is nonlinear Gaussian and to obtain a Gaussian filter distribution, the Extended Kalman Filter (EKF) has to be used leading to an approximation *(in general a very reliable approximation)*.

Pricing in a specific "affine" credit risk model

- *Interest rate* and *default intensity* are linear functions of the exponentials of the components of a stochastic factor process $\Psi_t \in \mathbb{R}^n$ such that:

$$d\Psi_t = diag\left(e^{-\Psi_t}\right)\left[Ae^{\Psi_t} + b - \frac{1}{2}\mathbf{1}\right]dt + diag(e^{\frac{1}{2}\Psi_t})dw_t$$

where w is an n-dimensional $(\mathcal{F}_t, Q)$-Wiener process.

$\longrightarrow \quad \Psi_t = \log \Phi_t$ *with* Φ_t *satisfying a CIR model.*

"Affine" credit risk model

- Putting

$$\Phi_t := \exp(\Psi_t)$$

by Itô's formula we obtain

$$d\Phi_t = (A\Phi_t + b)\, dt + diag\left(\sqrt{\Phi_t}\right) dw_t$$

namely Φ satisfies a multivariate CIR model in "canonical form" (in the terminology of Dai-Singleton 2000).

- Usual admissibility conditions *(Feller test for explosions, condition A in Duffie-Kan)*:
with $A = \{a_{ij}\}_{i,j=1,\dots,n}$, $b = (b_i)_{i=1,\dots,n}$ we require $a_{ij} \geq 0$ for $j \neq i$ and $b_i > \frac{1}{2}$, for $i = 1, \dots, n$.

  $\longrightarrow$ *Guarantees also existence and uniqueness of a strong solution.*

"Affine" credit risk model

- Define interest rate (r_t) and the default intensities (λ_t^j) as:

$$\begin{cases} r_t = a + b e^{\Psi_t} = a + b\Phi_t \\ \lambda_t^j = c^j + d^j e^{\Psi_t} = c^j + d^j \Phi_t \end{cases}$$

with a, c^j nonnegative constants and b, d^j n-dimensional row vectors of nonnegative constants.

$\longrightarrow$ *This setup implies positive rates even if Ψ_t is not restricted to be positive and allows for correlation between interest rate and default intensities, which (see Schoenbucher) is a desirable property for a stochastic credit risk model.*

Affine credit risk model; Bond prices

Default-free 0-coupon bond

$$\Pi_{DF}(t,T) = \mathbb{E}\left[e^{-\int_t^T r_s ds} | \mathcal{G}_t\right] = \exp\left[A(t,T) - B(t,T)\, e^{\Psi_t}\right]$$

Defaultable 0-coupon 0-recovery bond

$$\Pi(t,T) = E\left\{e^{-\int_t^T r_s ds} \mathbf{1}_{\tau > T} \mid \mathcal{G}_t\right\} = \mathbf{1}_{\tau > t} E\left\{e^{-\int_t^T (r_s + \lambda_s) ds} \mid \mathcal{Y}_t\right\}$$
$$= \mathbf{1}_{\tau > t} \exp\left[\tilde{A}(t,T) - \tilde{B}(t,T)\, e^{\Psi_t}\right]$$

- $A(t,T), B(t,T), \tilde{A}(t,T), \tilde{B}(t,T)$ satisfy ODEs with coefficients depending on those of the factor dynamics and in $\lambda_t^j = c^j + d^j e^{\Psi_t}$.

Affine credit risk model

Bond prices

- More generic credit-risky products, such as *corporate bonds* and *CDS spreads*, can be expressed by means of these two basic elements.
- Viceversa, a default-free and a defaultable term structure can be reconstructed from the more liquid *corporate bonds* prices and *CDS spreads (for the latter the link with default events is much clearer than for other products)*.

  $\longrightarrow$ $0-$*coupon default free bonds and* $0-$*coupon* $0-$*recovery defaultable bonds can be considered as “building blocks” for more complex instruments*.

Affine credit risk model

Yields and credit spreads

Yield of a 0-coupon default-free bond

$$YL(t,T) := -\frac{1}{T-t}\log \Pi_{DF}(t,T) = -\frac{A(t,T)}{T-t} + \frac{B(t,T)}{T-t}\, e^{\Psi_t}$$

Spread of a 0-coupon, 0-recovery defaultable bond w.r.to a default-free bond (same face value and maturity)

$$CS(t,T) := -\frac{1}{T-t}\log\left[\frac{\Pi(t,T)}{\Pi_{DF}(t,T)}\right]$$
$$= \frac{A(t,T) - \tilde{A}(t,T)}{T-t} + \frac{\tilde{B}(t,T) - B(t,T)}{T-t} e^{\Psi_t} \qquad t < \tau \wedge T$$

- *Yields* and *credit spreads* are affine functions of e^{Ψ_t}.

Incomplete information
The investor filtration

Assume, w.l.o.g., that all components of the factor process Ψ_t are unobservable (not precisely known).

- However, the investor can observe market data, in part. the *interest rate* (proxy), a number p of *yields* and a number q of *credit spreads*.
- The default indicator process (H_t) is indirectly contained in the *credit spreads*.

Investor filtration

$$\mathcal{Y}_t = \sigma\{r_s, YL(s, T_i), CS(s, T_j) : s \leq t, i = 1, \cdots, p; j = 1, \cdots, q\} \vee \mathcal{H}_t$$

and thus $\mathcal{H}_t \subset \mathcal{Y}_t \subset \mathcal{G}_t$.

Incomplete information
A filter-based pricing model

- Objective: evaluate (in the investor filtration) an *OTC* credit risky-product, the price of which under complete information is given by $\Pi(t, T; \Psi_t)$
- Main tool: filter distribution of Ψ_t with respect to the investor filtration $\mathcal{Y}_t$ and under the pricing measure Q

Price in the investor filtration

$$\hat{\Pi}(t, T) = \mathbb{E}\left[\Pi(t, T; \Psi_t) \,|\mathcal{Y}_t\right]$$

- $\hat{\Pi}(t, T)$ is an *arbitrage-free* price, since $r_t \in \mathcal{Y}_t$
- $\hat{\Pi}(t, T)$ is coherent with the observations of market data, since the latter are the input to the filtering problem

Incomplete information

Noise terms affecting the observations

- All observable processes are linear functions of the exponentials of the unobserved factors.
- Therefore, if $1 + p + q > n$ the values of the factors can be determined from the observations and so the filtering problem degenerates.

This setting is not very realistic: *yields and credit spreads* are reconstructed from *corporate bonds* and *CDS spreads* and are affected by bid-ask spread etc and , therefore, cannot be considered as perfectly observable.

Introduce (Gombani, Jaschke, R.-05) ℓ further unobserved factors, on which r_t and λ_t^j do not depend, but which represent additive noise terms affecting the observations $YL(t, T_i)$ and $CS(t, T_j)$ and such that $n + \ell > 1 + p + q$.

Incomplete information
The observation system

- Augment Ψ_t (dim. n) to $\Psi_t^* = (\Psi_t, \bar{\Psi}_t)$ (dim. $n+\ell$) by adding the ℓ noise factors to Ψ_t *(assumed to be independent* $(\mathcal{Y}_t, Q)-$*Brownian motions)*:

Observation system

$$\begin{cases} r_t & = & a + b e^{\Psi_t} & \\ YL(t, T_i) & = & \alpha_t^i + \beta_t^i e^{\Psi_t} + \bar{\beta}_t^i \bar{\Psi}_t & i = 1, \ldots, p \\ CS(t, T_j) & = & \gamma_t^j + \delta_t^j e^{\Psi_t} + \bar{\delta}_t^j \bar{\Psi}_t & j = 1, \ldots, q \end{cases}$$

where $\alpha_t^i, \gamma_t^j, \beta_t^i, \delta_t^j, \bar{\beta}_t^i, \bar{\delta}_t^i$ consist of deterministic functions of time that depend on the model parameters.

- Let

$$\mathcal{F}_t^Y := \sigma\{r_s, YL(s, T_i), CS(s, T_j) \ : s \leq t, i = 1, \cdots, p; j = 1, \cdots, q\}$$

so that $\mathcal{Y}_t = \mathcal{F}_t^Y \vee \mathcal{H}_t$

The filtering problem

- The filter distribution $\pi(\Psi_t^* \mid \mathcal{F}_t^Y)$ degenerates

  $\longrightarrow$ One can find a surrogate/auxiliary state process X_t, of lower dimension than Ψ_t^*, and solve equivalently the filtering problem for (X_t, Y_t).

  $\longrightarrow$ *For appropriate matrices Γ_t and Δ_t and appropriate μ_t one has in fact*

$$\Phi_t = \Gamma_t e^{X_t} + \Delta_t(Y_t - \mu_t)$$

- The choice of such a process X_t is not unique.

The filtering methodology

The pair (X_t, Y_t), satisfies a system of the form

$$\begin{cases} dX_t &= F(e^{X_t}, Y_t)\,dt + G(e^{X_t}, Y_t)dw_t + H(e^{X_t}, Y_t)d\bar{\Psi}_t \\ dY_t &= R(e^{X_t}, Y_t)\,dt + S(e^{X_t}, Y_t)dw_t + \bar{M}_t d\bar{\Psi}_t \end{cases}$$

with coefficients having a special structure as functions of (e^{X_t}, Y_t).

The filtering methodology

The system for (X_t, Y_t) is a nondegenerate nonlinear filter system to which the Extended Kalman Filter (EKF) can be applied leading to a Gaussian filter distribution for the factors.

$\longrightarrow$ Between default times one has

$$p_{X_t|\mathcal{F}_t^Y} = p_{X_t|\mathcal{Y}_t}$$

($\mathcal{Y}_t$ was defined as the “investor filtration”)

The filtering methodology

Filter at a default time

- Recall that we had assumed $\lambda_t^j = c^j + d^j e^{\Psi_t} = c^j + d^j \Phi_t$
- Put

$$\lambda_t^j(X_t, Y_t) = c^j + d^j\left(\Gamma_t e^{X_t} + \Delta_t(Y_t - \mu_t)\right) =: c_0^j(t) + \sum_{i=1}^{n} c_i^j(t)\, e^{X_t^i}$$

- Suppose that at $t = \tau_j$ one observes the default of firm j. Then

$$p_{X_t|\mathcal{Y}_t}(dx) = \frac{\lambda_t^j(x, Y_t)\, p_{X_t|\mathcal{F}_t^Y}(dx)}{\int \lambda_t^j(x, Y_t)\, p_{X_t|\mathcal{F}_t^Y} dx} \qquad \text{for} \quad t = \tau_j$$

$\longrightarrow$ *Gaussianity is thus preserved also at a default time and the only (reliable) approximation is due to the EKF.*

The filtering methodology
Remarks

- The price to be paid for having Gaussianity also at a default time is that, for each incoming Gaussian distribution, the outgoing distribution is a mixture of $\tilde{n}$ Gaussian distributions.

  $\longrightarrow$ *Parallel filters have to be run, one for each component of the mixture*.

- Between default times one has a continuous update of the “filtered default intensities”. At a default time they undergo a jump with size depending on the riskiness of the defaulted firm *(information induced contagion)*.

Parameter estimation and *EM* algorithm

General description of the *EM* algorithm

- Let θ be the vector of the model parameters

The *EM* algorithm is based on the iterative maximization, w.r.t θ for a fixed θ', of the following function Q:

$$Q(\theta, \theta') = \mathbb{E}_{\theta'}\left[\log \frac{dP^{\theta}}{dP^{\theta'}} | \mathcal{F}_t^Y\right]$$

The *EM* algorithm iterates through the two following steps

1. (*Expectation*): compute $Q(\theta, \theta')$ for given θ' (a conditional expected value)
2. (*Maximization*): maximize $Q(\theta, \theta')$ w.r.t θ

- The *maximization* step leads to a system of equations obtained by putting $\frac{\partial Q(\theta,\theta')}{\partial \theta} = 0$

Alternating iterative *EM* algorithm: an example

- Let $n = 3, p = 1, q = 2,$ (two defaultable issuers)
- $\Psi_t = (\Psi_t^1, \Psi_t^2, \Psi_t^3)$; $\Psi_t^* = (\Psi_t, \bar{\Psi}_t^1, \bar{\Psi}_t^2)$

$$\begin{cases} d\Psi_t^i & = \left[a^i + e^{-\Psi_t^i}\left(b^i - \frac{1}{2}\right)\right] dt + e^{\frac{1}{2}\Psi_t^i} dw_t^i, \quad (i = 1, 2, 3) \\ \bar{\Psi}_t^j & = \bar{w}_t^j, \quad (j = 1, 2) \end{cases}$$

- $w_t = (w_t^1, w_t^2, w_t^3, \bar{w}_t^1, \bar{w}_t^2)$ Wiener with independent components.
- For the rates we put

$$\begin{cases} r_t & = \Phi_t^1 + \Phi_t^2 \\ \lambda_t^A & = \lambda^A \left(\Phi_t^1 + \Phi_t^3\right) \\ \lambda_t^B & = \lambda^B \left(\Phi_t^2 + \Phi_t^3\right) \end{cases}$$

Alternating iterative *EM* algorithm: an example

- For the coefficients of the additional noise factors Ψ_t^j in the observation dynamics, namely in

$$\begin{cases} r_t & = & a + be^{\Psi_t} & \\ YL(t,T_i) & = & \alpha_t^i + \beta_t^i e^{\Psi_t} + \bar{\beta}_t^i \bar{\Psi}_t & i = 1,\dots,p \\ CS(t,T_j) & = & \gamma_t^j + \delta_t^j e^{\Psi_t} + \bar{\delta}_t^j \bar{\Psi}_t & j = 1,\dots,q \end{cases}$$

  assume

$$\bar{\beta}_t^1 = [v, 0] \quad , \quad \bar{\delta}_t^1 = [0, \rho^A] \quad , \quad \bar{\delta}_t^2 = [0, \rho^B]$$

  with (v, ρ^A, ρ^B) additional parameters to be estimated.

Alternating iterative *EM* algorithm: an example

- It turns out that one can choose as X_t any of the $\Psi_t^1, \Psi_t^2, \Psi_t^3$.
  - → This leads to three possible systems, each of which depends only on a single factor to be taken as the respective X_t.
  - → *In each one only some of the parameters are estimated*.
- Three further systems for the specific estimation of $(\lambda^A, \lambda^B, \nu, \rho^A, \rho^B)$.

Alternating iterative *EM* algorithm: an example

- The six systems are analogous to one another. As example consider the first one which is of the form

$$\begin{cases} d\Psi_t^1 & = & \left[a^1 + e^{-\Psi_t^1}\left(b^1 - \frac{1}{2}\right)\right] dt + e^{\frac{1}{2}\Psi_t^1} dw_t^1 \\ YL(t,T) & = & (\alpha_t + \beta_t r_t) + \gamma_t e^{\Psi_t^1} + v\,\bar{w}_t^1 \\ CS^A(t,T) & = & (f_t + g_t r_t - h_t CS^B(t,T)) + k_t e^{\Psi_t^1} + \rho\,\bar{w}_t^2 \end{cases}$$

$\longrightarrow$ Only $X_t = \Psi_t^1$ enters this system and the observations are $YL(t,T)$ and $CS^A(t,T)$. The coefficients depend on the various parameters as well as on $CS^B(t,T)$, with this system we estimate however only (a^1, b^1).

Alternating iterative *EM* algorithm: an example

- After a time discretization and linearization of the coefficients *around the most recent estimate of* Ψ_t^1 *(EKF)*, the system takes the form

$$\begin{cases} \Psi_{t+\Delta}^1 &= \bar{a}_t + \bar{b}_t \Psi_t^1 + \bar{c}_t Z_{t+\Delta}^1 \\ YL(t,T) &= \bar{\alpha}_t + \bar{\beta}_t r_t + \bar{\gamma}_t \Psi_t^1 + v\,\bar{w}_t^1 \\ CS^A(t,T) &= \bar{f}_t + g_t r_t + \bar{h}_t CS^B(t,T) + \bar{k}_t \Psi_t^1 + \bar{\rho} w_t^2 \end{cases}$$

with Z_t^1 an i.i.d. sequence of standard Gaussian random variables and where the “bar” over the coefficients indicates that they now depend also on the most recent estimate of Ψ_t^1.

Alternating iterative *EM* algorithm

Algorithm

0. Initialize the algorithm with a guess $\hat{\theta}$ for the entire vector θ and, setting $j = 0$, put $\theta_j = \hat{\theta}$;
1. Apply in parallel on each of the systems 1 and 6 the EM algorithm to estimate (a^1, b^1) and (λ^B, ρ^B) while keeping the other parameters fixed at their previously estimated values $(a_j^2, b_j^2, a_j^3, b_j^3, \lambda_j^A, \nu_j, \rho_j^A)$. The algorithm iterates through the two EM steps (expectation and maximization) until a stopping criterion is met, thereby producing estimates $(a_{j+1}^1, b_{j+1}^1, \lambda_{j+1}^B, \rho_{j+1}^B)$

Alternating iterative *EM* algorithm

Algorithm

2. Apply in parallel on each of the systems 2 and 5 the EM algorithm to estimate (a^2, b^2) and (λ^A, ρ^A) while keeping the other parameters fixed at their previously estimated values $(a^1_{j+1}, b^1_{j+1}, a^3_j, b^3_j, \lambda^B_{j+1}, \nu_j, \rho^B_{j+1})$. The algorithm iterates through the two EM steps until a stopping criterion is met, thereby producing estimates $(a^2_{j+1}, b^2_{j+1}, \lambda^A_{j+1}, \rho^A_{j+1})$;
3. Apply in parallel on each of the systems 3 and 4 the EM algorithm to estimate (a^3, b^3, ν) keeping all others parameters fixed at their previously estimated values. The algorithm iterates through the two EM steps until a stopping criterion is met, thereby producing estimates $(a^3_{j+1}, b^3_{j+1}, \nu_{j+1})$;

Alternating iterative *EM* algorithm
Algorithm

4. Put $\theta_{j+1} = (a^1_{j+1}, b^1_{j+1}, \ldots, a^3_{j+1}, b^3_{j+1}, \lambda^A_{j+1}, \lambda^B_{j+1}, \nu_{j+1}, \rho^A_{j+1}, \rho^B_{j+1})$ and, setting $j = j + 1$, return to step 1. Terminate the entire algorithm as soon as a global stopping criterion is met.

$\longrightarrow$ This is similar to the “Space Alternating Generalized EM” (SAGE) (see Fessler and Hero, 1994).

Alternating iterative *EM* algorithm

Simulation results

- Maturity T = 10 years, both for *default-free* and *defaultable bonds*
- $\Delta = 0.02$ (weekly observations)
- For “true” θ generate $(\Psi_k^1, \Psi_k^2, \Psi_k^3)$ for $k = 0, \cdots, 500$. Draw τ^A, τ^B and generate an observation sequence.
- Apply the Algorithm with θ_0 generated randomly.
- Stop individual iterations as soon as the difference between successive values of all the parameters $< 10^{-5}$ up to a maximum of 500.

Alternating iterative *EM* algorithm

Simulation results

Parameter	True value	Estimate	Std. dev.
a^1	-0.15	-0.15092	0.02980
b^1	0.60	0.60141	0.05163
a^2	-0.20	-0.19960	0.02529
b^2	0.70	0.69943	0.04205
a^3	-0.25	-0.24826	0.06377
b^3	0.80	0.80812	0.09398
λ^A	0.10	0.09981	0.02204
λ^B	0.30	0.30996	0.02533
ν	0.005	0.00509	0.00053
ρ^A	0.01	0.01001	0.00055
ρ^B	0.02	0.01978	0.00106

Table: Means and standard deviation of the estimates from 50 independent runs of the algorithm).

An extension of the model

Risk premia as further unobserved factors and *rating*-based information

- Consider the model under the physical-historical probability measure.
- Specify the *risk-premia*, which characterize the change of measure, as further unobserved stochastic processes to be included in the filtering system.
- In this setting we can consider also the information coming from the *rating scores*, which represent historical information.
- We can compute filtered estimates of default probabilities, on the basis of the information deriving from both the financial market and the *rating score*.

  $\longrightarrow$ C.Fontana, "Credit risk and incomplete information: a filtering framework for pricing and risk management. Preprint 2010.

Thank you for your attention

Appendix

Intensity-based model

The *default intensity* process

- $\mathcal{G}_t$-intensity λ_t^j of the $\mathcal{G}_t$-stopping time τ_j:

$$M_t^j := H_t^j - \int_0^{t\wedge\tau_j} \lambda_s^j ds \quad \text{is a } (P, \mathcal{G}_t)\text{-martingale}$$

- Assume τ_j conditionally independent, doubly stochastic random times with respect to $\mathcal{F}_t$

$$\downarrow$$

- λ_t^j is given by the $\mathcal{F}_t$-conditional hazard rate process of τ_j (McNeil-Frey-Embrechts-05) ($\mathcal{F}_t-$intensity of τ_j in Bielecki-Jeanblanc-Rutkowski-04).

In fact, according to the definition, for $\lambda_t^j \in \mathcal{F}_t$ and for independent exponential random variables ξ^j (parameter $= 1$ and independent of $\mathcal{F}_\infty$) one has

$$\tau_j = \inf\left\{ t \geq 0 \ : \ \int_0^t \lambda_s^j ds \geq \xi^j \right\}.$$

Consequently (with $T > t$)

$$P\{\tau_j > T|\mathcal{G}_t\} = \mathbf{1}_{\{\tau_j > t\}} \mathbb{E}\left[e^{-\int_t^T \lambda_s^j ds} | \mathcal{F}_t \right]$$

and

$$P\left\{\tau_j \in (t, T] | \mathcal{G}_t\right\} \underset{(T \to t^+)}{\approx} \lambda_t^j (T - t)$$

return