# Operational risk management

# and new computational needs in banks.

Abstract:

Basel II banking regulation introduces new needs for computational schemes. Indirectly they will involve both optimal stochastic control, and large scale simulations of decision processes of preventing low-frequency high loss-impact events.

Different approaches can be envisaged to derive sound strategies, from budgetary policy to capital allocation policy from a HJB equation. While this complex, high dimensional problem can be resolved by usual sophisticated methods (taking Galerkin approach, imposing Merton form solutions, finding viscosity solutions, or using ad-hoc utility functions that provide closed form solutions, etc.), the main interest of this model lies in exploring the scenarios in an adaptive learning framework (MDP, partially observed MDP, Q-learning, neuro-dynamic programming, greedy algorithm, etc.).

The benefits of such a computational approach is a consistent strategy to managing budgets, as opposed to policies of operational risk management made up from disconnected expenditures.

Keywords: Optimal Control, Modeling techniques, Operational risks, Management modeling, HJB equation, Reinforcement Learning.

## INTRODUCTION

Losses qualify as operational loss if they result from failed or inadequate systems, process or people, or from external sources. 7 loss categories are formally defined ranging from terrorist strikes or internal fraud to virus attacks and employment policy lawsuits. The New Basel Accord will require large international banks to evaluate through a model their own operational risks, then hold capital against them. This passive provisioning against loss causes great opportunity costs in banks and is a strong motivation for developing statistical models. The regulatory approach does not however provide motivation or obligation on modeling the active side of bank management, notably for modeling investments in risk prevention, or advances in risk reduction skills of managers. This other aspect of modeling *optimal use of investment capital* for operational risk reduction is an issue ignored by most practitioners in banking and regulatory circles. As a consequence, no actual proposed approach provides a framework that can discriminates in quantitative terms, consistent choices from arbitrary choices in operational risk budget management.

First, behaviour of choosing actions in a probabilistic reward space will be described. Then, models of small frequent risks and rare catastrophic events are introduced as to describe stochastic wealth accumulation. Budgets for improving internal controls on risk, and insurance expenditure are next modelized as stochastic controls. Resolution of the ensuing HJB equation in classical ways is next discussed along with new methods involving Adaptive Learning.

## 1. Modeling managerial risk reduction actions through budgets

Risk management is primarily about adopting coherent and optimal policies (hold cash and increase resiliency) to face potential disaster. It tries to draw from a global vision of the organisation, the environment, and the means at disposal, to derive intelligent behaviour while facing risks. Behaviour will be defined here as a set of actions linked together by the same rationale, motivations or preferences, either economic or psychological.

In managing operational risk, focus are on important questions such as :

- How does senior management define a sound and coherent policy ?
- How does middle management make sure its actions are efficient in reducing risk over time, and stay on course in the risk mitigation project ?
- And how does Capital allocate itself rationally and optimally across business lines, and over time ?

The answer lies in the implementation of a corporate-wide model that will link the search of an optimal line of conduct to reality representative variables.

## 1.1. Modeling intelligent operational risk actions as an optimization problem

### 1.1.1. Linking system centered control variables to external set of probabilities

In modeling optimal behaviour, we have to introduce actions of risk management. At time $t$, let $a_t$ be one possible action in a set of possible choices while facing a risk situation characterized by the probability state variable $x_t$. We define an admissible policy $a_t$ as the sequence of actions adapted to $x_t$

$$a \in \pi = \{a_t\} = \{a_t(x_t)\}$$

Choosing $a_t$ will result in a "reward" (or penalty) $R$, which in a probabilistic context, is the expectation of a system's response $r$ as a function of $x$ and $a$, $r$ being of course given by its probability distribution.

$$R(x,a) = E[r(x,a)]$$

Reward has to do with Value, as it increments the riches held in the system. A risk policy $\pi$ is taken as one time- and state- dependent set of possible actions. Supposing null terminal value at infinite $t$ , the value function is the total of what can be expected in the future, in continuous time setting :

$$V(x,\pi(x_t),t) = \int_{t=0}^{\infty} r(x,\pi(x_t),t)dt$$

In a budgetary context, we will adopt discrete time framework and introduce a discount rate $\gamma$ and taking the expected value, starting from $x_0$ :

$$V^{\pi}(x) = E\left[\sum_{t=0}^{\infty} \gamma^t r(x_t,\pi(x_t)) \middle| x_0 = x\right]$$

A value $V^{\pi}$ for the global period is thus associated to each risk policy $\pi$.

### 1.1.2. Solving either for strategies or for value of risk

The rationality of the risk manager is to seek maximum of value for the net worth of the bank, starting from initial risk state $x_0$ at time $t$, to "learn" policy $\pi^*$ maximising $V$ over the set $A$ of admissible actions :

$$V^{\pi^*}(x_0) = \max_A [V^{\pi}(x_0)]$$

where $V^{\pi}(x)$ is the the consequence of following policy $\pi$ from initial risk situation $x$ going into state $y$

$$V_{t+1}(x) = R(x,\pi_{t+1}) + \gamma.\sum_y P_{x,y}.V_t(y)$$

Value for a given strategy is sum of immediate reward and discounted flow of possible future rewards, depending on the probabilistic response $r$ which makes the transition :

$$V^{\pi}(x) = R(x,\pi(x)) + E\left[\sum_{t=0}^{\infty} \gamma^t r(x_t,\pi(x_t))\right]$$

Solving for optimal policy requires dealing with nonlinear equation

$$\pi^*(x) = \arg\max_{\pi}\left\{R(x,\pi(x)) + E\left[\sum_{t=0}^{\infty} \gamma^t r(x_t,\pi(x_t))\right]\right\}$$

### 1.2. Solving for risk with learning techniques

Solving for strategies is non-linear and is generally difficult. One approach to bypass this is to turn to solving for value. In the general case, taking $G$ to be terminal value, one solves for :

$$V(x,t) = \max_{u \in U}\left[V(f(x), t+1) + G(x,u,t)\right]$$

We will consider below terminal value $G$ to be nil.

For more convenience, we can consider alternately, the distribution $P$ on arriving at discrete states $y$ as possible outcomes of initial state $x$, (instead of steps in time $t$) and introduce the impact of chosen action $a$ :

$$V^* = \max_a \left\{ r(a,x) + \sum_y P(x,y)V(y) \right\}$$

with of course : $\sum_y P(x,y) = 1 \quad \forall x \quad , \quad P(x,y) \geq 0$

We iterate on $V$ since the problem is linear. In practical situation, let $\zeta_t$ be the proxy for $V$ at time $t$ ; The solution for $V$ is obtained under favorable conditions by iterating :

$$\omega_{t+1}(x) \longleftarrow \max \left( r(x,\pi) + \gamma.\sum_y P(x,y)\omega_t(y) \right)$$

- Going into further simplification, we have Q-learning algorithm, which is easy to set up and is model-free (we do not need to know the whole range of $P(x,y)$, only the one transition ahead) however, we have to re-introduce time steps $t$

$$Q_{t+1}(s,a) = r(s,a) + P(s,y).Q_t(y,a)$$

- In Temporal Differences algorithms, the eligibility vector $r_t$ serves as a memory of past states and corrections.

$$d_t = r(t) + \gamma V(t+1) - V(t)$$

$$r_t \longleftarrow r_t + \gamma.d_t z_t$$

We will come back to using this strategy for solving the dynamic program for risk budget after deriving, in Part 2, the stochastic wealth accumulation process in a risk environment.

## 2. Operational risks and risk reduction actions as dynamic processes

### 2.1. Viewing operational loss as a process, not as a static distribution.

A quick analysis of frequency of occurrence versus size of loss reveals there are 2 major categories of risks. On one hand, it is not necessary, in a long term perspective, to consider banks that have frequently large catastrophic losses, as they will inevitably disappear. On the other hand, it is unnecessary to reckon very rare, small, losses. Indeed, we are left with a category of frequent losses of small amounts of money and a another category of rare but catastrophic losses. Most of the 7 categories[1] of risk events in the Basel II Accord contain both of these risks. For

[1] Internal fraud ; external fraud ; employment practices workplace security ; clients products business practices ; damage to physical assets ; business disruption systems failure ; execution delivery process management.

example, petty cash theft and rogue trading that runs up to the billion dollar both belong to Internal Fraud events. However, their processes differ and must be described differently.

For the simplicity of notation, the discussion below relates to a business unit (that can alternately designate one branch, one business line or the whole banking group).

#### 2.1.1. Small, frequent operational risks.

Small losses frequently occur in banking processing, due for instance to inevitable complexity errors, like mistyping a check's amount, causing its cost of rejection or mispayments. They augment with business volume and with increases in business activity, and are "regular" enough for their *return* to vary in a Gaussian way. The *return* of "losses" can be either positive or negative.

For this category of losses, between time $t$ and $t + \Delta t$, the increment of loss is proportional to the current business volume, and to a degree of variability which is a sort of volatility.

$$dX_t^A = X_t^A \sigma_t dB_t$$

where $B_t$ is the Brownian motion.Typically, the bulk of these losses form the "body" of the loss distribution curve. They are often modelized as lognormal or Gamma distributions.

#### 2.1.2. Rare, large losses.

Extreme events like earthquakes or bird flu impacts on business can be modeled as independent events, each having a constant probability of occurring in time, and causing losses according to a time-invariant probability distribution.

It is reasonable to assume for these risks that :

- the probability of catastrophes does not vary over time in current affairs and indeed there is no reason to suppose that earthquakes , or viral attacks have an inherent time-dependent structure : the lapse of time between 2 catastrophes follow the same probability distribution law in the past, in the present time and in the future.
- The losses in dollar value due to these catastrophes follow distributions that are not dependent on each other .

These "stationary and independent increments" hypotheses define classes of processes encompassing a large proportion of losses. On these grounds, insurers have long been using Poisson laws (in which time need not flow uniformly) to represent the number of incoming insurance claims N as a function of time t and compound it with an independent loss size distribution $\{(z_k)\}$.

The total loss up to time t can thus be written :

$$X_t^B = \sum_{j=1}^{N(t)} x_j = \sum_{k=1}^{d} z_k N_t^k$$

where the $(z_k)_{k=1\ldots d}$ are different positive numbers representing $d$ classes of discrete amount of dollar losses (each equal to $z_k$), and the $d$ related Poisson processes $(N^k)_{k=1\ldots d}$ are independent. In this representation, a 100.000$ loss may follow a Poisson law with $\lambda = 1.235$ , wheras a 25.000$ loss law would have a $\lambda = 57.67$ etc. This is essential to the "Loss Distribution Approach" which is becoming increasingly popular in banks seeking to comply with the Basel II Accord by building their own Advanced Measurement Approach (AMA) models.

In continuous form, the jumps in cumulated losses can be written under the form:

$$dX_t^B = \beta X_{t-}dt + X_{t-}\int_{\mathbf{R}*}\left(e^z - 1\right)\widetilde{N}(dt,dz)$$

where $\widetilde{N}(dt,dz) = N(dt,dz) - dt \times \nu(dz)$ is the compensated Poisson random measure [see Lamberton *et al.*,1997]. The deterministic measure $\nu(dz)$ is the Lévy measure, and $dt \times \nu(dz)$ designates the intensity of the jumps between $t$ and $t+dt$ with the technical condition $\int_{\mathbf{R}}\left|e^z - 1\right|\nu(dz) < \infty$ for integrability. These parameters of frequency and of severity of loss can be fiited using collected databases of observed loss history .

In general there is no particular reason why catastrophic losses should follow a trend, so $\beta = 0$ is assumed. The normal inverse Gaussian distribution (NIG) is frequently used because it naturally provides the integral form of Lévy jump processes without the Brownian term [in Elghanjaoui *et al.*, 2002].

#### 2.1.3. Synthesizing a loss function : using Levy processes

These 2 kinds of risks can now be used to describe a general process where coexist both the fat tails (the Pareto distributions in Extreme Value Theory), and the business generated, roughly Lognormal, "expected losses" (described in the Basel terminology). Collecting these components, the total loss is obtained :

$$dX_t = dX_t^A + dX_t^B$$

The losses process, as a combination of pure jumps Poisson process with a Brownian motion with drift, thus follows a Levy process whose characteristic triplet is given by $\left(0,\sigma^2,\nu(dz)\right)$ for example in [Bertoin (1996)]. Its representation can be expressed with the Lévy-Khintchine decomposition [see Framstad *et al.*1999]:

$$dX_t = \sigma_t X_t dB_t + X_{t-}\int_{\mathbf{R}*}\left(e^z - 1)\right)\widetilde{N}(dt,ds)$$

### 2.2. Modeling management action plans as processes

These losses can be reduced by better quality management, and by raising risk awareness and improving risk reporting through voluntary expenses planned in a series of budgets. By taking risk budgets as a control variables, quality of management is naturally introduced into the wealth evolution equation. The stochastic volatility of the Gaussian component of risk, as well as the size of jumps in the Poisson component, can be subordinated to mitigation effects arising from higher quality management. The industry's so-called "key risk indicators", and the scorecards approaches to risk reporting, provide early warning systems that can reduce the impact of losses. Their impacts are modeled below as coefficients modulating the amplitude of the draw-downs.

#### 2.2.1. Qualitative self-assessment : where they fit in

The New Basle Accord has made room for qualitative assessments, in management and in environment. Banks often use ***scorecards*** techniques, which runs through a list of questions often based on values of a series of ***key risk indicators*** or KRI's) and synthesize the result into a discrete score ("red, amber, green" or a grade from 1 to 5). These Risk Control Self Assessments (or RCSA) are a mandatory feature in using a bank's own model.

Intuitively, quality of risk management should introduce a differentiation effect on the required capital: a more proactive mitigation of risk should require less capital than no action is taken to reduce risks.

This effect is here introduced through control variables $\zeta(t)$ and $\xi(t)$ that represent two types of budgets (improvement and insurance) per unit time.

We suppose that senior management has decided to levy these budgets on the income per unit time $\mu$.

### 2.2.2. Impact of budget $\zeta$ as measured by Scorecards on Internal Control $\eta(t)$

Let us suppose that spending an amount of money $\zeta(t)$ to ensure better process management will reduce the standard deviation $\sigma_t$ of small, frequent risks by a factor of $\eta(\zeta)$ as perceived by risk managers.

$$. \quad F_{\varsigma}(\sigma_t) = \sigma_t \eta(\zeta) \quad 0 < k_1 < \eta \leq 1$$

However, the transformation of this expense into a risk mitigant cannot have linear effects indefinitely:

$$\eta''(\zeta) \leq 0 \quad \eta'(\infty) = 0$$

where $\eta(\zeta)$ is the perceived loss reduction factor obtained by spending budget amount $\zeta$ (more fraud detection, personnel, etc.)

Reduced risks on expected losses appear as :

$$X_t^A = X_0 \, exp \sigma_t \eta(\zeta) B_t$$

<u>Key risk indicators on business environment θ</u>

It is commonly observed there are some periods where business activity is more risky than at other times, in some places more than others. Examples are economic downturns, atmospheric or geologic hazard (typhoons zone or period of airline controller strikes…) . Let $\theta$ stand for the business environment indicator, a sort of "temperature" of "hot" or "cool" atmosphere to do business in. $\theta$ should increase, ceteris paribus, if more business is expanding aggressively. $\theta$ is assessed through scorecards as well ( Business environnement factor).

This affects the losses in the following way :

$$F_{\varsigma}(\sigma_t) = \sigma_t \eta(\zeta) \theta(\mu) \text{ with } \theta \geq 1 \quad ; \quad \theta'(\dot{\mu}) \geq 0$$

### 2.2.3. Impact of contingency planning on catastrophic risks

The impact of risk budget cannot always be readily assessed, because some of the threats being hedged against, will occur once in a thousand years. Here, utility function showing aversion to risk can be used to create a component of the value function that will be sensitive to another type of budget control variable.

Reducing impact of catastrophic events through insurance, or recovery plans, at cost $\xi_t$ ,

$$F_{\xi}(X_t^B) = \sum_{0}^{N(t)} K_{\xi}(x_j^B) = g(\xi_t) \int_{\mathbf{R}+} (e^z - 1) \tilde{N}_t(dx)$$

with $1 > k_2 > g(\xi_t) > 0 \quad g''(\xi_t) \leq 0 \quad g' \geq 0$ where individual losses $x_j$ may be capped or reduced , where $k_2$ represents the waiver, a fixed threshold.

## 2.3. Revenue and wealth accumulation

Let $W_t$ be the level of all available assets in the branch or business unit under consideration at time $t$. That "wealth" will be put to work to generate revenues and, on the other hand, will be subject to risks.

The Revenue process is deterministic, with rate μ, and proportional to total cash available for the period:

$$dR_t = \mu W_t dt$$

Cash for the period is made up of losses and gains :

$$dW_t = dR_t + dX_t$$

Cash evolution is described by :

$$dW_t = W_t\left(\mu_t - (\varsigma_t + \xi_t)\right)dt - W_t\sigma_t\eta_\varsigma\theta\, dB_t - g(\xi_t)W_t\int_{\mathbf{R}}\left(e^z - 1\right)\widetilde{N}_t(dx)$$

$dW_t$ can be allowed to be (temporarily) negative, provided $W_t$ stays positive. The cumulative level of cash up to time $t$ is given by :

$$W_t^\pi = \int_0^t dW(s, \mu_t, \sigma_t, \pi_t, N, \nu)$$

where the dependency of cash on policy $\pi(t)$ as the given pair $(\zeta(t)\,,\ \xi(t))$ is highlighted, both through the expense that cuts into business income and the savings from losses. It is of course also dependent on the parameters of the two types of risks.

### 2.4. Enhancing realistic modeling with Utility functions

The implicit rationality in the New Basle Accord is that banks will try to minimize their needs for Capital requirements, as they represent non-interest producing, "non-productive" capital. It is interesting take a broader view here, that Management will not only try to put aside money to pay for the operational losses risks, but will also spend budgets to best protect the process of generating Gross income. The accumulated wealth $W_t$ of a business unit, in a banking group, is to be maximized (under the constraints of non-bankruptcy), with its Lévy-driven evolution equation taking into account contrary effects of stronger income and higher risks and rising risk-prevention spending. In maximizing the objective function, which is a sum across the lines of businesses in a banking group, or across the subsidiaries in a geographically diversified group, natural trade-offs will occur and represent the transverse capital flow due to optimal allocation and risk diversification.

However, since the pure Lévy jumps are unbounded, the effect of a reduction in risk may not perceptible, in mathematical simulation. Its impact is just as if the random variables drawings for catastrophe simulation have yielded smaller results. In once-in-a-century catastrophe categories, having taken preventive/evasive actions should bring its “reward” even if the average short term pay-back for the investment is negligible or vastly delayed. The simple wealth effect $U_1$ resulting from holding of instantaneous wealth $W_t$ is insufficient to describe this “risk anxiety” effect. It is necessary to add an (psychology, risk averse) utility function $U_2$ that will “score” positively whenever risk prevention expense $\xi$ is done, regardless of immediate returns. Determination of $U_2$ can be done through interviews of Top management using classic techniques of mapping risk aversion.

$$V_t = \max_\pi \mathbf{E}\left[\int_0^t e^{-rs}\, U_1\left[W_s^\pi\right]ds + U_2(\xi)\right]$$

The Hamilton Jacobi Bellman equation is then in the following form :

$$0 = \max_{\pi=(\zeta,\xi)}\left[\begin{array}{r} U_1(x) - rt + U_2(\xi_t) + x\left[\mu - (\zeta_t + \xi_t)\right]v(x) - \dfrac{1}{2}x^2\sigma^2\theta^2\eta^2\left(\dfrac{\partial^2 v}{\partial x^2}\right) \\ - \int_{\mathbf{R}^*}\left[v(x(e^z - 1)) - v(x)x(e^z - 1)\right]g(\xi_t)\widetilde{N}_t(ds) \end{array}\right]$$

In other settings, one can also choose other types of objective functions (average past wealth, terminal wealth etc.).

Having at our disposal this value function $V_t$, we will use diverse methods, including the Part 1 techniques of Value iteration strategy in dynamic programming to solve for the risk reduction budgets $(\zeta(t), \xi(t))$.

# 3. Solving the optimization problem

There are many ways to solve this stochastic optimization problem, but not all of them are adequate or useful.

Indeed, depending on the degree of confidence in the model, and on how the model should be used, it may not be necessary or desirable to resolve the complete equation. That is fortunate, as the above HJB equation is generally difficult to solve, due both to its non-linearity and its high degree ( *a priori* integro-differential of the second degree) in the case of Levy processes.

It would it make sense, either, to seek much precision in the management world where so much still depend on human uncertainties and in the face of once-in-a-millenium events. According to the use of the model, many simplifications can be envisaged. In the case of risk management, unlike solving problems in options pricing or credit rating, the goal pursued is an in-depth view of the possible ramifications of actions undertaken in preventive measures, - by rough assessment of their impacts - , or the effect of external shocks on the bank, and how the best course of actions can be designed in response. Observations of the many interacting human, systems and processes show many possible alternative best solutions. Therefore, within the scope of the model, it is reasonable to seek a maximum of insight from a high level viewpoint, then decide what granularity and which control sets are to be primarily considered. Secondary objectives may be then accommodated from a management point of view, that is, solve separately sub-optimal aggregates within the margins left by the global first level solution.

One first way consist in solving the Hamilton Jacobi Bellman equation, through linearization of the HJB equation or the objective function and arrive at a classic, analytic form that allow application of known solutions. One can choose to simplify parts of the model that is not vital to realism, and push further by focusing on interesting variables or factors elsewhere. The goal is to make Markov decision processes appear, and one can then call upon an important arsenal of new developments to get at least one significant result from the equations. In Finance, the additivity of quantities in dollars (one rarely observes units of squared dollars !) will often provide the justification for Galerkin linearisation approximation.

Alternatively, one can choose to heuristically split the Gaussian part from the Lévy part. Considering that the large jumps are in finite number over the optimization horizon, one makes the common sense assumption (to be verified *ex post*) that very rare catastrophes, are a separate contingency problem to solve.

## 3.1. General resolution approaches

### 3.1.1. The Linear Quadratic case

One point of view, in clear inspiration from market theories, is to envisage risk as volatility. Indeed, in the working papers from the early debates of the Basle committee, one can see operational risk sometimes described as the effect by which non-financial assets undergo large variability in their evaluation. According to this view, buildings, corporate organizations or software systems allow the generation of income, but their value can be undermined by hazards such as earthquakes, real estate crises, or human fraud. Conceptually, such physical assets can be represented just like any other financial asset, as generating a revenue but enduring from time to time losses in their values so that the (downside only) variance can stand for risk in the equations. In the objective function, a quadratic penalty representing risk appears, consistently with more-than-linear, observed aversion to risk. Global accumulated wealth follows the same evolution equation as above.

If risk is seen as volatility, characterized by variance, then the objective of growing wealth with minimum risk can be approximated by a linear-quadratic (LQM) markov process. The resulting Riccati equation method works and a closed form solution can be obtained as usual [see Beard *et al,*(1995)].

#### 3.1.2. The Merton consumption-investment case.

The problem here can be analyzed as Merton's consumption-investment problem: losses are due to the risky asset part of the investment, and the steady income is provided by the riskless part, generated by the bank's normal operations. Consumption either can be taken as nil, or it can be viewed as the remainder of income revenue after deduction of the risk business.

Intrinsically, some well-chosen utility functions will give closed form solutions. If a Hyperbolic Absolute Risk Aversion function is used as in [Benth *et al,*(2001)], a closed form solution can be obtained. If it is a Constant Absolute Risk Aversion framework, [Eberlein *et al,*(1999)] has demonstrated that another closed form solution can be derived.

And finally, even when no closed solution is readily obtained, generalized Hamilton Jacobi Bellman and the use of viscosity solution will lead, even if tediously, to an optimum. This approach somewhat justifies the Basel framework of breakdown of banking activities into eight independent lines of businesses : mathematically, the HJB operator is to be seen as a projector in a Lipschitz contraction space, whose basis is the set of Galerkin approximation vectors. This set of non-colinear vectors form can be interpreted as the 8 Basel lines of business.

For a case of application of viscosity solution in optimal risk control, see for example [Mnif *et al.*,2001].

#### 3.1.3. Insurance company's viewpoint: Ruin theory

Another viewpoint comes from the insurance world. Here, the bank is seen as generating an income that can be compared to the premium collecting rate of an insurer, and instead of losing money in paying claims, it incurs losses with every « attack » on the values of its physical or immaterial assets : a worker strike consequential to an inadequate employment practice does indeed diminish revenue directly due to decreasing sales opportunity or due to loss of customer fidelity.

This is referred to as the Cramer-Lundberg model. Optimization aims at maximizing the time of ruin, when "cash" as exposed above becomes null for the first time. Numerical examples are given by Castillo *et al,(*2002). The problem in a more financial form is solved by Hojgaard *et al.,*(1998).

### 3.2. Solving with MDP and Learning processes

#### 3.2.1. Advantages of this approach

The approaches cited above may be inoperative in the real business world, for at least 3 reasons :

- expensive resources are needed both in terms of computers, specialists of HJB and viscosity solution, software etc.
- the results may not converge between the Monte Carlo drawing of conditions and paths, due to the complexity of the parameters
- most of all, business bosses will not buy in, because they will not get the insight of the process.

The insight into risk management must be the drive to set up a framework that will allow through simulation, the exploration of the dynamics of the whole systems, people, processes under internal or external shocks, either under mild uncertainties or under exceptional catastrophes.

This is the viewpoint of automated learning. In this context some aspects of online sampling even appears as a means of generating experience to learn from. Reinforcement learning, or adaptive learning, has more chance to appeal to top management than outright closed form solution derivation. The latter will be generally seen as "math jargon", while the former can be understood by the laymen as the exhaustive testing of generated catastrophe scenarios.

For these reasons, Q-learning or Temporal Difference or other techniques that are both simplified and more intuitive should be favoured over heavy approaches.

### 3.2.2. Learning processes

Application of Adaptive learning theory reformulates the full HJB optimization problem into a simplified Bellman program where probability states replace the temporal dimension. Solving is just looking for the best sequence of actions starting from any state (and, in the case of the Greedy algorithm, looking only for the next step). For an overview, see [Moody (2003)]. Though in theory, establishing transition probability matrices is very difficult due to scarce data, in practice, expert panels and bayesian rules can help build a practical framework for exploration.

Adaptive Learning, which uses quick sampling of state space rewards, is ideal in situations where data is scarce, and there is no model, to learn in real time from responses obtained from exploration of state space. Even when the response from the external world is not immediate, techniques of delayed reinforcement learning allow learning. It requires much less feedback than when supervised learning is concerned.

Using simple greedy control, and reinforcement learning combined with bootstrapping and Monte Carlo generation of input sets, one can derive enough state-response pairs that will prove helpful in bank risk management policies. Q-learning or TD control, which feature model-free based learning, are powerful tools here. Adaptive learning, by updating parameter estimates online, is very fitting in this type of situation..

In neuro-dynamic programming, some of the parameters that are difficult to obtain through statistically fitted methods will become the subject of "learning" procedures. These features of the model are singled out, and neural techniques applied in the inverse formulation of the problem. In risk management, many banks can use numerous "from the top of the head" estimates of frequencies and severity of losses from live business experts. These estimated values can be candidates as learning sets for supervised learning. The models are put to learn to match their output with these values through imprinting and modifying their neurones' weights accordingly by trial-and-error runs. When learning is completed, subcomponents based on neural nets will serve as function approximators for parts of the loss model. [Pham-Hi (1997)] reported how a feedforward 3-layer neural network was used as a stand-in for volatility surface models in the pricing of 3-months LIBOR forward calls on interest rate.

## New computational needs for risk models

The New Basel Accord prescribes extensive back test and stress test of any models specifically derived to calculate these operational risk capital requirements. A question that generally arises is how to exhaustively test the models, and cover the important contingencies. Without a general framework, like Hamilton Jacobi Bellman optimization, it is indeed very difficult to fit into a same "big picture", disparate issues such as shop floor fire prevention and quantification of internal control capacities.

Keeping in mind that one of the main goals of these computing schemes is to help determine best policies, the practitioner in a banking environment will not pursue abstract modeling. His endeavour is to numerically, with the help of computing devices, with the help of mathematical language and tools, explore and unearth facts that are buried in the fuzzy world of rare events probability.

Coping with risk reduction optimized policies is a challenge both for banking institutions and for computational finance specialists. Approaches that draw on learning, adaptation, online sampling in a model-free methodology should bring about solutions favored by business people because it gives them insight. Such optimization processes shed light on the relationship between the variables and parameters, and their relevance to risk policy determination.

It also provides a tentative remedy scarceness of loss data, a big hurdles for banks in the early stages of New Basel Accord application, as well as opening systematic and consistent ways to explore risk scenarios, e.g. by reinforcement learning or, human expertise supervised learning.

**References**

Beard, R.W., G.N. Saridis and J.T. Wen (1995) Approximate Solutions to the Time-Invariant Hamilton-Jacobi-Bellman equation, Journal of Optimization theory and Applications,

Bertoin, J. (1996) *Lévy processes*. Cambridge University Press, Cambridge.

Benth, F.E., K.H. Karlsen and K.Reikvam(2001). Merton's portfolio optimization problem in a Black & Scholes market with non-Gaussian stochastic volatility of Ornstein-Uhlenbeck type. *Pre-print*.

Castillo, M.T, G. Parrocha, (2002) Stochastic Control theory for Optimal Investment, *Pre-print.*

Eberlein, E., S.Raible (1999). Term structure driven by general Levy processes. *Mathematical Finance,* **9,** 31-53

Elghanjaoui S., K.H. Hvistendahl (2002) A Markov chain approximation scheme for singular investment-consumption problem with Lévy driven stock prices. *Preprint.*

Framstad, N.C., B. Oksendahl and A.Sulem (1999), Optimal consumption and portfolio in a jump diffusion market with proportional transaction costs, Rapport de recherche n° 3749, Projet MATHFI, INRIA Rocquencourt, Paris.

Hojgaard, B., M.Taksar (1998). Optimal proportional reinsurance policies for diffusion models, Scandinavian Actuarial Journal, 81,166-180.

Lamberton, D., B. Lapeyre (1997) *Introduction au Calcul Stochastique appliqué à la Finance*, chap.7, Ellipses, Paris

Moody, J. (2003) Risk, Reward & Reinforcement, AMS Workshop- Machine learning , Statistics & discovery, presentation slides, Utah.

Mnif, M., A. Sulem (2001). Optimal risk control under excess of loss reinsurance,Rapport de recherche n° 4317, Projet MATHFI, INRIA Rocquencourt, Paris.

Pham-Hi, D. (1997). Using neural networks to gain insight into stochastic differential equations, (abstracts) *Proceedings of the Society for Computational Economics 1997,* Stanford University.