



DIPARTIMENTO DI INFORMATICA  
E SISTEMISTICA ANTONIO RUBERTI



SAPIENZA  
UNIVERSITÀ DI ROMA

## **Computing optimal recovery policies for financial markets**

Fred E. Benth  
Geir Dahl  
Carlo Mannino

**Technical Report n. 20, 2010**

# Computing optimal recovery policies for financial markets

Fred E. Benth\*    Geir Dahl†    Carlo Mannino‡

December 18, 2010

## Abstract

The current financial crisis motivates the study of correlated defaults in financial systems. In this paper we focus on such a model which is based on Markov random fields. This is a probabilistic model where uncertainty in default probabilities incorporates expert's opinions on the default risk (based on various credit ratings). We consider a bilevel optimization model for finding an optimal recovery policy: which companies should be supported given a fixed budget. This is closely linked to the problem of finding a maximum likelihood estimator of the defaulting set of agents, and we show how to compute this solution efficiently using combinatorial methods. We also prove properties of such optimal solutions. A practical procedure for estimation of model parameters is also given. Computational examples are presented and experiments indicate that our methods can find optimal recovery policies for up to about 100 companies. The overall approach is evaluated on a real-world problem concerning the major banks in Scandinavia and public loans. To our knowledge this is a first attempt to apply combinatorial optimization techniques to this important, and expanding, area of default risk analysis.

*Keywords: Financial models, discrete optimization, bilevel programming, Markov random field.*

---

\*Department of Mathematics, CMA, University of Oslo, Norway. [fredb@math.uio.no](mailto:fredb@math.uio.no)

†Department of Mathematics and Department of Informatics, CMA, University of Oslo, Norway. [geird@math.uio.no](mailto:geird@math.uio.no)

‡University of Rome, La Sapienza, Department of Computer Science, Rome, Italy, and visiting scientist at CMA. [mannino@dis.uniroma1.it](mailto:mannino@dis.uniroma1.it)

# 1 Introduction

In financial transactions there is always a risk that the counter-party can not fulfill the obligations, and that the investment goes to default. In view of the current crisis, the world financial markets have experienced companies with high credit rating to run into distress, and even defaulting. Examples include apparently well-managed companies like the investment bank Lehman Brothers that defaulted due to large exposure in subprime loans, and the insurance giant AIG that was rescued by the Government after huge losses in the credit derivatives sector. Credit risk modeling is of vital importance for investors in order to understand and manage their counter-party risk in financial operations. In particular, understanding how the default risk of one company is linked to others is crucial.

In this paper we study a Markov random field model for the default risk of each company, modelled as a Bernoulli random variable, which may be dependent on the default risk of other companies in the market. The whole market is represented as vertices in a graph, with the directed arcs (edges) between the vertices describing dependency in terms of default risk. Moreover, we include uncertainty in the default probabilities, which may be interpreted as a collection of expert's opinions on the default risk, coming from various credit rating companies. The role of these companies in the marketplace were claimed to be one of the sources for creating the financial crisis starting in 2007/2008. The special case of undirected graph model was recently analyzed in [10].

Our modelling approach is within the class of what is frequently referred to as *Bernoulli mixture models* (see [12]). We do not suppose that the default probabilities and correlations come from any structural model (like for instance the Merton model, see [17]), but consider a reduced-form model instead. In a Bayesian context, we are interested in the *posterior* default probability distribution of the market, given the *prior* default probability distribution and conditioning on expert judgements introduced as probability distributions on the default probabilities.

The main question we focus on in this paper is to use the graph approach to derive optimal recovery policies. A characteristic of the current financial crisis is that governments world-wide try to resolve the turmoils by pumping in money to recover various sectors which are believed to be crucial for the market as a whole. We use our model to look for optimal policies in the sense of recovering as much of the potential loss as possible, given a fixed budget for the government. Closely related to the recovery problem is the "default geometry" of the market in the sense of the most probable set of

defaulted companies under the model assumed. This seems to be a question which is rarely investigated in the literature, but turns out to be important in order to understand the derivation of optimal recovery policies. Moreover, it is interesting in its own since it will provide investors guidelines for which companies will most likely default, and which are more “safe havens”. Of course, from the default probabilities of each firm this is simple to answer on an individual basis, but after introducing dependencies among companies the picture may be complicated significantly.

The two problems posed above result in discrete optimization problems that in general are highly non-trivial to solve. However, as we show in the paper as one of our main results, finding the most likely set of default companies can be transformed into a well-known combinatorial optimization problem called the *minimum cut problem*. There exist very efficient combinatorial algorithms for solving such problems, so that large-scale instances can be solved extremely fast; see [1] for a discussion of such algorithms, their complexity and implementations. It is therefore computationally feasible to study markets consisting of many companies with complex dependency structures. Furthermore, we are also able to prove a uniqueness-result for the most likely default set in the Markov random field model. Among all the most probable configurations of companies defaulting, there is a unique one of maximal cardinality. This maximal set contains all the other most likely defaulting sets.

In the recent literature on credit risk modelling, the focus has been on dynamic portfolio credit risk models. Dynamical models based on point processes have been used in order to model the exposure to credit risk in a portfolio, with the aim of computing the loss distribution. One supposes a conditionally independent default structure among the firms in the portfolio. Dependencies in the credit situation of the different firms are modelled by the default intensities being correlated stochastic processes (see Jarrow and Turnbull [14] and Duffie and Singleton [9]). Since typically it is hard to compute the loss distribution, approximations through limit theory or simulation are called for. For example, Deng, Giesecke and Lai [8] propose an importance sampling technique to estimate rare events probabilities in a credit risk portfolio. The method is based on a change of measure and resampling to approximate the zero-variance importance measure connected to the rare events.

The alternative to *intensity-based models* is *structural modelling* where credit events of a particular firm is modelled via the evolution of the firm’s value and its capital structure. Typically one is interested in the firm’s default, triggered by the value of the firm hitting a certain threshold. Recently,

Carmona, Fouque and Vestal [5] presented an efficient numerical approach to compute rare credit portfolio losses in a structural model based on interacting particle systems.

Our model provides an explicit way to model statistically the impact of a default of one company to another. In this respect, our approach places itself in the stream of literature on *credit contagion*. For example, Giesecke and Weber [13] study a dynamic reduced-form credit risk model based on the so-called voter model in the theory of interacting particle systems. Similar to our approach, they consider a lattice of companies which may be in two liquidity states, “high” and “low”. To model contagion effects, a company  $i$  in the lattice have a neighborhood of business partners  $j \in N(i)$  (being other firms in the lattice, of course), where  $N(i)$  is the set of labels of the partners. The transition between states of company  $i$  is a Poisson event with transition rate depending on the states of the business partners  $j \in N(i)$ . Although the contagion effects in the lattice is dynamic and modelled differently than ours, the idea is somewhat similar to the approach proposed and analysed in this paper. Giesecke and Weber [13] provide an explicit Gaussian approximation the portfolio loss distribution, and analyse the effects of credit contagion based on this.

In our approach, we allow explicitly for impacts between business partners which are not necessarily mutual. We may have that a firm  $i$  is depending on the credit situation of firm  $j$ , but not the other way around. Also, we mix in the possibility of firm’s exposure to macro-economic factors, as we assume these to be vertices in the graph. In this way, our model takes into account both *cyclical default dependence* and *credit contagion* (see Giesecke and Weber [13] for more discussions on this). Different than most of the literature on credit risk, we focus on providing an efficient computational method for the model applied to studying recovery policies and most likely market scenarios, given impacts between business partners and from macro-economic variables.

Our modeling approach and analysis are illustrated with numerous examples. In particular, we show how our graph representation of companies can be used to include variables modelling the “state of the economy”, that is, macro-economic factors. For example, we consider a graph consisting of a bank with several clients borrowing money. The clients are divided into two groups, one being “good” clients with low default probability and the other “bad” clients with high default probability. But, in addition, we suppose that the borrowers are dependent on a vertex in the graph representing the “state of the economy”, being for instance the interest rate in the market, or inflation. The state is either good or bad, and influencing the clients di-

rectly. In this way we can investigate the effect of, for instance, a change in the interest rate from “normal” to “high”, and study the impact on the bank. The example may be viewed as a simplified model for banks with subprime loans in their portfolio.

Another important issue that we study through examples is how the loss distribution, defined as the number of defaulted companies, is dependent on the introduction of additional default uncertainty through expert judgments.

Finally, the overall approach is evaluated on a real-world problem concerning the major banks in Scandinavia and public loans. We discuss how the data are obtained from ratings and annual reports and demonstrate how the parameters of the model are estimated. Then we compute most likely default sets and optimal recovery policies for the actual situation and for a more pessimistic scenario.

The paper is organized as follows: In Section 2 our model is introduced and put into a proper statistical context. The analysis of the model and some main results are presented in Section 3 where we discuss most likely default sets (of companies) and the connection to the minimum cut problem. The model includes certain parameters representing default probabilities and dependencies. The role of these parameters is discussed in Section 4 and a procedure for estimating the parameters is presented in Section 5. Our main issue concerning optimal recovery policies, is treated in Section 6, while Section 7 presents computational experiments and some examples of our methodology and on a real-world case. Finally, some concluding remarks and future directions are given.

## 2 The model

The purpose of this section is to introduce our mathematical model and to define associated optimization problems. Our model belongs to the class of Markov random fields (MRF), which has proved to be useful in several areas, for instance, in image segmentation ([2], [3], [6], [11], [20]), and statistical physics (the Ising model) [16].

The model is stochastic and represents a financial system consisting of different economic agents (companies, banks) and economic dependencies between pairs of such agents. The main purpose of the model is to be able to determine a most likely set of agents that will be defaulted at some specified future time. The financial system is represented by a directed graph  $G = (V, E)$  with vertex set  $V$  and arc set  $E$ . The vertices represent

the agents, and the arcs represent a dependency between pairs of agents. This dependency will be explained below. Define  $n := |V|$  and  $m := |E|$ . An arc from vertex  $u$  to vertex  $v$  is denoted by  $(u, v)$  or  $uv$ . We consider a one-step dynamical model, today is time  $t_0$  and  $t_1$  represents a specified future time (e.g. one month ahead). Associated with each  $v \in V$  is a random Bernoulli variable  $x_v \in \{0, 1\}$  where  $x_v = 1$  means that agent  $v$  defaults at time  $t_1$ , while  $x_v = 0$  means that it survives. The vector  $x = (x_v : v \in V)$  is a random variable, and we refer to  $x$  as the *default state vector*.

At time  $t_0$  we have available certain information about the system which specifies the probability distribution of the random variable  $x$ . This information includes a real number  $z_v \in [0, 1]$ , for each vertex  $v$ , which may be considered as an approximation to  $x_v$ . Thus,  $z_v$  may be the result of an analysis – made at time  $t_0$  – of the financial situation of the agent  $v$ . For instance, one may think of  $z_v$  as the proportion of a group of analysts that think agent  $v$  will be default at time  $t_1$ . We will treat  $z = (z_v : v \in V)$  as a random vector and call it the *judgement vector*. The connection between  $x$  and  $z$  is given by a statistical model where the conditional probability distribution  $f(z|x)$  of  $z$  given  $x$  is specified. We assume conditional independence between agents in the sense that

$$f(z|x) = \prod_{v \in V} f_v(z_v|x_v) \quad (1)$$

for some conditional probability functions  $f_v$  (which we assume to be strictly positive). This means that the judgement value  $z_v$  depends on the future status  $x_v$ ; this reflects the knowledge of the experts. However, for a given default state vector  $x$ , the judgements  $z_v$  for the agents ( $v \in V$ ) are independent.

The model also includes dependencies among the agents as represented by the arcs in the directed graph  $G$ . This is done using a so-called Ising model which gives a *prior* probability distribution for the default state vector  $x$

$$\pi(x) = \frac{1}{s} \exp\left\{ \sum_{v \in V} [\alpha_v^1 x_v + \alpha_v^0 (1 - x_v)] + \sum_{(u,v) \in E} \beta_{uv} I(x_u \leq x_v) \right\}. \quad (2)$$

Here  $\beta_{uv} \geq 0$  is a parameter indicating how dependent agent  $v$  is on agent  $u$  (for  $(u, v) \in E$ ); the larger  $\beta_{uv}$  is, the less likely are all events where  $u$  defaults but  $v$  does not. The constant  $s$  is a normalization constant making  $\pi(x)$  a proper probability distribution (i.e. total probability is 1) and the indicator function  $I(x_u \leq x_v)$  equals 1 if  $x_u \leq x_v$  and 0 otherwise. The parameters  $\alpha_v^1$  and  $\alpha_v^0$  reflect a priori information on the financial position

of agent  $v$ , independent of the expert judgement  $z_v$ . Actually, we interpret  $(1/s)e^{\alpha_v^1}$  (resp.  $(1/s)e^{\alpha_v^0}$ ) as an “a priori probability” that agent  $v$  defaults (resp. survives) whenever  $\beta_{uv} = 0$ . We may here assume that  $\alpha_v^1, \alpha_v^0 \leq 0$  since subtracting a (suitable large) number, say  $b$ , from each  $\alpha_v^1$  and  $\alpha_v^0$  in (2) results in the exponential term being multiplied by the constant  $e^{-nb}$ , and this number will be incorporated in the normalization constant  $s$ . As we note later, the value of  $s$  is irrelevant for our method below.

We should remark that the Ising model (e.g. in image analysis) is most frequently used in connection with undirected graphs. Then the “correlation” term in (2) would be  $\beta_{uv}I(x_u = x_v)$  summed over all edges in the graph. In our model, however, a directed graph is crucial for being able to model asymmetric markets which should be more the rule than an exception.

The parameters in our model may therefore be summarized as follows:

- (i)  $\alpha_v^1, \alpha_v^0 \leq 0$ : prior parameters for agents
- (ii)  $\beta_{uv} \geq 0$ : prior parameters reflecting dependencies between agents
- (iii)  $z_v \in [0, 1]$ : expert judgement of the robustness of agent  $v$
- (iv)  $f_v(\cdot|\cdot)$ : conditional probability functions used to describing  $x_v$  and  $z_v$ .

The *posterior* distribution  $\pi(x|z)$  for  $x$  given  $z$  may be determined through Bayes formula, and this gives

$$\pi(x|z) = \phi(z)\pi(x)f(z|x)$$

where  $\phi$  is a suitable normalization function. We now consider the judgement vector  $z$  as fixed, and therefore  $\phi(z)$  is a constant. The Bayesian paradigm ([2]) is to base all inference on the posterior distribution  $\pi(x|z)$ . A natural estimate of  $x$  is the *maximum a posteriori solution*  $\hat{x}$ . This solution  $\hat{x}$  maximizes  $\pi(x|z)$ , or equivalently,  $\pi(x)f(z|x)$  (since  $\phi(z)$  is a constant). Thus, the maximum a posteriori solution is an optimal solution of the discrete optimization problem

$$(MLDS) \quad \max\{F(x) : x \in \{0, 1\}^n\} \quad (3)$$

where

$$F(x) = \exp\left\{\sum_{v \in V} [\alpha_v^1 x_v + \alpha_v^0 (1 - x_v)] + \sum_{(u,v) \in E} \beta_{uv} I(x_u \leq x_v)\right\} \cdot \prod_{v \in V} f_v(z_v | x_v)$$

An optimal solution  $\hat{x}$  in MLDS is the incidence vector of a subset  $S \subseteq V$  (i.e., a  $(0, 1)$ -vector with ones in the positions of  $S$ ) and we call  $S$  a *most likely default set*, defending the attribution (MLDS). Thus, the MLDS problem is to find a most likely default set. Note that – in MLDS – the judgement vector  $z$  is given. An instance of MLDS will be denoted by the tuple  $(G, f, z, \alpha, \beta)$  where all these components are described above.

Note that one may also interpret the MLDS problem more directly as finding a maximum likelihood estimator  $\hat{x}$  of the (unknown) default state vector  $x$  for the given probability distribution  $F$  (as mentioned, multiplying  $F$  by a suitable constant gives a true probability distribution).

We remark that our prior distribution  $\pi$  coincides with the graphical model investigated in [10]. A main result in [10] is that there is a one-to-one correspondence between the model parameters and certain probabilities: the marginal probabilities of default of agents and the probabilities of pairs of agents defaulting simultaneously. Hence, the model can be fully specified by these probabilities. The computation of the parameters, given these marginal probabilities, is however a non-trivial task. Our main contributions are:

1. We extend the prior model so as to incorporate additional expert judgement (via the posterior). Moreover, we consider a directed model. We also propose a sound statistical procedure for estimating model parameters.
2. We show how to *compute efficiently* a maximum likelihood estimator in the model, both in the prior  $\pi$  and in the posterior  $F$  (in which case we find an MLDS solution). This is done by a transformation into the minimum cut problem; a basic combinatorial optimization problem for which highly efficient algorithms exist.
3. We use the basic model as the core in another optimization model whose purpose is to find optimal recovery policies in financial systems.

### 3 Most likely default sets and minimum cuts

We now demonstrate that the MLDS problem may be reformulated as a minimum cut problem in a (directed) graph derived from  $G$ . This means that the problem may be solved efficiently by known algorithms, see e.g. [1].

Consider again the function  $F$  in (3) which is to be maximized. Equiva-

lently, we may maximize its logarithm which is

$$L(x) = \sum_{v \in V} [\alpha_v^1 x_v + \alpha_v^0 (1 - x_v)] + \sum_{(u,v) \in E} \beta_{uv} I(x_u \leq x_v) + \sum_{v \in V} \log f_v(z_v | x_v) \quad (4)$$

Sometimes this function (or its negative) is referred to as the energy function of the MRF.

We now introduce an auxiliary graph  $G' = (V', E')$  based on  $G$ . Let the vertex set be  $V' = V \cup \{s, t\}$  where the vertices  $s$  and  $t$  are called terminal vertices. The arc set  $E'$  consists of all arcs in  $E$  and, in addition, arcs of the form  $(s, v)$  and  $(v, t)$  for each  $v \in V$ . Thus,  $G'$  is obtained from  $G$  by adding two terminal vertices  $s$  and  $t$  as well as arcs connecting each original vertex to the terminal vertices. We associate weights to the arcs in  $G'$  as the following table indicates:

arc $e$	weight $w_e$
$(u, v) \in E$	$w_{uv} := -\beta_{uv}$
$(s, v) (v \in V)$	$w_{sv} := \alpha_v^0 + \log f_v(z_v   x_v = 0)$
$(v, t) (v \in V)$	$w_{vt} := \alpha_v^1 + \log f_v(z_v   x_v = 1)$

It is important to note that all these weights are nonpositive:  $f(z_v | x_v = 1), f(z_v | x_v = 0) \leq 1$  as these are discrete probabilities,  $\beta_{uv} \geq 0$  and  $\alpha_v^1, \alpha_v^0 \leq 0$ . This fact is crucial for the efficient solvability of the MLDS problem.

Consider a partition of the vertex set  $V'$  into two sets  $S$  and  $T = V' \setminus S$  where  $s \in S$  and  $t \in T$ . The set of arcs  $e = (u, v)$  in  $G'$  with  $u \in S$  and  $v \in T$  is denoted by  $\delta^+(S)$ , and it is called an *st-cut* (or simply a *cut*). The *weight* of the cut  $\delta^+(S)$ , denoted by  $w(\delta^+(S))$ , is the sum of the weights of the arcs in the cut, i.e.,

$$w(\delta^+(S)) = \sum_{e \in \delta^+(S)} w_e.$$

Let  $x \in \{0, 1\}^n$  and define  $S_x = \{s\} \cup \{v \in V : x_v = 1\}$  and  $T_x = \{t\} \cup \{v \in V : x_v = 0\}$ . Then we have

$$\begin{aligned} w(\delta^+(S_x)) &= \sum_{v \in S_x \setminus \{s\}} w_{vt} + \sum_{v \in T_x \setminus \{t\}} w_{sv} + \sum_{(u,v) \in E \cap \delta^+(S_x)} w_{uv} \\ &= \sum_{v \in S_x \setminus \{s\}} (\alpha_v^1 + \log f(z_v | x_v = 1)) + \sum_{v \in T_x \setminus \{t\}} (\alpha_v^0 + \log f(z_v | x_v = 0)) \\ &\quad + \sum_{(u,v) \in E \cap \delta^+(S_x)} (-\beta_{uv}) \\ &= \sum_{v \in S_x \setminus \{s\}} (\alpha_v^1 + \log f(z_v | x_v = 1)) + \sum_{v \in T_x \setminus \{t\}} (\alpha_v^0 + \log f(z_v | x_v = 0)) \\ &\quad + \sum_{(u,v) \in E: x_u \leq x_v} \beta_{uv} - \sum_{(u,v) \in E} \beta_{uv} \\ &= L(x) - \beta^* \end{aligned}$$

where  $\beta^* = \sum_{(u,v) \in E} \beta_{uv}$  is a constant.

Thus, up to a constant, the weight of the cut  $\delta^+(S_x)$  equals the value of the function  $L$  at  $x$ . The calculation also shows that, conversely, the weight of cut  $\delta^+(S)$  equals  $L(x) - \beta^*$  where  $x \in \mathbb{R}^n$  is the incidence vector of  $S$  in  $V$ , i.e.,  $x_v = 1$  if  $v \in S \cap V$  and  $x_v = 0$  otherwise.

This discussion shows that the MLDS problem is equivalent to finding a maximum weight cut in the graph  $G'$  with weights as defined above. As noted above all these weights are nonpositive. Therefore the maximum weight cut problem is equivalent to finding a minimum weight cut in  $G'$  where the weights are nonnegative and given by  $w'_e := -w_e$  ( $e \in E'$ ).

We have therefore shown the following theorem (see [18] for algorithmic complexity).

**Theorem 1** *The MLDS problem is solvable in polynomial time as a minimum  $st$ -cut problem in the directed graph  $G'$  with nonnegative weights  $w'_e$  ( $e \in E'$ ).*

The important consequence of this theorem is that we can solve MLDS efficiently for any given graph  $G$ . We refer to Section 7 for some computational experiments where MLDS is solved for certain specific instances, of varying structure and size.

If  $\delta^+(S)$  is an optimal solution to the above minimum  $st$ -cut problem, then  $S \setminus \{s\}$  is an MLDS, a most likely default set. A relevant question concerns the *uniqueness* of the most likely default sets. In general, we may have several optimal solutions to the minimum  $st$ -cut problem, each corresponding to different most likely default sets. However, the following property holds:

**Theorem 2** *Let  $S^*$  be a maximum cardinality most likely default set. Then  $S^*$  is unique, and it contains every other most likely default set.*

**Proof.** In what follows, for each  $S \subseteq V$  we denote by  $f(S)$  the weight of the  $st$ -cut  $\delta^+(S \cup \{s\})$  (using the nonnegative weights  $w'_e$ ,  $e \in E'$ ). It is well known (see [18]) that  $f$  is a submodular set function, that is, for every  $S_1, S_2 \subseteq V$  we have

$$f(S_1 \cup S_2) + f(S_1 \cap S_2) \leq f(S_1) + f(S_2).$$

Observe that if  $\delta^+(S_1 \cup \{s\})$  and  $\delta^+(S_2 \cup \{s\})$  are minimum  $st$ -cuts, with  $f(S_1) = f(S_2) = m$ , then we have  $f(S_1 \cup S_2) \geq m$  and  $f(S_1 \cap S_2) \geq m$ .

Thus,  $2m \leq f(S_1 \cup S_2) + f(S_1 \cap S_2) \leq f(S_1) + f(S_2) = 2m$  which implies that  $f(S_1 \cup S_2) = m$  (and  $f(S_1 \cap S_2) = m$ ).

Let now  $S^*$  be a maximum cardinality most likely default set, and let  $\bar{S}$  be a most likely default set such that  $\bar{S} \not\subseteq S^*$ . By optimality and submodularity,  $f(S^*) = f(\bar{S}) = f(S^* \cup \bar{S})$ . This implies that  $S^* \cup \bar{S}$  is a most likely default set. Since  $\bar{S} \not\subseteq S^*$ , we have  $|S^* \cup \bar{S}| > |S^*|$ , a contradiction. So, every most likely default set is contained in  $S^*$  and the theorem follows.  $\square$

It is not difficult to modify the original arc weights so that the minimum  $st$ -cut is unique and therefore it corresponds to the maximum cardinality most likely default set  $S^*$  discussed in Theorem 2. This can be done by increasing, for each  $v \in V$ , the weight  $w'_{sv}$  of arc  $(s, v)$  by a suitable small quantity  $\epsilon > 0$ , and we have the following result.

**Proposition 3** *Given an instance of the MLDS problem, let  $G'$  be the corresponding auxiliary graph, with weights  $w'$ . Then we can find weights  $w''$  so that the minimum  $st$ -cut in  $(G', w'')$  is unique (and corresponds to the maximum cardinality most likely default set).*

**Proof.** Consider the auxiliary graph  $G'$  with weights  $w'$ , let  $m$  be the minimum cut weight, and let  $\tilde{m}$  be the weight of a second best minimum  $st$ -cut, meaning that  $\tilde{m} > m$ , and no other cut has weight in between (take  $\tilde{m} = +\infty$  if all cuts have equal value). It suffices to take  $0 < \epsilon < (\tilde{m} - m)/|V|$ , and define  $w''_{sv} = w'_{sv} + \epsilon$ , for  $v \in V$ ,  $w''_{uv} = w'_{uv}$  otherwise. Then, for any  $st$ -cut  $\delta^+(S)$  we have  $w''(\delta^+(S)) = w'(\delta^+(S)) + |V \setminus S| \cdot \epsilon = w'(\delta^+(S)) + |V| \cdot \epsilon - |S| \cdot \epsilon = K + w'(\delta^+(S)) - |S| \cdot \epsilon$ , where  $K$  is constant.

First observe that every non-minimum  $st$ -cut  $D = \delta^+(\bar{S})$  in  $(G', w')$  (i.e.  $w'(D) \geq \tilde{m} > m$ ) is also a non-minimum  $st$ -cut in  $(G', w'')$ . In fact, let  $C = \delta^+(\hat{S})$  be a minimum  $st$ -cut in  $(G', w')$ . Then  $w'(D) - w'(C) \geq \tilde{m} - m$ . It follows that  $w''(D) - w''(C) = K + w'(D) - \epsilon \cdot |\bar{S}| - (K + w'(C) - \epsilon \cdot |\hat{S}|) \geq \tilde{m} - m - \epsilon \cdot (|\bar{S}| - |\hat{S}|) \geq \tilde{m} - m - \epsilon \cdot |V| > 0$ , implying that  $D$  is not a minimum  $st$ -cut of  $(G', w'')$ .

So every minimum  $st$ -cut in  $(G', w'')$  is also minimum in  $(G', w')$ .

Now, let  $B^* = \delta^+(S^*)$  and  $C = \delta^+(\hat{S})$  be two distinct minimum  $st$ -cut in  $(G', w')$ , and suppose  $|S^*|$  is the largest possible. By Theorem 2 we have  $|S^*| > |\hat{S}|$ . It follows that  $w''(B^*) - w''(C) = K + w'(B^*) - \epsilon \cdot |S^*| - (K + w'(C) - \epsilon \cdot |\hat{S}|) = \epsilon \cdot (|\hat{S}| - |S^*|) < 0$ , and  $B^*$  is the unique minimum  $st$ -cut in  $(G', w'')$ .  $\square$

In practice, weights are approximated and represented by rational numbers; so they can always be scaled up to positive integers and  $\tilde{m} - m \geq 1$ .

Then it suffices to take  $\epsilon < 1/|V|$ .

For the sake of simplicity, in the rest of the paper we will assume that the minimum  $st$ -cut in the auxiliary graph  $G'$  with weights  $w'$  is unique.

## 4 The role of the parameters

In order to better understand the role of the parameters  $\alpha_v^1$ ,  $\alpha_v^0$  and  $\beta_{uv}$  in our model, we study analytically the special case where the graph  $G$  has two vertices. Then it is possible to calculate explicit expressions for the prior probabilities  $\pi$ . (The calculation of  $F$  also involves the functions  $f_v$ , so it is more complicated; for this reason we focus on the prior.)

So, assume that the graph  $G$  consists of two vertices  $v_1$  and  $v_2$  and arcs  $(v_1, v_2)$  and  $(v_2, v_1)$ . Define  $p_{v_i}^0 = e^{\alpha_{v_i}^0}$  and  $p_{v_i}^1 = e^{\alpha_{v_i}^1}$  ( $i = 1, 2$ ). Then the prior (2) is given by

$$\begin{aligned}\pi(0, 0) &= (1/s)p_{v_1}^0 p_{v_2}^0 e^{\beta_{12} + \beta_{21}} & \pi(0, 1) &= (1/s)p_{v_1}^0 p_{v_2}^1 e^{\beta_{12}} \\ \pi(1, 0) &= (1/s)p_{v_1}^1 p_{v_2}^0 e^{\beta_{21}} & \pi(1, 1) &= (1/s)p_{v_1}^1 p_{v_2}^1 e^{\beta_{12} + \beta_{21}}.\end{aligned}$$

where

$$s = (p_{v_1}^0 p_{v_2}^0 + p_{v_1}^1 p_{v_2}^1) e^{\beta_{12} + \beta_{21}} + p_{v_1}^0 p_{v_2}^1 e^{\beta_{12}} + p_{v_1}^1 p_{v_2}^0 e^{\beta_{21}}$$

Then the marginal probability  $p(\beta_{12}, \beta_{21})$  that  $x_{v_2} = 1$  ( $v_2$  defaults) becomes

$$p(\beta_{12}, \beta_{21}) = \frac{p_{v_1}^0 p_{v_2}^1 e^{-\beta_{21}} + p_{v_1}^1 p_{v_2}^1}{(p_{v_1}^0 p_{v_2}^0 + p_{v_1}^1 p_{v_2}^1) + p_{v_1}^0 p_{v_2}^1 e^{-\beta_{21}} + p_{v_1}^1 p_{v_2}^0 e^{-\beta_{12}}}$$

which is an increasing function of  $\beta_{12}$  (for each  $\beta_{21}$ ) as one would expect. Now, fix  $\beta_{21}$  and take the limit as  $\beta_{12} \rightarrow \infty$ , so there is large influence of  $v_1$  on  $v_2$ :

$$\lim_{\beta_{12} \rightarrow \infty} p(\beta_{12}, \beta_{21}) = \frac{p_{v_1}^0 p_{v_2}^1 e^{-\beta_{21}} + p_{v_1}^1 p_{v_2}^1}{(p_{v_1}^0 p_{v_2}^0 + p_{v_1}^1 p_{v_2}^1) + p_{v_1}^0 p_{v_2}^1 e^{-\beta_{21}}}$$

Further, if we  $\beta_{21}$  tends towards 0, this gives

$$\lim_{\beta_{21} \rightarrow 0} \lim_{\beta_{12} \rightarrow \infty} p(\beta_{12}, \beta_{21}) = \frac{p_{v_1}^0 p_{v_2}^1 + p_{v_1}^1 p_{v_2}^1}{(p_{v_1}^0 p_{v_2}^0 + p_{v_1}^1 p_{v_2}^1) + p_{v_1}^0 p_{v_2}^1}$$

A more interesting case is when  $\beta_{21}$  tends towards infinity:

$$\lim_{\beta_{21} \rightarrow \infty} \lim_{\beta_{12} \rightarrow \infty} p(\beta_{12}, \beta_{21}) = \frac{p_{v_1}^1 p_{v_2}^1}{p_{v_1}^0 p_{v_2}^0 + p_{v_1}^1 p_{v_2}^1}$$

So for the highly correlated case, where  $\beta_{12}$  and  $\beta_{21}$  are large, this limit probability of  $v_2$  defaulting is equal to the “a priori” conditional probability of both defaulting given that  $x_{v_1} = x_{v_2}$  (recall the interpretation of  $(1/s)p_{v_1}^1$  as the a priori probability of  $v_1$  defaulting, see Section 2).

It is also possible to compute the prior analytically when  $G$  is a star, and thereby analyze different questions, but we do not go into this here. Finally, we mention an observation concerning the parameters in the prior: (i) if  $\alpha_v^1 \geq \alpha_v^0 + \beta|\delta^+(v)|$  where  $|\delta^+(v)|$  is the number of arcs leaving  $v$ , then there is a maximizer  $x$  of the prior  $\pi$  with  $x_v = 1$ . (This follows by a simple variation of the single variable  $x_v$  in (2)). So, assuming  $F \approx \pi$ , a similar property is likely for a MLDS solution. Also, we see that if  $\alpha_v^0 \geq \alpha_v^1 + \beta|\delta^-(v)|$ , then there is a maximizer  $x$  of  $\pi$  with  $x_v = 0$ .

## 5 Parameter estimation

For the practical usefulness of our model it is required to determine suitable values for the parameters ( $\alpha$ 's and  $\beta$ 's). In particular, the agent dependency parameters  $\beta_{uv}$  may be hard to set “manually”. We here suggest a methodology for parameter estimation which may be used in concrete situations to calibrate the model.

First we discuss the situation when the number  $n$  of agents is “moderate”, so that all  $(0,1)$ -vectors of length  $n$  may be enumerated. If some prior knowledge of default probabilities of agents is known, this may be reflected in a choice of  $\alpha_v^0$  and  $\alpha_v^1$  (see Section 2); otherwise they may all be set to e.g. 0. Next, the expert judgement  $z_v$  ( $v \in V$ ) is set based on the credit risk rating of each bank  $v$  by assigning a default probability to each of the different ratings used (typically very few classes). See Section 7 for details on this approach. Values  $z_v$  for other agents than banks must be set by some other/manual procedure for risk evaluation.

It remains to estimate the dependency parameters  $\beta_{uv}$ 's. For this purpose one may specify certain conditional probabilities, based on our model. From an economic analysis one may find how “exposed” an agent  $v$  is to another agent  $u$ . More specifically, say that  $v$  is a bank that has public loans of size  $p$  million EURO, and let agent  $u$  correspond to “public loans”. Also assume bank  $v$  has total assets of value  $q$  million EURO. Then we consider the fraction  $p_{uv}^* := p/q$  as an estimate of the conditional probability that bank  $v$  defaults given default of the public loans  $u$ . A similar idea may be used when  $u$  is another bank with loans in bank  $v$ . Thus, for each pair  $(u, v)$  of agents for which  $u$  may have such a negative influence on  $v$ , so  $\beta_{uv}$  could

be positive, we compute a number  $p_{uv}^*$  using the same economic reasoning. This leads to the problem of choosing  $\beta_{uv} \geq 0$  for each  $(u, v) \in E$  such that

$$g(\beta) := \sum_{(u,v) \in E} (\mathbb{P}^\beta(x_v = 1|x_u = 1) - p_{uv}^*)^2$$

is minimized. Here  $\beta$  denotes a vector containing all the  $\beta_{uv}$ 's, and  $\mathbb{P}^\beta(\cdot|\cdot)$  denotes the conditional probability in our model, indicating the dependence on  $\beta$  (the dependence on  $\alpha$ 's and  $z$  is not indicated).

In summary, economic analysis gives certain conditional probabilities  $p_{uv}^*$  ( $(u, v) \in E$ ) that one believes are correct, and the  $\beta_{uv}$ 's are adjusted to fit these  $p_{uv}^*$ 's best possible, using a least squares approach. The problem of minimizing the function  $g$ , subject to the constraint that  $\beta$  is nonnegative, is a nonlinear constrained optimization problem. The computation of the function value  $g(\beta)$  may be done by enumeration of all possible state vectors  $x \in \{0, 1\}^n$  (recall,  $n$  is assumed moderate) and by suitable summations one computes the conditional probabilities  $\mathbb{P}^\beta(x_v = 1|x_u = 1)$ . The function  $g$  is very complicated, but there are few variables, so a good general optimization code should be able to solve the problem. In Section 7 we report this approach for a realistic test case, and where we use the routines in the Matlab Optimization Toolbox for minimizing  $g$ .

We now discuss the case when  $n$  is too large for the approach above to be appropriate. Then one might split the graph into smaller pieces and use the procedure above on each part. After this one may proceed with an adjustment procedure which we now discuss in detail. This procedure is an adaption of the parameter estimation method described in [3] and [4] for Markov Random Field models in the image segmentation area. Let  $\theta$  be a vector containing all  $\alpha$ 's and  $\beta$ 's to be estimated, after (perhaps) setting some parameters manually and grouping agents so that fewer parameters are needed. Again, we assume the judgement vector  $z$  is given (see above). Note that finding maximum likelihood estimator (MLE) in our model, when  $z$  is given, is not possible because the default state vector  $x$  is also a part of our model (1) and (2). We shall instead use the so-called pseudo-likelihood function.

Consider the following procedure where, throughout,  $z$  is the fixed judgement vector:

1. Choose an initial estimate  $\hat{\theta}$  of  $\theta$ .
2. For the given  $z$  and the current parameter  $\hat{\theta}$ , solve the MLDS problem to obtain the maximum a posteriori solution  $\hat{x}$ .

3. For the given  $z$  and the current vector  $\hat{x}$ , find a maximum pseudo-likelihood (MPLE) estimate  $\hat{\theta}$  of  $\theta$  in the model (1), (2).
4. Return to Step 2 until convergence or a maximum number of iterations is reached.

The MPLE problem in Step 3 may be treated as follows. Let  $f(x, z; \theta)$  denote the joint probability density of the random variable  $(x, z)$  for the given parameter  $\theta$ . Then from (1), (2) (where dependency of  $\theta$  is indicated in the notation)

$$\begin{aligned} f^\theta(x, z) &= \pi^\theta(x) f(z|x) \\ &= (1/s^\theta) \cdot e^{\psi^\theta(x)} \cdot f(z|x) \end{aligned} \tag{5}$$

where

$$\psi^\theta(x) = \sum_{v \in V} [\alpha_v^1 x_v + \alpha_v^0 (1 - x_v)] + \sum_{(u,v) \in E} \beta_{uv} I(x_u \leq x_v).$$

Note, in particular, that the summation constant  $s = s^\theta$  in the prior  $\pi^\theta(\cdot)$  does depend on  $\theta$ . But the conditional distribution  $f(z|x)$  is independent of  $\theta$ . Let now  $z$  and  $x$  be fixed and consider the problem of maximizing  $f^\theta(x, z)$ . It is equivalent to maximize  $\psi^\theta(x) - \log(s^\theta)$  w.r.t.  $\theta$ . Since  $n$  is assumed large,  $s^\theta$  is hard to compute, and we adopt the maximum pseudo-likelihood approach (see [4]). The pseudo-likelihood  $p^\theta(x, z)$  is defined by

$$p^\theta(x, z) = \prod_{v \in V} f^\theta(x_v | x_u (u \neq v), z)$$

so this is the product of the conditionals of each variable  $x_v$  given the remaining variables  $x_u$  for  $u \neq v$  (and  $z$  which is fixed here). The pseudo-likelihood serves as an approximation to the likelihood function. A maximum pseudo-likelihood estimator (MPLE)  $\hat{\theta}$  maximizes  $p^\theta(x)$  (w.r.t.  $\theta$ ) for the given  $x$ . The main advantage of the MPLE approach is that one gets around the difficulty caused by the normalization constant  $s^\theta$ . In fact, we have from (5)

$$\begin{aligned} &f^\theta(x_v = 0 | x_u (u \neq v), z) \\ &= \frac{f^\theta(x_v=0, x_u(u \neq v), z)}{f^\theta(x_v=0; x_u(u \neq v), z) + f^\theta(x_v=1; x_u(u \neq v), z)} \\ &= \frac{f_v(z_v | x_v=0) \cdot e^{\psi^\theta(x_v=0, x_u(u \neq v))}}{f_v(z_v | x_v=0) \cdot e^{\psi^\theta(x_v=0, x_u(u \neq v))} + f_v(z_v | x_v=1) \cdot e^{\psi^\theta(x_v=1, x_u(u \neq v))}} \end{aligned}$$

which is independent of  $s^\theta$  (here  $1/s^\theta$  and  $\prod_{u \neq v} f_u(z_u|x_u)$  canceled). And, of course,  $f^\theta(x_v = 1|x_u(u \neq v), z) = 1 - f^\theta(x_v = 0|x_u(u \neq v), z)$ . So the conditional distribution can be computed efficiently (without knowing  $s^\theta$ ). This means that finding a MPLE becomes the nonlinear optimization problem of maximizing the logarithm of the pseudo-likelihood given above. Having in mind that there are rather few variables, and that a near-optimal solution is good enough, this problem should be solvable by standard techniques in nonlinear optimization. For further general comments and references on MPLE and the relation to MLE, see [4].

## 6 Optimal recovery policies

So far we have considered the problem of finding a most likely default set in our financial system model. In this section we build on this model in order to discuss an interesting extension of this problem: how can we “optimally” support the financial system by providing economic support to some subset of the agents. We present a mixed integer linear programming model for this problem of finding an *optimal recovery policy*.

First we discuss a variation of the most likely default set problem MLDS. Consider again MLDS in (3) and let  $R \subseteq V$ . The *R-saved MLDS problem* is

$$(MLDS(R)) \quad \max\{F(x) : x \in \{0, 1\}^n, x_v = 0 \text{ for all } v \in R\} \quad (6)$$

An optimal solution  $\hat{x}$  is the incidence vector of a subset  $S \subseteq V$  satisfying  $S \cap R = \emptyset$ , and we then call  $S$  an *R-saved most likely default set*. Thus, we “force” the agents in  $R$  to survive; otherwise the problem is similar to the MLDS problem.

The *R-saved MLDS problem* may also be solved as a minimum cut problem as discussed next. Consider a fixed set  $R \subseteq V$  and consider again the directed graph  $G'$  defined in Section 3. A cut in  $G'$  corresponds to a partition  $(S, T)$  of  $V$ , so the cut is  $\delta^+(S \cup \{s\})$ . If  $R \subseteq T$  we call such a cut an *sR-cut*. Then, with the same arguments as in Section 3, it is clear that the *R-saved MLDS problem* corresponds to finding a minimum weight *sR-cut*, and this problem reduces to a minimum  $s\hat{t}$ -cut problem in a smaller graph obtained from  $G'$  by shrinking  $R \cup \{t\}$  into a single vertex  $\hat{t}$ .

Moreover, the following result holds on the connection between the *R-saved MLDS problem* and the ordinary MLDS problem.

**Theorem 4** *Let  $S^*$  be the unique maximum cardinality most likely default set in the MLDS problem. Let  $R \subseteq V$ . Then there is a unique R-saved most likely default set  $S_R$  of maximum cardinality. Moreover,  $S_R \subseteq S^*$ .*

**Proof.** As in the proof of Theorem 2 we let  $f(S)$ , for each  $S \subseteq V$ , denote the weight of the  $st$ -cut  $\delta^+(S \cup \{s\})$  in  $G'$  (using the nonnegative weights  $w'_e$ ,  $e \in E'$ ). Suppose  $S_R \not\subseteq S^*$ . Since  $\delta^+(S_R)$  is an  $sR$ -cut,  $R \cap S_R = \emptyset$ . If  $R \cap S^* = \emptyset$  then  $S_R = S^*$  (since  $S^*$  is optimal without the restriction corresponding to  $R$ ), a contradiction. So, assume  $R \cap S^* \neq \emptyset$ . Clearly,  $f(S_R) \geq f(S^*)$ . By submodularity we have  $f(S^* \cup S_R) + f(S^* \cap S_R) \leq f(S^*) + f(S_R)$ . Both  $\delta^+(S^* \cup S_R \cup \{s\})$  and  $\delta^+((S^* \cap S_R) \cup \{s\})$  are  $st$ -cuts, and so  $f(S^* \cup S_R) \geq f(S^*)$  and  $f(S^* \cap S_R) \geq f(S^*)$  (by optimality of  $S^*$ ), which implies  $f(S^* \cup S_R) \leq f(S_R)$  and  $f(S^* \cap S_R) \leq f(S_R)$  (due to the mentioned submodularity). Now,  $\delta^+((S^* \cap S_R) \cup \{s\})$  is an  $sR$ -cut, since  $(S^* \cap S_R) \cap R = \emptyset$ . As a consequence  $f(S^* \cap S_R) = f(S_R)$  and thus (again by the submodularity and optimality of  $S^*$ )  $f(S^* \cup S_R) = f(S^*)$ . So  $S^* \cup S_R$  is a most likely default set. Since  $S_R \not\subseteq S^*$ , we have  $|S^* \cup S_R| > |S^*|$ , a contradiction. Uniqueness of  $S_R$  can be easily derived as in the proof of Theorem 2.  $\square$

This theorem is important because it says that fixing (forcing) certain agents to survive does not add any agents to the maximum cardinality set of most likely agents that default.

Now, to each agent  $v \in V$  we associate a *survival benefit*  $b_v$  and a *recovery cost*  $c_v$ . This means that “the administration” can assure that  $v$  survives by paying a cost of  $c_v$ . Moreover, for each surviving agent  $v$  there is a survival benefit  $b_v$ . Suppose the administration can allocate a limited amount of money to the system, say  $K$ . We want to find an optimal allocation, i.e., we want to establish which agents to be recovered (meaning: avoid defaulting) in order to maximize the social benefit. This *optimal recovery problem (ORP)* can be restated as the problem of finding a set  $R \subseteq V$  of agents to be recovered so that the total benefit of the *most likely* nondefaulting (surviving) set of agents is maximized, and the total cost does not exceed  $K$ . Clearly, if  $R$  is the set of recovered agents, and  $T$  the *most likely* set of surviving agents, then  $R \subseteq T$ .

We consider again the auxiliary graph  $G'$  defined in Section 3. A cut in  $G'$  corresponds to a partition  $(S, T)$  of  $V$ , so the cut is  $\delta^+(S \cup \{s\})$ . Note that in order to recover a set  $R$  of agents, we can restrict the attention to the case where  $R \subseteq T$ . The reader should keep in mind the interpretation that vertices in  $S$  default while vertices in  $T$  survive.

As already mentioned, for fixed  $R \subseteq V$ , the maximizer of the (logarithm of the) likelihood function  $L$  in (4) can be found by solving a minimum cut problem on a new graph obtained from  $G'$  by shrinking the vertices in  $R$  and vertex  $t$  into a single vertex. However, for our purposes it is convenient

to let  $G'$  remain unchanged and modify arc weights instead. We define new arc weights  $w^R \in \mathbb{R}^{E'}$  by letting  $w_{ut}^R = w'_{ut} + M$ , for  $u \in R$  and  $w_{uv}^R = w'_{uv}$  otherwise, where  $M$  is any constant satisfying  $M > \sum_{sv \in \delta^+(s)} w'_{sv} = \sum_{sv \in \delta^+(s)} w_{sv}^R$ . By this construction it is clear that no arc  $(u, t)$ , for  $u \in R$ , belongs to a minimum  $st$ -cut of  $(G', w^R)$ . In other words, if  $(S^*, T^*)$  is the partition of  $V$  induced by a minimum  $st$ -cut in  $(G', w^R)$ , then  $R \subseteq T^*$ .

Finally, similarly to Proposition 3, one can show that there exist weights  $w'$  such that, for any  $R$ , there is a unique minimum  $st$ -cut in  $(G', w^R)$ . From now on we assume that this condition is satisfied.

The following result says that the ORP problem is computationally hard.

**Theorem 5** *The optimal recovery problem is NP-hard.*

**Proof.** The proof is a reduction from the knapsack problem ([18]). Given non-negative real numbers  $b_1, \dots, b_n$  (*benefits*),  $c_1, \dots, c_n$  (*costs*) and  $K$  (*budget*), the knapsack problem amounts to finding a subset  $R^* \subseteq \{1, \dots, n\}$  such that  $\sum_{i \in R^*} c_i \leq K$  (i.e. the budget is not violated) and  $\sum_{i \in R^*} b_i$  (overall benefit) is maximized.

Let us define an instance  $G' = (V', E')$  of ORP as follows.  $V' = V \cup \{s, t\}$ , where  $V = \{1, \dots, n\}$  and  $E' = \{si : i \in V\} \cup \{it : i \in V\}$ . With each  $i \in V$  associate a cost  $c_i$  and a benefit  $b_i$ , and suppose the budget is  $K$ . Finally, associate with the arcs of  $G'$  the following weights: for  $i \in V$ , let  $w'_{si} = k_s$  and  $w'_{it} = k_t$ , with  $k_s > k_t > 0$ . With this setting, for every set  $R$ , there is a unique minimum  $sR$ -cut in  $G'$ , namely  $\{si : i \in R\} \cup \{it : i \in V - R\}$ . In other words, for every choice of  $R$ , the set  $T$  of surviving agents coincides with  $R$ . So, an optimal solution to ORP is simply a set  $R^*$  satisfying the budget constraint and maximizing the overall benefit, which solves the knapsack problem.  $\square$

Next, we formulate ORP as a bilevel optimization problem. To this end, let  $y_v$  be a binary variable which is 1 iff  $v \in R$  (i.e.  $v$  is recovered), and 0 otherwise. Also, let  $q$  be the incidence vector of an  $st$ -cut of  $G'$ . According to the modelling assumptions introduced in Section 3,  $q_{sv} = 1$  ( $q_{vt} = 1$ ) implies that  $v \in V$  survives (defaults); also,  $q_{sv} = 1 - q_{vt}$ . ORP can be formulated as the following bilevel program:

$$\begin{aligned}
& \max && \sum_{v \in V} b_v q_{sv}(y) \\
& \text{subject to} && \\
& && \sum_{v \in V} c_v y_v \leq K \\
& && y \in \{0, 1\}^n
\end{aligned} \tag{7}$$

where  $q(y)$  is the unique optimal solution to

$$\min_{q \in \mathcal{C}_{st}} \sum_{uv \in E'} w'_{uv} q_{uv} + \sum_{v \in V} M y_v q_{vt} \quad (8)$$

and  $\mathcal{C}_{st}$  is the family of incidence vectors of  $st$ -cuts of  $G'$ . In fact, if we denote by  $y^R$  the incidence vector of  $R \subseteq V$ , then  $q(y^R)$  is the incidence vector of a minimum  $st$ -cut in  $(G', w^R)$ , and it corresponds to a maximum cardinality  $R$ -saved most likely default set. Also, if  $y^R$  satisfies the budget constraint in (7), then the cost of  $R$  does not violate the budget and all agents in  $R$  can be recovered. Finally, the objective function in (7) amounts to maximizing the value of the surviving agents.

It is important to recall that, for each  $y \in \{0, 1\}^n$ , the vector function  $q(y)$  above is well-defined. Incidentally observe that our assumption on the weights  $w'$  corresponds to what is known as the *pessimistic paradigm* [7]. In fact, among all possible solutions to (6) with the same  $F$  value, we choose the one with the maximum number of defaulting agents.

Similarly to [15] we show how to reduce the above bilevel program to a mixed integer linear programming problem. First observe that the inner problem can be formulated as a linear programming problem (LP). To this end, suppose  $y$  is fixed to some  $(0, 1)$ -vector and introduce the vertex potential vector  $p \in \mathbb{R}^{V'}$ . Then (8) can be written as the following LP:

$$\begin{aligned} \min \quad & \sum_{uv \in E'} w'_{uv} q_{uv} + \sum_{v \in V} M y_v q_{vt} \\ \text{subject to} \quad & p_v - p_u + q_{uv} \geq 0 \quad (uv \in E') \quad (9) \\ & p_s = 1, p_t = 0 \\ & q \geq 0, p \in \mathbb{R}^{V'}. \end{aligned}$$

Let  $F_P$  be the feasible set of (9). Then  $F_P$  is a polyhedron and its vertices are  $(0, 1)$  vectors, in one-to-one correspondence with the  $st$ -cuts of  $G'$  (see [19]). Since all weights in the objective function are strictly positive, the (unique) optimal value of (9) is attained in a vertex  $(q^*, p^*)$  and  $q^*$  is the incidence vector of an  $st$ -cut of  $G'$  (i.e.  $q^* \in \mathcal{C}_{st}$ ). The dual of the above LP

is:

$$\begin{aligned}
& \max && \sum_{su \in \delta^+(s)} f_{su} \\
& \text{subject to} && \\
& && \sum_{uv \in \delta^+(u)} f_{uv} - \sum_{vu \in \delta^-(u)} f_{vu} = 0 \quad (u \in V) \quad (10) \\
& && 0 \leq f_{uv} \leq w'_{uv} \quad (uv \in E'') \\
& && 0 \leq f_{vt} \leq w'_{vt} + My_v \quad (v \in V)
\end{aligned}$$

where  $E'' = \{uv \in E' : v \neq t\}$ . Let  $F_D$  be the feasible set of (10). Then we know that two feasible solutions  $(q^*, p^*)$  in (9) and  $f^*$  in (10) are both optimal in the respective problems if and only if the corresponding objective functions coincide, i.e.

$$\sum_{uv \in E'} w'_{uv} q_{uv}^* + \sum_{v \in V} My_v q_{vt}^* = \sum_{su \in \delta^+(s)} f_{su}^*. \quad (11)$$

We can then rewrite the bilevel program as the following nonlinear integer program:

$$\begin{aligned}
& \max && \sum_{v \in V} b_v q_{sv} \\
& \text{subject to} && \\
& && \sum_{v \in V} c_v y_v \leq K \quad (12) \\
& && \sum_{uv \in E'} w'_{uv} q_{uv} + \sum_{v \in V} My_v q_{vt} = \sum_{su \in \delta^+(s)} f_{su} \\
& && y \in \{0, 1\}^V, (q, p) \in F_P, f \in F_D.
\end{aligned}$$

This program is nonlinear due to the presence of the term  $y_v q_{vt}$ . A standard way to linearize it requires the introduction of a new variable  $l_v = y_v q_{vt}$  for all  $v \in V$  (see [15]). However, we can take advantage of the special structure of our problem. In particular, let  $(y^*, q^*, p^*, f^*)$  be an optimal solution of (12), with  $q^* \in \mathcal{C}_{st}$ . Now, we have  $\sum_{su \in \delta^+(s)} f_{su}^* \leq \sum_{su \in \delta^+(s)} f_{su} \leq \sum_{su \in \delta^+(s)} w'_{sv} < M$ . This implies, by (11), that  $y_v^* q_{vt}^* = 0$ , which in turn can be expressed by the constraint  $y_v + q_{vt} \leq 1$ , for all  $v \in V$ .

Thus, we may remove the nonlinear term  $y_v^* q_{vt}^*$  in (12), and we have therefore obtained the following valid mixed integer linear programming for-

mulation of the optimal recovery problem:

$$\begin{aligned}
& \max && \sum_{v \in V} b_v q_{sv} \\
& \text{subject to} && \\
& && \sum_{v \in V} c_v y_v \leq K \\
& && \sum_{uv \in E'} w'_{uv} q_{uv} = \sum_{su \in \delta^+(s)} f_{su} \\
& && y_v + q_{vt} \leq 1 \quad (v \in V) \\
& && y \in \{0, 1\}^V, (q, p) \in F_P, f \in F_D.
\end{aligned} \tag{13}$$

The above program can be solved for medium-sized instances by means of the commercial optimization solver ILOG CPLEX. We provide an example in the next section.

## 7 Computational experiments

This section has two major purposes. On one hand, we construct a number of realistic but artificial test cases to analyze the ability of the model to predict future crises and handle it. Also, the impact of systematic variations of the model parameters on the overall response is investigated. On the other hand we analyze a real case involving the major Scandinavian banks to assess the ability of our approach to cope with real-world problems. We forecast most likely default agents and identify possible recovery policies in this example.

### 7.1 Experiments on realistic scenarios

In this subsection we report on a number of tests performed on some small, but interesting problems. The main purpose is to show that our approach can be used to indicate the risk of future crises and how to handle it. In addition, we also investigate the effect of systematic changes of the parameters to better understand the role played in the model.

We decided to set up different scenarios that reflect some typical features of realistic financial systems and their systemic risk, thereby showing the potential usefulness of our approach. The model parameters were set manually, but we also tested the parameter estimation procedure introduced in Section 5. We do not report extensive computational results for random test cases since, in our opinion, this has limited interest.

In order to build a suitable test-bed for our experiments, we generated some small-sized problems (instances), corresponding to some interesting scenarios. To generate such instances we performed the following steps:

1. Specify the conditional probability distribution  $f(z|x)$ .
2. Select the prior parameters  $\alpha$  and  $\beta$ .
3. Find a default state vector  $\bar{x}$  which maximizes, or nearly maximizes, the prior  $\pi$ .
4. Draw an expert judgement vector  $\bar{z}$  according to the conditional distribution  $f(z|x)$ .

The choice of  $\alpha$  and  $\beta$  in Step 2 is discussed below for different scenarios. In Step 1 we used again two continuous distributions (for each agent  $i \in V$ ), namely  $f(z_i|x_i = 0) = ce^{-\gamma z_i}$  and  $f(z_i|x_i = 1) = ce^{\gamma(z_i-1)}$ , where  $c = \gamma/(1 - e^{-\gamma})$  and  $\gamma$  is a given constant. To draw the observation vector  $\bar{z}$  in Step 4 we first draw with uniform probability a state  $\bar{x}$  among the 6 most likely ones according only to the prior probability distribution  $\pi$  in (2). Note that, since we are dealing with small instances, the number of status vectors  $2^n$  is not too large, and may be enumerated. The state  $\bar{x}$  and the conditional probability distributions are then exploited to draw  $\bar{z}$ .

Next we discuss more in detail our test cases. In all cases, the state  $x^*$  corresponding to the most likely default set is obtained by a single min-cut computation. We will present the following three scenarios:

- (a) The bad path scenario
- (b) The effect-vertex scenario
- (c) The effect-vertex-many-banks scenario

**(a) The bad path scenario.** This example illustrates how defaults may develop in a chain (path) of agents, each dependent only on its predecessor. The path consists of 10 vertices, and the graph is shown in Figure 1.

The default prior probability  $p_v = e^{\alpha_v}$  for  $v = 2, \dots, 9$  is drawn independently with uniform distribution in the interval  $[0.4, 0.5]$ . Also, we let  $p_1 = 0.995$  and  $p_{10} = 0.2$ . Finally, we let  $\beta_{4,5} = k$  whereas all other  $\beta$ s are  $2k$ . For low values of  $k$ , only vertex 1 defaults. When  $k$  increases (up to approx. 3), then vertices 1 to 4 default. Finally, for larger values of  $k$  all vertices default.



Figure 1: The bad path scenario.

**(b) The effect-vertex scenario.** We consider 18 agents: 16 clients, one bank and one *effect vertex* (see Figure 2). Each client  $v$  obtains a loan from the bank; denote by  $\beta^l$  the (constant) dependency of the bank on every client. The effect vertex represents the influence of the interest rate on the clients, which are partitioned into 6 *good clients* and 10 *bad clients*. The dependency of the good clients on the interest rate is given by  $\beta^g$  and is the same for all good clients. The dependency of the bad clients on the interest rate is given by  $\beta^b$  and is the same for all bad clients.

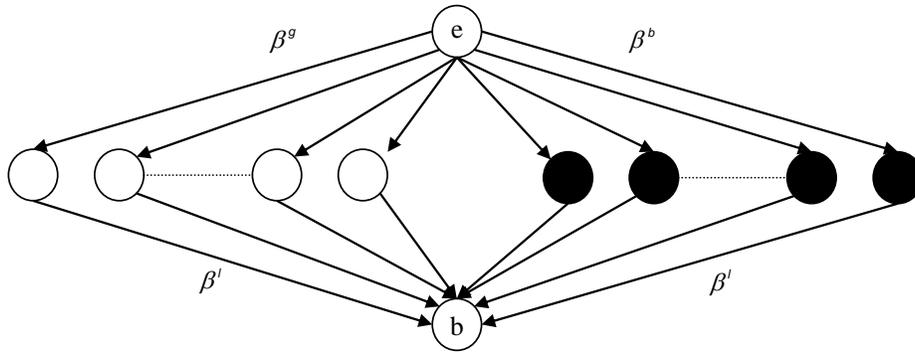


Figure 2: The effect-vertex scenario.

The default prior probability  $p_v = e^{\alpha_v}$  for each client  $v$  is drawn independently with uniform distribution in the interval  $[0.3, 0.5]$  for the good clients, whereas is set equal to 0.5 for the bad ones. High interest rates are represented by a high prior default probability for the effect vertex (0.995). The bank vertex has a low prior default probability (0.2). Also we let  $\beta^l = 0.25$ , whereas  $\beta^g = 0.1$ .

In order to analyze the bank reaction to increasing influence of the interest rate on the bad clients, we let  $\beta^b$  grow from 0 to a large value. Then, for low values of  $\beta^b$ , only the effect vertex is default. When  $\beta^b$  reaches a threshold value (here 0.15), then all bad clients and the bank are default, while for intermediate values only a subset of the bad clients default.

As a variation of scenario (b), we tested the parameter estimation procedure described in Section 5. To this end we replace  $\beta^l$  (which denotes the constant dependency of the bank on the clients) with two different parameters, namely  $\beta_g^l$  and  $\beta_b^l$ , denoting the dependency on the good and the bad clients, respectively. We consider as fixed all parameters except for  $\beta^b$  and  $\beta_b^l$ ; the parameters involving bad clients. The values for the fixed parameters are those given above. As initial values on the two unknown parameters we used  $\beta^b = 0.1$  and  $\beta_b^l = 0.25$  and then the parameter estimation procedure of Section 5 was applied.

With this choice we obtained the MLE estimate  $\beta_b^l = 0.08$  and  $\beta^b = 0$ . The procedure converged in one step. The resulting default state vector  $\hat{x}$  (the MLDS solution) has only 5 defaulting clients (all “bad” ones), and the bank is non-defaulting. (The default state vector was the same for these two parameter vectors).

Since we have only two “variable parameters” to be determined, it is possible to visualize the corresponding pseudo-likelihood function  $p^\theta(z, \hat{x})$ , as a function of  $\beta^b$  and  $\beta_b^l$  for the fixed  $x$  (just found). The values of this function is shown in Figure 3 for  $\beta_b^l \in [0, 0.002]$  and  $\beta^b \in [0, 0.2]$ .

**(c) The effect-vertex-many-banks scenario.** In this case we have 12 clients, 3 banks and one *effect vertex* (see Figure 4). Each client obtains a loan from one bank; denote by  $\beta^l$  the (constant) dependency of the banks on every client. Again, the effect vertex represents the influence of the interest rate on the clients, which are partitioned into *good clients* and *bad clients*. The dependency of the good (bad) clients on the interest rate is given by  $\beta^g$  ( $\beta^b$ ). All the bad clients got loans from the same bank (*bad bank*). Finally, the three banks are also (strongly) interrelated, and we denote by  $\beta^i$  the corresponding value of the parameter (which is equal for all pairs).

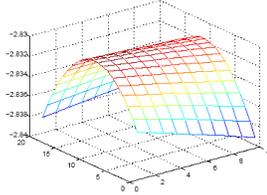


Figure 3: Pseudo-likelihood

As for the effect-vertex scenario, the default prior probability  $p_c = e^{\alpha_c^1}$  for each client  $c$  is drawn independently with uniform distribution in the interval  $[0.3, 0.5]$  for the good clients, and is equal to 0.5 for the bad ones. High interest rates are represented by a high prior default probability for the effect vertex (0.995). The bank vertex have a low prior default probability, equal to 0.2 for all banks. We have fixed initial values for  $\beta^b = 0$ ,  $\beta^g = 0.1$ , and  $\beta^i = 0.3$  whereas  $\beta^l$  is chosen randomly (uniformly) in the interval  $[0.25, 0.5]$ . The fraction of bad clients is 50 %.

In order to understand the banks reaction to increasing influence of the interest rate on the bad clients, we let  $\beta^b$  and  $\beta^i$  grow. For low values of  $\beta^b$  and  $\beta^i$ , only the effect vertex is default; when  $\beta^b$  reaches a first threshold value (0.15), then all bad clients and the bad bank are also default. Finally, for increasing values of  $\beta^i$ , all banks will default.

Figure 5 shows more in detail the posterior probability distribution for the first 16 states, where, for each good (bad) client  $\beta^g = 0.3$ ,  $\beta^b = 0.6$  and  $\beta^i = 1.2$ .

Each bar corresponds to one of the 16 states. The  $y$ -axis represents the (posterior) probability, whereas on the  $x$ -axis the pair  $(nc : nb)$  below each bar denotes the number of defaulting clients ( $nc$ ) and banks ( $nb$ ), respectively. States are ordered from 1 to 16 according to decreasing probabilities. The picture clearly shows that the first two states, with 2 and 3 defaulting clients and no defaulting banks, respectively, are much more likely than the remaining ones. In other words, the probability drops very quickly and it is concentrated in the first state. Also, State 1 and State 2 are very similar to each other, as they only differ by one client (i.e. the *Hamming distance* is

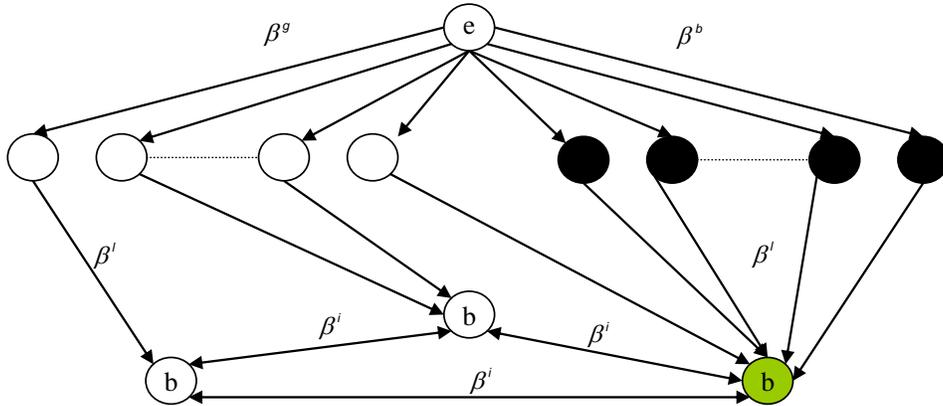


Figure 4: The effect-vertex-many-banks scenario.

one). It is interesting to compare the above figure with the one corresponding to the prior probabilities of the same 16 states.

Observe that state 5 is the most likely state according to the prior distribution. Comparing Figure 6 with Figure 5, we observe that the effect of the “experts judgment” is twofold. By one side, the ranking imposed by the prior knowledge is substantially modified by the posterior information. By the other side, the overall distribution gets more “peaked”, concentrating just on at most the first two states. Finally, in Figure 7 we show the loss probability distribution, i.e. the probability that a given number  $q$  of agents (clients or banks) defaults. Observe that the most likely are obtained for  $q = 4$  with  $q = 3$  following next.

Again, the probability is represented on the  $y$ -axis, while the  $x$ -axis is the number of defaulting clients. No banks are defaulting in this example.

The next three figures are analogous to the three previous ones, but with  $\beta^b = 0.7$  and  $\beta^i = 1.4$ . We can see that these values dramatically affects the results of our simulation. In particular, due to the tighter connection between the bad bank and the bad clients, the bad bank is likely to default: as a consequence all other banks are also likely to default.

This last scenario provided the basis for an ORP scenario (optimal recovery problem). In particular we associated with each client a benefit of 20 units and a recovery cost equal to 5, whereas the benefit associated with each bank is 100 and its recovery cost is 80. The recovery cost of the effect-vertex

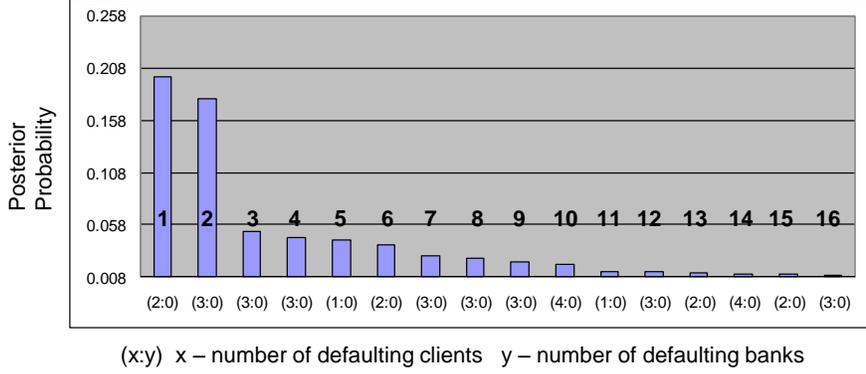


Figure 5: Posterior probability distribution for each of the 16 possible states of the effect-vertex-many-banks scenario, with  $\beta^b = 0.6$ ,  $\beta^i = 1.2$

is set to a very large value and we fixed the budget  $K = 20$ . Interestingly, by recovering a suitable subset of only 4 bad clients (with total cost equal 20), also the three banks got rescued.

Finally, in order to assess the possibility of handling larger ORP instances by direct application of CPLEX to its mixed integer formulation, we solved the effect-vertex-many-banks case for increasing number of agents (banks and clients). We fixed the number of banks to  $1/5$  of the total number of agents, the remaining ones were clients (bad and good in equal share). Since the number of different states is very large, we had to slightly modify the drawing of the observation state  $\bar{z}$  in Step 4 of our instance generation procedure. In fact, the state  $\bar{x}$  (from which  $\bar{z}$  is actually derived through the conditional probability distributions  $f(z|x)$ ) is now simply the most likely default set, which is computed by a single min-cut computation.

For different budget levels, CPLEX was able to solve to optimality instances up to 100 agents. For 150 agents, after 1 hour running time the solution found by CPLEX was at most 2% worse than the optimum one. However, the pernicious *tailing off* phenomenon occurs, with very little hope for further gap reductions within reasonable computing time. This shows that ORP can indeed be attacked for medium-size instances by a direct application of CPLEX. Larger instances will require a different solution ap-

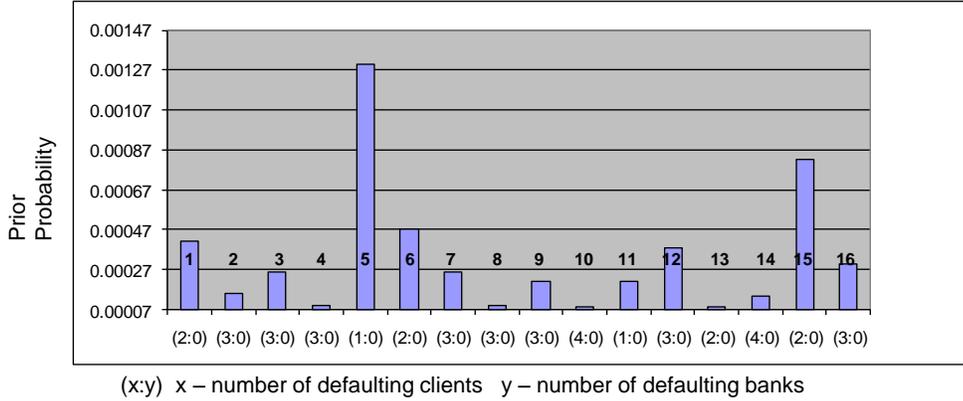


Figure 6: Prior probability distribution for the effect-vertex-many-banks scenario, with  $\beta^b = 0.6$ ,  $\beta^i = 1.2$

proach.

## 7.2 A real-world example

We consider here a realistic case involving the five larger commercial Scandinavian banks: Swedbank, Nordea, Danske Bank, DnBNOR and Handelsbanken. Swedbank and Handelsbanken are Swedish banks, while Danske Bank is Danish and DnBNOR is Norwegian. Nordea is a bank with branches in each of the Scandinavian countries, but with mother office in Sweden. All the banks are operating in the Scandinavian market, but mainly concentrated in their home countries. We suppose in our study that the banks are all operating in a Scandinavian market for simplicity. The banks are exposed to public loans which we model as the sixth agent in our case study. All the banks are dependent on each other through mutual lending.

The corresponding graphical model contains 6 vertices and 36 directed arcs. To compute the corresponding  $\beta_{uv}$  parameters, we apply the first parameter estimation procedure discussed in Section 5. In order to do this, we need to estimate default conditional probabilities and get hold of a suitable expert judgment vector  $z$ . For this purpose we base our procedures on publicly available data from the annual reports of 2009 along with credit ratings from Standard and Poor's (S&P).

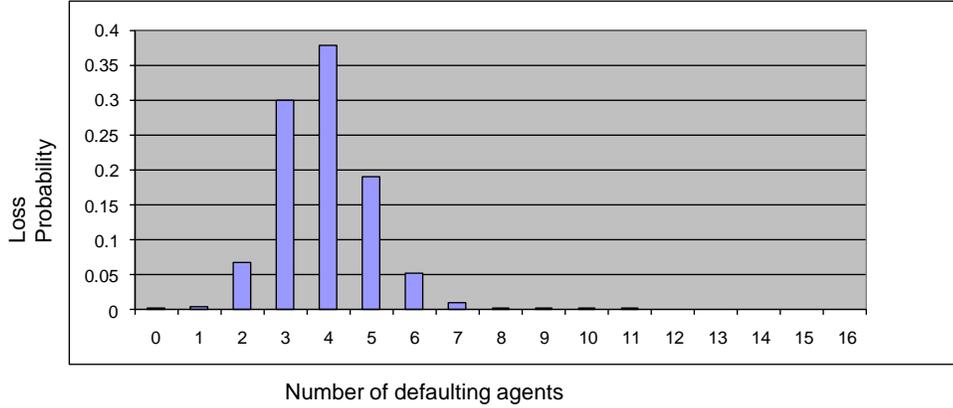


Figure 7: Loss probability distribution for the effect-vertex-many-banks scenario, with  $\beta^b = 0.6$ ,  $\beta^i = 1.2$

### Background financial figures.

The long-term S&P credit rating can be found on the web pages of the respective banks, and are shown in Table 1.

Bank	Swedbank	Nordea	Danske B.	DnBNOR	Handelsb.	Public
Rating	A	AA-	A	A+	AA+	C
Def. Prop.	0.008	0.006	0.008	0.007	0.006	0.25

Table 1: S&P ratings and corresponding default probability for the 5 major banks in Scandinavia and the Public Loans

For the Public Loans we have associated a credit rating C, in order to model a largely uncertain situation where households may start to default due to a crack in the property market. The actual situation in Scandinavia was overall rather stable, but in this example we would like to mimic an unrestful financial situation.

Unfortunately, we did not have accessible the default probabilities associated with the S& P credit ratings for the banking sector. Our figures in the second row labeled *Def. Prob.* are motivated from public reports made

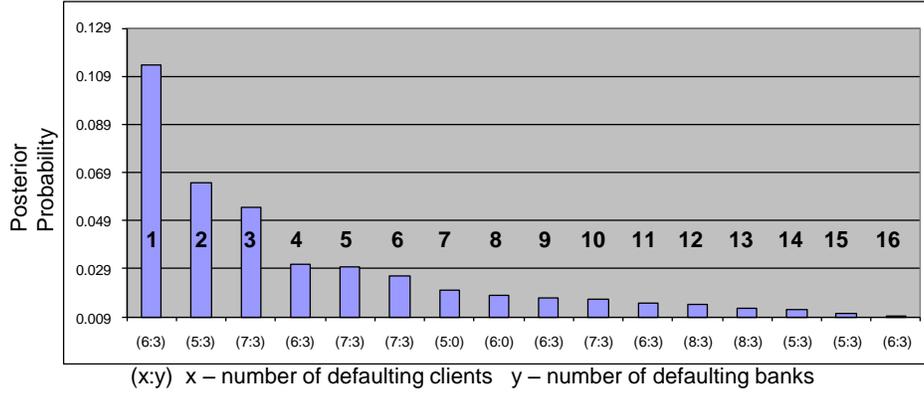


Figure 8: Posterior probability distribution for the effect-vertex-many-banks scenario, with  $\beta^b = 0.7$  and  $\beta^i = 1.4$

available through the web page of S& P giving historical figures for different sectors. These probabilities are interpreted as the expert judgment vector  $z$  in our context.

Table 2 shows, for each bank, some relevant figures taken from the annual report 2009, which can be easily downloaded from the bank website (all figures are expressed in Millions).

Bank	Swedbank	Nordea	Danske B.	DnBNOR	Handelsb.
Currency	SEK	EUR	DKK	NOK	SEK
Total Asset Value	1,794,687	507,544	3,098,477	1,823,453	2,122,843
Public Loans	1,290,667	282,411	1,127,142	1,114,886	1,477,183
Credit to Institutions	92,131	18,555	101,178	62,317	168,100
Central Deposit	37,879	11,500	101,178	31,859	49,882

Table 2: Figures from the annual reports of the 5 major banks in Scandinavia (in Millions)

The *Central Deposit* is the amount of money the bank has deposited in the Central Bank of their respective country. We point out that the figures

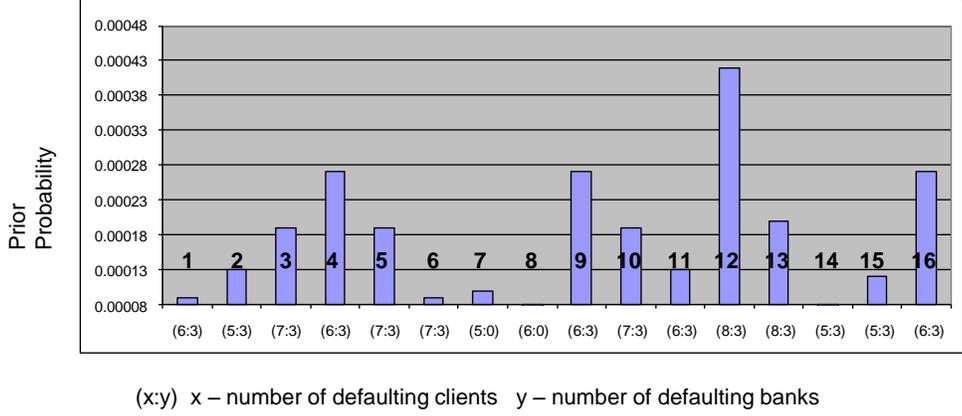


Figure 9: Prior probability distribution for the effect-vertex-many-banks scenario, with  $\beta^b = 0.7$  and  $\beta^i = 1.4$

in the two last rows for Danske Bank are obtained by equally distributing the cumulative figure given in the annual report, since they have reported their credit to institutions and the central bank in one cumulative number. This is of course just chosen by us as a proxy.

### Parameter estimation

The figures in Table 2 are exploited to derive suitable conditional default probabilities. These are in turn used as data for determining the coefficients of our model, according to the methodology described in Section 5.

First, for each bank  $B$ , the conditional probability for  $B$  to default given a default in Public Loans is computed as the ratio between the bank's exposure in Public Loans and its total asset value. Next, we use the figures in *Credit to institutions* as a proxy for the bank's exposure to the other banks, and we assume a uniform distribution of this figure among the four other banks: in other words, each bank has a fully diversified loan portfolio in the other banks. So, for instance, Swedbank's total exposure is 92,131 mill. SEK, which then under our assumptions gives an exposure of 23,032.75 mill. SEK with each of the four other banks. To measure the vulnerability of Swedbank to the four others, we compute the ratio between its exposure with its deposit in the central bank. For the case of Swedbank, this results

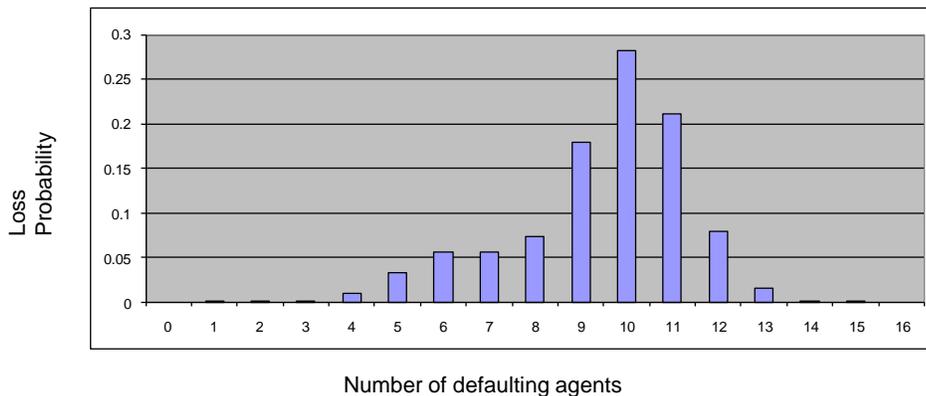


Figure 10: Loss probability distribution for the effect-vertex-many-banks scenario, with  $\beta^b = 0.7$  and  $\beta^i = 1.4$

in  $23,032.75/37,879 \simeq 0.608$ , which we use as a proxy for the conditional probability of Swedbank defaulting, given that one of the other default (for example, the default probability of Swedbank given that Handelsbanken defaults, is 0.608). We compute these conditional default probabilities for all the banks, and collect the figures in Table 3

Although the exposure in each of the other banks are minor compared to the total asset holdings, we think of the lending to credit institutions as part of the liquid portfolio of the bank, which makes it highly vulnerable if bigger parts of this fall out in the sense of another bank defaulting.

Bank	Swedbank	Nordea	Danske B.	DnBNOR	Handelsb.
Public Loans def. prob.	0.719	0.556	0.364	0.611	0.696
Other Bank def. prob.	0.608	0.403	0.750	0.489	0.842

Table 3: Conditional default probabilities

The conditional probabilities of Table 3 and the "expert judgments" provided by S&P and reported in Table 1 (Row *Def. Prob.*) are used as the input data to our parameter estimation procedure (Section 5). Since we do not have any *a priori* knowledge on the banks and the Public Loans de-

fault probability, we let  $\alpha_i^0 = \alpha_i^1 = 0$  for  $i = 1, \dots, 6$ : this makes vectors  $\alpha^0$  and  $\alpha^1$  irrelevant to determine the most likely defaulting set. Also, to define the conditional probability distribution  $f(z|x)$  we used two continuous distributions (for each agent  $i \in V$ ), namely  $f(z_i|x_i = 0) = ce^{-\gamma z_i}$  and  $f(z_i|x_i = 1) = ce^{\gamma(z_i-1)}$ , where  $c = \gamma/(1 - e^{-\gamma})$  and  $\gamma$  is a given constant.

### Computing most likely default sets and optimal recovery policies

By solving the MLDS problems described in Section 2 by using the current S&P judgment depicted in Table 1 and the corresponding estimated parameters, it turns out that the most likely default set is empty. This could be interpreted as a healthy banking sector in Scandinavia, despite of a rather low credit rating of Public Loans, as set by us. Although we have no reasons to believe that Danske bank is in a critical situation, we are motivated by the fact that there have been several defaults in the Danish bank sector to let them have the highest default probability. The Norwegian bank DnBNOR is supposed to have the smallest default probability since the financial situation in Norway is relatively much better than the other Scandinavian countries.

Things change dramatically when we investigate an artificial and much more pessimistic expert judgments vector like the one depicted in Table 4. This could be considered as a scenario in an extraordinary turbulent financial situation, possibly mimicing what happened to the Icelandic bank sector around 2008.

Bank	Swedbank	Nordea	Danske B.	DnBNOR	Handelsb.	Public
Prob.	0.55	0.46	0.75	0.27	0.36	0.65

Table 4: A pessimistic rating scenario

The corresponding most likely default set contains in this case all agents, namely all banks and the Public Loans. We then solve the associated ORP problem with decreasing budget  $K$  by solving the corresponding MILP programs (6). To this end, we set the survival benefit of each bank equal to its total asset value, whereas the Public Loans benefit equals the total amount of loans from the banks. Finally we let the recovery costs be 5% of the corresponding benefit values. All figures (expressed in Millions EUR) are shown in Table 5. For the figures below, we note that the exchange rate for Euro to NOK was 8.315, Euro to SEK 10.2667 and Euro to DKK 7.442. All figures are for January 1, 2010, and obtained from valutakurser.no.

Bank	Swedbank	Nordea	Danske B.	DnBNOR	Handelsb.	Public
Survival Benefit	174,806	507,544	416,350	219,296	206,769	837,544
Recovery Cost	8,740	25,377	20,817	10,964	10,338	41,877

Table 5: Benefit values and recovery costs

When the budget  $K = 21,000$  mill. EUR, then it is allocated to save Danske Bank only, but this is sufficient to rescue all the agents and so the final most likely default set is empty. Similarly, when  $K = 11,000$  mill. EUR,  $K = 10,500$  mill. EUR and  $K = 9,000$  mill. EUR, the budget is allocated to save DnBNOR, Handelsbank and Swedbank, respectively, and again the final most likely default set is empty! Observe that since the MLDS model does not try to minimize payout, the rescued bank(s) may in general not be the cheapest solution. For example, with  $K = 21,000$  any bank among Danske Bank, DnBNOR, Handelsbank and Swedbank could be rescued obtaining an empty most likely default set, but the solution returned by the solver was actually rescuing Danske Bank. Interestingly, one can show that forcing the allocation of a higher budget to save Nordea does not produce a similar effect. Indeed, in this case the most likely default set contains all other agents, and no agent is saved but Nordea itself.

## 8 Final comments

We mention some possible directions of future work in this area.

Concerning the model it would be natural to investigate the case when negative parameters  $\beta_{uv}$  are allowed. This models decreased risk of defaulting when an agent defaults. Note, however, that negative  $\beta_{uv}$ 's lead to computational difficulties (as the MLDS problem with arbitrary weights is NP-hard). Another important extension would be a dynamic model where companies might default at different times and the risk of defaulting would vary with time.

An interesting topic is a further study of the optimal recovery problem ORP. We showed how to reformulate the ORP into a valid mixed integer linear programming problem (13). We used standard methods to solve this problem, but one could investigate more advanced approaches where the specific structure of the problem is exploited in order to solve larger problems. Finally, ORP is very natural to study in a dynamic extension of our model.

**Acknowledgments.** We appreciate several very useful comments from the referees, in particular for improving the analysis given in Section 4. The authors also wish to thank Anders Øksendal (DnBNOR) and Geir Storvik (University of Oslo) for very useful discussions and suggestions in connection with Markov Random Field models.

## References

- [1] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Englewood Cliffs, New Jersey, Prentice-Hall, 1993.
- [2] J. Besag, Towards Bayesian image analysis, *Journal of Applied Statistics*, 16(3) (1989), 395–407.
- [3] J. Besag, On the statistical analysis of dirty pictures, *Journal of Royal Statistical Society, Series B* 48(3) (1986), 259–302.
- [4] J. Besag and J. Tantrum, Likelihood analysis of binary data in space and time, In P.J. Green, N.L. Hjort and S. Richardson (Eds.) *Highly Structured Stochastic Systems.*, Oxford Statistical Science Series, **27**, Oxford University Press, Oxford, 2003. (p.289–295)
- [5] R. Carmona, J.-P. Fouque and D. Vestal, Interacting particle systems for the computation of rare credit portfolio losses, *Finance Stoch.*, vol. 13 (2009), 613–633
- [6] G. Dahl, G. Storvik and A. Fadnes, Large-scale integer programs in image analysis, *Operations Research*, Vol.50 (2002), No.3, 490–500.
- [7] S. Dempe, *Foundation of Bilevel Programming*, Kluwer Academic (2002).
- [8] S. Deng, K. Giesecke and T. L. Lai, Sequential importance sampling and resampling for dynamic portfolio credit risk, *Manuscript*, (2010)
- [9] D. Duffie and K. J. Singleton, Modeling term structures of defaultable bonds *Rev. Financial Studies*, Vol. 12 (1999), 687-720.
- [10] I.O. Filiz, X. Guo, J. Morton, B. Sturmfels, Graphical Models for Correlated Defaults, *Technical report*, Sept. 2008, arXiv:0809.1393. To appear in *Mathematical Finance*.

- [11] S. Geman and D. Geman, Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 6(6) (1984), 721–741.
- [12] K. Giesecke, Credit risk modeling and valuation: an introduction, Technical report, Cornell University, Oct. 2004.
- [13] K. Giesecke and S. Weber, Credit contagion and aggregate losses, *J. Econ. Dyn. Control*, vol. 30 (2006), 741–767
- [14] R. A. Jarrow and S. Turnbull, Pricing Derivatives on Financial Securities Subject to Credit Risk, *J. Finance*, vol. 50, March (1995), 53–85
- [15] M. Labbé, P. Marcotte, and G. Savard, A Bilevel Model of Taxation and Its Application to Optimal Highway Pricing, *Management Science*, Vol 44 (1998), No. 12, 1608–1622
- [16] K.R. Mecke, D. Stoyan (Eds.) *Statistical Physics and Spatial Statistics*. Series: Lecture Notes in Physics, Vol. 554, Springer, 2000.
- [17] R.C. Merton, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29 (1974), 449–470.
- [18] G.L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*, Wiley, 1988.
- [19] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley, Chichester, 1998.
- [20] G. Storvik and G. Dahl, Lagrangian based methods for finding MAP solutions for MRF models, *IEEE Transactions on Image Processing*, Volume 9, No. 3 (2000), 469–479.