Contagious default: application of methods of Statistical Mechanics in Finance

Wolfgang J. Runggaldier

University of Padova, Italy

www.math.unipd.it/runggaldier

based on joint work with :

Paolo Dai Pra, Elena Sartori (University of Padova) and Marco Tolotti (SNS Pisa and Bocconi University Milan)

20th Anniversary of the Department of Pure and Applied Mathematics

Outline

- *The financial problem :* credit risk and the modeling of contagion

- *The interacting particle system model*

- *The main results for the particle system*

  i) Asymptotics when the number of particles $N \to \infty$
  ii) Equilibria of the limiting dynamics
  iii) Finite volume approximations

- *Back to Finance :* large portfolio losses in a credit risky environment with contagion and default clustering.

The financial problem : Credit risk

- Risk faced by a financial institution holding a portfolio of positions issued by a (large) number of firms that may default.

→ *Problem compounded by the fact that :*

i) *Default* may be *contagious*
ii) There may be a *clustering of defaults*

Losses may therefore be large and we want to address this problem in the above context *(contagion and clustering)*.

General modeling context

- *Default can be seen from two viewpoints :*

i) From the management of a firm that has more detailed information about the financial health of the firm and may thus view default through the mechanism that leads to it *(structural approach)*

ii) From the market to which the default may come as a "surprise" *(reduced-form or intensity-based approach)*.

$\longrightarrow$ Here we follow the latter (modeling default as a *point process*).

Point processes and intensities

- A *point process* describes events *(here : the default of a given firm)* that occur randomly over time.

  It can be described in two equivalent ways :

  i) $0 = T_0 < T_1 < T_2 < \cdots$
  with T_n the $n-$th instant of occurrence of the event;

  ii) $N_t = n$ if $t \in [T_n, T_{n+1}) \quad \leftrightarrow \quad N_t = \sum_{n \geq 1} 1_{T_n \leq t}$
  which is the associated counting process.

- A **Poisson process** is a point process where $N_t - N_s$ is Poisson distributed with parameter $\int_s^t \lambda_u du$

  $\rightarrow$ λ_t : *intensity* of the Poisson process;

  $\rightarrow$ $E\{N_t\} = \int_0^t \lambda_u du$ *(the larger the intensity, the more events can be expected in a given time interval)*.

- The *intensity* may itself be a *stochastic process (doubly stochastic Poisson process)* and it is characterized by

  $N_t - \int_0^t \lambda_u du$ a zero-mean martingale $\rightarrow E\{N_t\} = \int_0^t E\{\lambda_u\} du$

Contagion *(interacting intensities)*

- *To describe propagation of financial distress in a network of firms linked (directly or indirectly) by business relationships one possibility is via interacting intensities.*

  $\longrightarrow$ A natural way to obtain interacting intensities is to let the default intensities depend on a *common exogenous macroeconomic factor process* X_t, i.e. for the generic $j-$th firm one postulates

$$\lambda_t^j = \lambda^j(X_t)$$

- Given $\lambda_t^j = \lambda^j(X_t)$

  i) If X_t is observable and has jumps in common with the point process counting the defaults $\longrightarrow$ *direct contagion (counterparty risk)*

  ii) If X_t is unobservable, but its distribution is successively updated on the basis of the observed default history $\longrightarrow$ *information induced contagion*.

$\longrightarrow$ *Interacting intensity models are currently those mostly investigated and they are motivated by the empirical observation that default intensities are correlated with macroeconomic factors.*

- However (quoting from Jarrow and Yu (2001)):

  "A default intensity that depends linearly on a set of smoothly varying (exogenous) macroeconomic variables is unlikely to account for the clustering of defaults around an economic recession".

- Furthermore, one might also want to describe the general health of a network of firms by *endogenous financial indicators* thereby viewing *a credit crisis as a microeconomic phenomenon* and so possibly also arrive at explaining *default clustering*.

  → *Interacting particle system models from Statistical mechanics may allow to adequately address the above issues.*

The interacting particle system model

- *A mean-field interacting model of the Curie-Weiss type*; a simple model to describe dynamically the credit quality of firms.

- *The "credit state" of each firm is identified by two variables* (σ, ω) $((\sigma_i, \omega_i)$: *state of* $i-$*th firm* $i = 1, \cdots, N$*)*.

  → σ : a *"rating class indicator"* *(a low value reflects a bad rating class, i.e. a higher probability of not being able to pay back obligations)*.

  → ω : a *"liquidity indicator" or "sign of the cash balances"* *(a more fundamental indicator of the financial health of the firm; it is typically not directly observable from the market)*.

- At a first level assume $(\sigma_i, \omega_i) \in \{-1, +1\}^2$ *(generalization to a generic finite number of possible values rather straightforward)*

- No explicit "default state" (could be $\sigma_i = -1$).
  Always need a positive probability that the firm can exit from the state where σ_i takes its lowest possible value.

- For the time evolution on a generic interval $[0, T]$ of the "state" of the particle system, i.e. $(\sigma_i(t), \omega_i(t))_{i=1,\cdots,N} \in \mathcal{D}^{2N}[0, T]$ we need to specify the stochastic dynamics for the transitions $\sigma_i \rightarrow -\sigma_i, \; \omega_i \rightarrow -\omega_i$.

- *The mean-field assumption leads to letting the interaction depend on the global health indicator (endogenous global factor)*

$$m_N^{\sigma}(t) := \frac{1}{N} \sum_{i=1}^{N} \sigma_i(t)$$

- The vehicle of interaction/contagion is given by

$$\underset{fundam.\ indic.}{\omega_i} \quad \rightarrow \quad \underset{rating\ class}{\sigma_i} \quad \rightarrow \quad \underset{global\ health\ indic.}{m_N^{\sigma}} \quad \rightarrow \omega_j$$

Continuous time finite state Markov processes

- *A simple example. A point process (in form of counting process): at an event time T_n only a transition $n-1 \to n$ can take place.*

- *More generally :* consider a system that may be in one of a finite number of possible states, making transitions from one state to another at random times. Assume that, given an ordered pair of states, the transition from one to the other corresponds to a Poisson process and that these Poisson processes are independent *(a transition from one state to another is independent of the past)* $\longrightarrow$ *finite state Markov process*

  $\longrightarrow$ *Instead of a single transition intensity λ, a matrix of transition intensities*.

Transition intensities for the particle system

$$\begin{cases} \sigma_i \to -\sigma_i & with\ intensity\ \ \lambda_i\ := e^{-\beta\sigma_i\omega_i}, \quad \beta > 0 \\ \omega_j \to -\omega_j & with\ intensity\ \ \mu_j\ := e^{-\gamma\omega_j m_N^\sigma}, \quad \gamma > 0 \end{cases}$$

β, γ are parameters indicating the strength of the interaction.

$\longrightarrow$ The resulting transition intensity matrix can be taken as *infinitesimal generator* L of a *continuous-time Markov chain* with state space $\{-1,+1\}^{2N}$ that acts on $f : \{-1,1\}^{2N} \to \mathbb{R}$ as

$$Lf(\sigma,\omega) = \sum_{i=1}^{N} \lambda_i \nabla_i^\sigma f(\sigma,\omega) + \sum_{j=1}^{N} \mu_j \nabla_j^\omega f(\sigma,\omega)$$

where $\nabla_i^\sigma f(\sigma,\omega) = f(\sigma^i,\omega) - f(\sigma,\omega)\,; \quad \nabla_j^\omega f(\sigma,\omega) = f(\sigma,\omega^j) - f(\sigma,\omega)$ and $\sigma^i = (\sigma_1, ..., \sigma_{i-1}, -\sigma_i, \sigma_{i+1}, ..., \sigma_N)$; analogously for ω^j.

Dynamics of the system for large N

- Unlike many mean field models in Statistical mechanics our *model is non-reversible*.

$\longrightarrow$ *An explicit formula for the stationary (in time) distribution is not available.*

$\longrightarrow$ We then shall rather

A. Look for the *limit* $(N \to \infty)$ *dynamics* of the system on the path space *(via a LLN based on a Large Deviations Principle)*;

B. Study the *equilibria of the limiting dynamics*;

C. Describe *"finite volume approximations"* *(for large but finite* N*) via a Central Limit type result.*

A. Limit for $N \to \infty$ (Law of Large Numbers)

- Let $(\delta_{\{\cdot\}}$ denotes the Dirac measure)

$$\rho_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\{\sigma_i[0,T],\omega_i[0,T]\}}$$

be the sequence of empirical (random) measures on the space $\mathcal{M}_1(\mathcal{D}^2[0,T])$ endowed with the weak convergence topology.

- For a probability measure $q \in \mathcal{M}_1(\{-1,1\}^2)$ let

$$m_q^\sigma := \sum_{\sigma,\omega=\underline{+}1} \sigma\, q(\sigma,\omega)$$

(expected health under q).

Theorem 1. Let $(\sigma(t), \omega(t))$ be the Markov process corresponding to the generator $Lf(\sigma, \omega)$ and with initial distribution s.t. $(\sigma_i(0), \omega_i(0))$, $i = 1, \cdots, N$ are i.i.d. with law ℓ.

i) There exists $Q^* \in \mathcal{M}_1(\mathcal{D}^2[0,T])$ s.t. $\rho_N \to Q^*$ a.s. in the weak topology;

ii) if $q_t \in \mathcal{M}_1(\{-1,1\}^2)$ is the marginal distribution of Q^* at time t, then it is the unique solution of the *McKean-Vlasov equation (MKV)*

$$\begin{cases} \frac{\partial q_t}{\partial t} & = & \mathcal{L} q_t \,, \quad t \in [0,T] \\ q_0 & = & \ell \end{cases}$$

with $\mathcal{L} q(\sigma,\omega) = \nabla^\sigma \left[e^{-\beta\sigma\omega} q(\sigma,\omega) \right] + \nabla^\omega \left[e^{-\gamma\omega m_q^\sigma} q(\sigma,\omega) \right]$

B. Large time behavior of the limiting $(N \to \infty)$ *dynamics*

- A measure μ on $\{-1,1\}^2$ is completely specified by

$$m_\mu^\sigma := \sum_{\sigma,\omega=\pm 1} \sigma\,\mu(\sigma,\omega),\, m_\mu^\omega := \sum_{\sigma,\omega=\pm 1} \omega\,\mu(\sigma,\omega),\, m_\mu^{\sigma\omega} := \sum_{\sigma,\omega=\pm 1} \sigma\omega\,\mu(\sigma,\omega)$$

Write $m_t^\sigma = m_{q_t}^\sigma$ (analogously for m_t^ω, $m_t^{\sigma\omega}$)

- *(MKV) can be reduced to determining a solution of*

$$(\dot m_t^\sigma, \dot m_t^\omega) = V(m_t^\sigma, m_t^\omega) \qquad \text{(mkv)} \quad \text{with}$$

$$V(x,y) := (2\sinh(\beta)y - 2\cosh(\beta)x, 2\sinh(\gamma x) - 2y\cosh(\gamma x))$$

$\rightarrow$ *To analyze in (MKV) equilibria and their stability it suffices to analyze (mkv)*

Theorem 2.

i) Suppose $\gamma \leq \frac{1}{\tanh(\beta)}$. Then equation (mkv) has $(0,0)$ as a unique equilibrium solution, which is *globally asymptotically stable*, i.e. for every initial condition (m_0^σ, m_0^ω), we have

$$\lim_{t\to+\infty}(m_t^\sigma, m_t^\omega) = (0,0).$$

ii) For $\gamma < \frac{1}{\tanh(\beta)}$ the equilibrium $(0,0)$ is *linearly stable*, i.e. $DV(0,0)$ *(the Jacobian matrix)* has strictly negative eigenvalues. For $\gamma = \frac{1}{\tanh(\beta)}$ the linearized system has a neutral direction, i.e. $DV(0,0)$ has one zero eigenvalue.

iii) For $\gamma > \frac{1}{\tanh(\beta)}$ the point $(0,0)$ is still an equilibrium for (mkv), but it is a saddle point for the linearized system, i.e. the matrix $DV(0,0)$ has two nonzero real eigenvalues of opposite sign. Moreover (mkv) has *two linearly stable solutions* (m_*^σ, m_*^ω), $(-m_*^\sigma, -m_*^\omega)$, where m_*^σ is the unique strictly positive solution of the equation

$$x = \tanh(\beta)\tanh(\gamma x),$$

and

$$m_*^\omega = \frac{1}{\tanh(\beta)} m_*^\sigma$$

iv) For $\gamma > \frac{1}{\tanh(\beta)}$, the *phase space* $[-1,1]^2$ *is bi-partitioned* by a smooth curve Γ containing $(0,0)$ such that $[-1,1]^2 \setminus \Gamma$ is the union of two disjoint sets Γ^+, Γ^- that are open in the induced topology of $[-1,1]^2$. Moreover

$$\lim_{t \to +\infty} (m_t^\sigma, m_t^\omega) = \begin{cases} (m_*^\sigma, m_*^\omega) & \text{if } (m_0^\sigma, m_0^\omega) \in \Gamma^+ \\ (-m_*^\sigma, -m_*^\omega) & \text{if } (m_0^\sigma, m_0^\omega) \in \Gamma^- \\ (0,0) & \text{if } (m_0^\sigma, m_0^\omega) \in \Gamma. \end{cases}$$

- The fact that the limiting $(N \to \infty)$ dynamics may have multiple stable equilibria implies that our system exhibits what is called *phase transition*.

  $\to$ *One obtains different domains of attraction corresponding to each of the stable equilibria.*

  $\to$ The *effects of phase transition* for the system with finite N can be seen on *different time scales* as follows:

i) *On a long time-scale,* of the order of a power of e^N, one may observe what in Statistical mechanics is referred to as *metastability:*

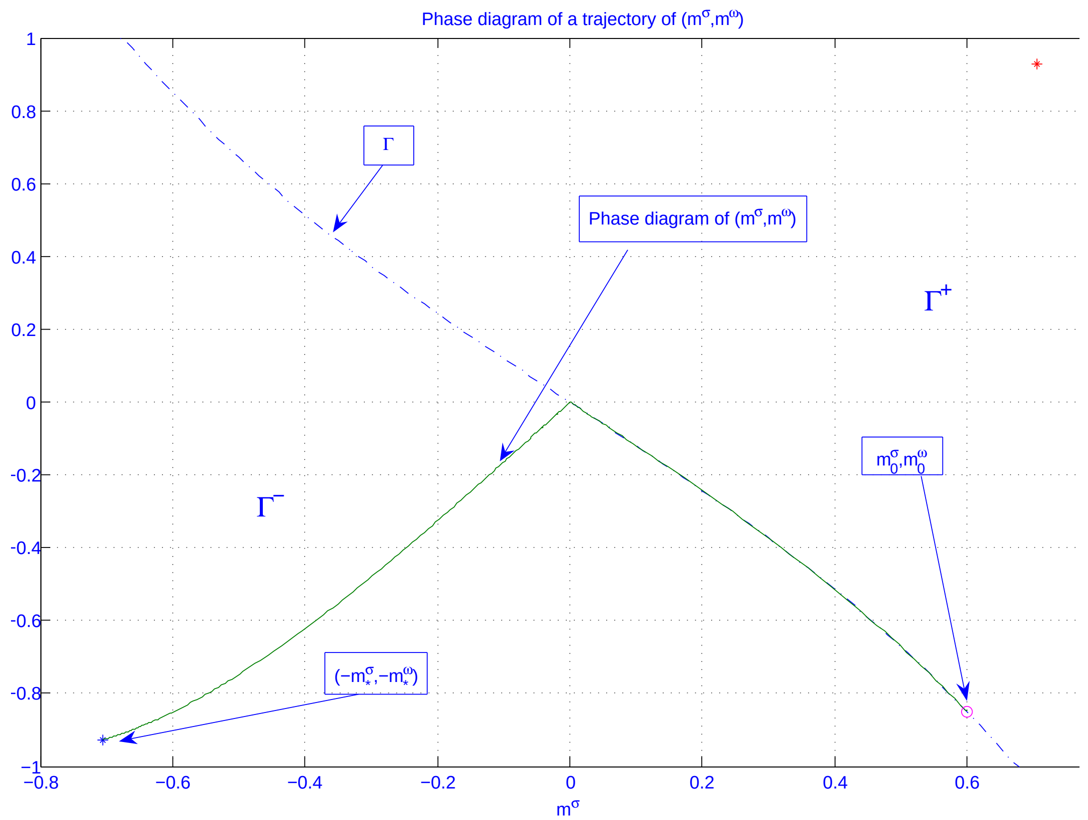
The system may spend a very long time in a small region of the state space around a stable equilibrium of the limiting dynamics and then switch relatively fast to another region around a different stable equilibrium.

$\longrightarrow$ the time scale is too large to be of interest in financial applications.

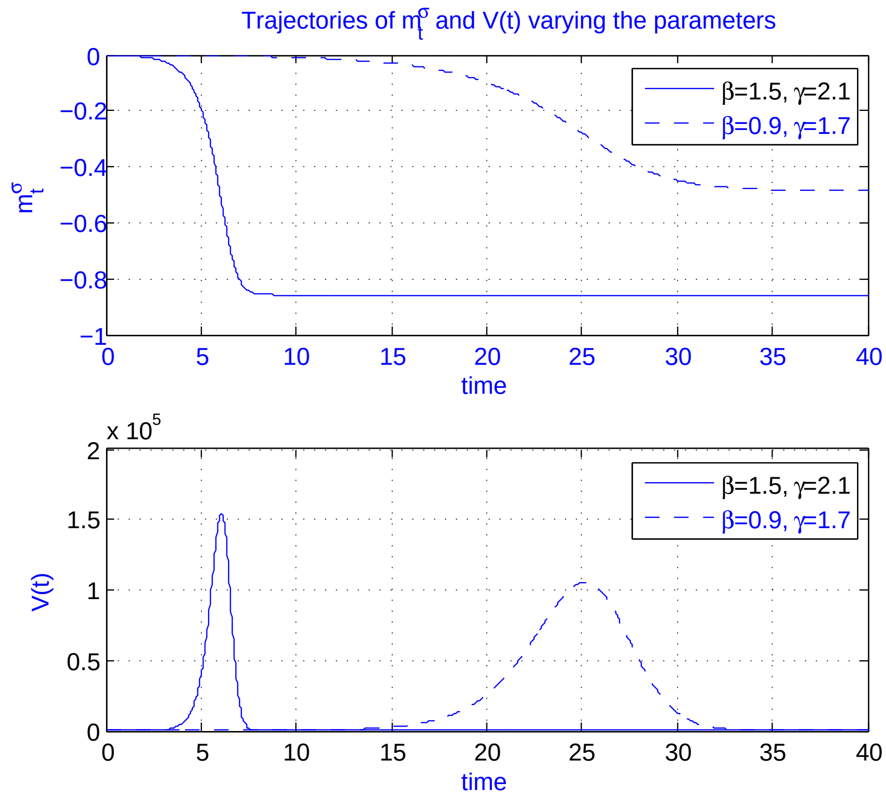
ii) *On a time-scale of order O(1)* the following occurs : *for certain values of the initial condition the system is driven towards the asymptotic symmetric neutral equilibrium state* $(0, 0)$ *where half of the firms are in good financial health.*

After a certain time (depending on the initial condition) the system is captured by an unstable direction of this neutral equilibrium and moves towards a stable asymmetric equilibrium. *During this transition the volatility of the system (will be defined below) increases sharply* before decaying to a stationary value.

$\longrightarrow$ *This phenomenon can be interpreted as a credit crisis and may account for default clustering.*



Phase diagram of a trajectory of (m^σ,m^ω)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
−1
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
m^σ
Γ
Phase diagram of (m^σ,m^ω)
Γ^+
Γ^−
m^σ_0,m^ω_0
(−m^σ_*,−m^ω_*)



Trajectories of m_t^σ and V(t) varying the parameters
m_t^σ
0
−0.2
−0.4
−0.6
−0.8
−1
0
5
10
15
20
25
30
35
40
time
β=1.5, γ=2.1
β=0.9, γ=1.7
x 10^5
2
1.5
1
0.5
0
V(t)
0
5
10
15
20
25
30
35
40
time
β=1.5, γ=2.1
β=0.9, γ=1.7

C. Analysis of the fluctuations

- Concerns the asymptotic distribution of $(\rho_N - Q^*)$.

  $\rightarrow$ *Recall that $\rho_N(t)$, being a measure on $\{-1,1\}^2$, is characterized by*

$$m^{\sigma}_{\rho_N}(t),\;\; m^{\omega}_{\rho_N}(t),\;\; m^{\sigma\omega}_{\rho_N}(t)$$

- With $A(t), D(t)$ appropriate matrices depending on β, γ and $m^{\sigma}_t, m^{\omega}_t, m^{\sigma\omega}_t$ one has the following

Theorem 3. Let

$$\begin{cases} x_N(t) &=& \sqrt{N}\left(m^{\sigma}_{\rho_N}(t) - m^{\sigma}_t\right) \\ y_N(t) &=& \sqrt{N}\left(m^{\omega}_{\rho_N}(t) - m^{\omega}_t\right) \\ z_N(t) &=& \sqrt{N}\left(m^{\sigma\omega}_{\rho_N}(t) - m^{\sigma\omega}_t\right) \end{cases}$$

Then $(x_N(t), y_N(t), z_N(t)) \overset{N\to\infty}{\longrightarrow} (x(t), y(t), z(t))$ in the sense of weak convergence of stochastic processes, where $(x(t), y(t), z(t))$ is a centered Gaussian process, unique solution of the linear SDE

$$\begin{pmatrix} dx(t) \\ dy(t) \\ dz(t) \end{pmatrix} = A^*(t) \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} dt + D(t) \begin{pmatrix} dB_1(t) \\ dB_2(t) \\ dB_3(t) \end{pmatrix}$$

where B_1, B_2, B_3 are independent Brownian motions and $(x(0), y(0), z(0))$ is a centered Gaussian.

$\longrightarrow$ *The asymptotic, for $N \to \infty$, distribution of $(x_N(t), y_N(t), z_N(t))$ is thus, for each fixed t, a centered Gaussian with covariance matrix Σ_t - the volatility referred to earlier - satisfying (asymptotics in t depend upon γ)*

$$\frac{d\Sigma_t}{dt} = A(t)\,\Sigma_t + \Sigma_t\,A^*(t) + DD^*(t)$$

Corollary 1: $\sqrt{N}\ \left[m^{\sigma}_{\rho_N}(t) - m^{\sigma}_t\right] \ \xrightarrow{D}\ \mathcal{N}(0, \Sigma^x_t)$ so that (notice that $m^{\sigma}_N(t) = m^{\sigma}_{\rho_N}(t)$)

$$P(m^{\sigma}_N(t) \geq \alpha) \approx \Phi\left(\frac{\sqrt{N}m^{\sigma}_t - \sqrt{N}\alpha}{\sqrt{\Sigma^x_t}}\right)$$

$(\Phi(\cdot)$ cumulative standard Gaussian).

Portfolio losses

- A bank holds a portfolio of financial positions issued by the N firms.

- *Random loss* for the $i-th$ position at time t:

$$L_i(t) \in \mathbb{R}^+ ; \quad i = 1, ..., N$$

- *Aggregated losses* are $L^N(t) = \sum_{i=1}^{N} L_i(t)$

- More specifically, let

$$G_x(u) := P\{L_i(t) \leq u \mid \sigma_i(t) = x\}\,, \quad x \in \{-1, +1\}$$

(*homogeneity with respect to* i *and* t) and

$$\ell_1 := E\{L_i(t) \mid \sigma_i(t) = 1\} < E\{L_i(t) \mid \sigma_i(t) = -1\} := \ell_{-1}$$

$\rightarrow$ *one expects to loose more when in financial distress.*

Furthermore,

$$v_1 := Var\{L_i(t) \mid \sigma_i(t) = 1\}; v_{-1} := Var\{L_i(t) \mid \sigma_i(t) = -1\}$$

Example 1

- Portfolio consisting of N positions of 1 unit due at time T *(defaultable bonds)*.

$$L_i(T) = L(\sigma_i(T)) = \begin{cases} 1 & if \quad \sigma_i(T) = -1 \\ 0 & if \quad \sigma_i(T) = 1 \end{cases}$$

$$\longrightarrow \qquad L^N(T) = \sum_{i=1}^{N} \frac{1-\sigma_i(T)}{2} \; = \; \frac{N(1-m_N^{\sigma}(T))}{2}$$

$$\longrightarrow \qquad P\{L^N(T) \geq \alpha\} = P\left\{m_N^{\sigma}(T) \leq 1 - \frac{2\alpha}{N}\right\}$$

apply Corollary 1

A further result

- Let

$$L(t) := \frac{(\ell_1 - \ell_{-1})}{2} m_t^\sigma + \frac{(\ell_1 + \ell_{-1})}{2}$$

$$V(t) := \frac{(\ell_1 - \ell_{-1})^2 \Sigma_t^x}{4} + \frac{(1 + m_t^\sigma)\, v_1}{2} + \frac{(1 - m_t^\sigma)\, v_{-1}}{2}$$

Theorem 4: When the distribution of $L_i(t)$ depends on $\sigma_i(t)$,

$$\sqrt{N}\left(\frac{L^N(t)}{N} - L(t)\right) \xrightarrow{D} \mathcal{N}\left(0, V(t)\right)$$

Corollary 2: In the setting of Theorem 4 it follows

$$P\left\{L^N(T) \geq \alpha\right\} \sim \Phi\left(\frac{NL(T) - \alpha}{\sqrt{N}\sqrt{V(T)}}\right)$$

Example 2 (Bernoulli mixture model)

- As before but with

$$L_i(T) = L(\sigma_i(T); \Psi) = \begin{cases} 1 & \textit{with prob} \quad P(\sigma_i(T); \Psi) \\ \\ 0 & \textit{with prob} \quad 1 - P(\sigma_i(T); \Psi) \end{cases}$$

where Ψ is an exogenous random factor.

$\rightarrow$ $\ell_1 = P(1; \Psi), v_1 = P(1; \Psi)(1 - P(1; \Psi))$ (analogously for ℓ_{-1}, v_{-1})

- *A possible specification is*

$$P(\sigma; \Psi) = 1 - \exp\{-k_1 \Psi - k_2 (1 - \sigma)/2 - k_3\}$$

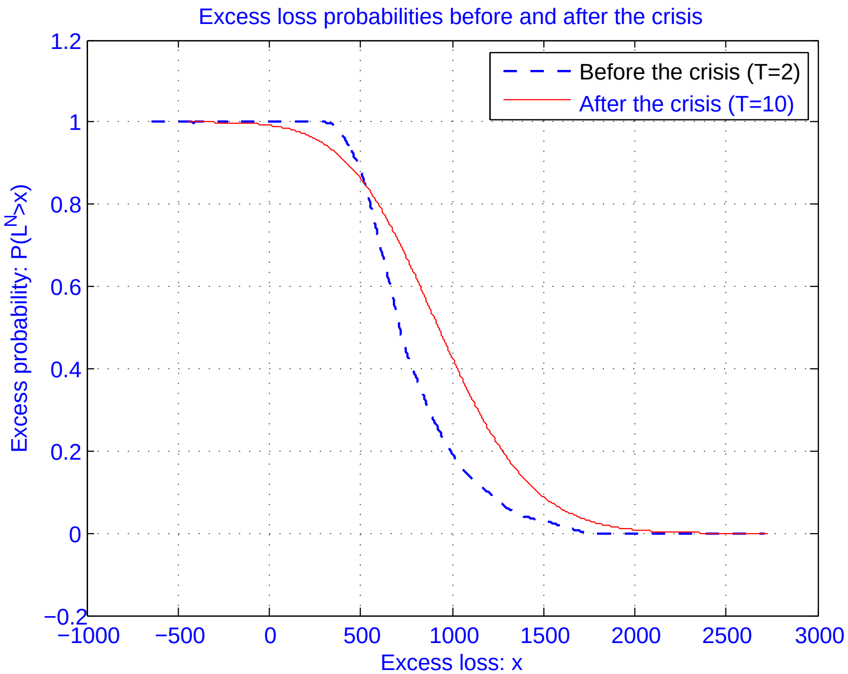
with $k_i \geq 0$ *and* $\Psi \sim \Gamma(\alpha; \kappa)$. *(The prob. for* $L_i(T) = 1$ *is bigger for* $\sigma_i(T) = -1$ *than for* $\sigma_i(T) = 1$*).*

- *Here* $\ell_1, \ell_{-1}, v_1, v_{-1}$ *and thus also* $L(t)$ *and* $V(t)$ *depend on the value* ψ *taken by the Gamma-type r.v.* Ψ*. Denote the latter by* $L(t;\psi)$, $V(t,\psi)$.

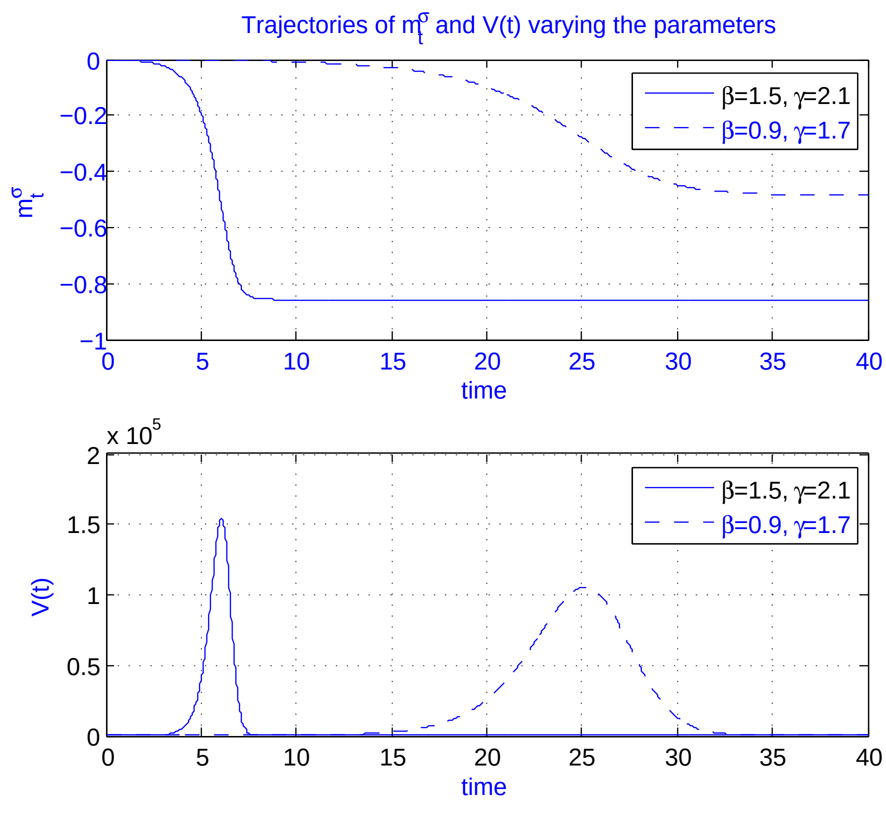
$\longrightarrow$ *by Corollary 2*

$$P\left\{L^N(T) \geq \alpha\right\} \sim \int \Phi\left(\frac{N\,L(T;\psi) - \alpha}{\sqrt{N}\sqrt{V(T;\psi)}}\right) df_\Psi(\psi)$$

with $f_\Psi(\cdot)$ the Gamma-density of Ψ.



Excess loss probabilities before and after the crisis
1.2
1
0.8
0.6
0.4
0.2
0
−0.2
−1000
−500
0
500
1000
1500
2000
2500
3000
Excess loss: x
Excess probability: P(L^N>x)
Before the crisis (T=2)
After the crisis (T=10)



Trajectories of m_t^σ and V(t) varying the parameters
m_t^σ
0
−0.2
−0.4
−0.6
−0.8
−1
0
5
10
15
20
25
30
35
40
time
β=1.5, γ=2.1
β=0.9, γ=1.7
x 10^5
2
1.5
1
0.5
0
V(t)
0
5
10
15
20
25
30
35
40
time
β=1.5, γ=2.1
β=0.9, γ=1.7