

Lifted Inference for Relational Continuous Models

Jaesik Choi and Eyal Amir

Computer Science Department
University of Illinois at Urbana-Champaign
Urbana, IL 61801
{jaesik,eyal}@illinois.edu

David J. Hill

Department of Civil and Environmental Engineering
Rutgers, The State University of New Jersey
Piscataway, NJ 08854
ecodavid@rci.rutgers.edu

Abstract

Relational Continuous Models (RCMs) represent joint probability densities over attributes of objects, when the attributes have continuous domains. With relational representation, they can model joint probability distributions over large numbers of variables compactly in a natural way. This paper presents the first exact inference algorithm for RCMs at a lifted level, thus it scales up to large models of real world applications. The algorithm applies to *relational pairwise models* which are (relational) products of potentials of arity 2. Our algorithm is unique in two ways. First, it is an efficient lifted inference algorithm. When Gaussian potentials are used, it takes only linear time while existing methods take cubic time. Second, it is the first exact inference algorithm which handles RCMs in a lifted way. The algorithm is illustrated over an example from Econometrics. Experimental results show that our algorithm outperforms both a ground-level inference algorithm and an algorithm built with previously-known lifted methods.

1 Introduction

Many real world systems are described by continuous variables and relations among them. Such systems include measurements in environmental-sensors networks, localizations in robotics, and economic forecasting in finance. Once a relational model among variables is given, inference algorithms can solve the problems of value prediction and classification.

At a ground level, inference with a large number of continuous variables is nontrivial. Typically, inference is the task of calculating a marginal over variables of interest. Suppose that a market index has a relationship with revenues of n banks. When marginalizing the market index out, the result is a function of n variables (revenues of banks), thus following marginalizations become harder. When n grows, the computation becomes expensive. When relations among variables follow Gaussian distributions, the computational complexity of the inference problem is cubic to the number of ground variables. Thus, computation with such models is limited to moderate-size models, preventing usage of such models for large, real-world applications.

To address these issues, Relational Probabilistic Languages (RPLs) (Ng and Subrahmanian 1992; Koller and Pfeffer 1997; Pfeffer et al. 1999; Friedman et al. 1999; Poole 2003; de Salvo Braz, Amir, and Roth 2005; Richardson and Domingos 2006; Milch and Russell 2006; Getoor and Taskar 2007) describe probability distributions at a relational level with the purpose of capturing larger models. RPLs combine probability theory for handling uncertainty and relational models for representing system structures. Thus, they facilitate construction and learning of probabilistic models for large systems. (Poole 2003; de Salvo Braz, Amir, and Roth 2005; Milch et al. 2008; Singla and Domingos 2008) showed that such models enable more efficient inference than possible with propositional graphical models, when inference occurs directly at a relational level.

Present exact lifted inference algorithms (Poole 2003; de Salvo Braz, Amir, and Roth 2006; Milch et al. 2008) and those developed in the efforts above are suitable for discrete domains, thus can in theory be applied to continuous domains through discretization. However, the precision of discretizations deteriorates exponentially with the number of dimensions in the model, and the number of dimensions in relational models is the number of ground random variables. Thus, discretization and usage of discrete lifted inference algorithms is highly imprecise.

Here, we propose the first exact lifted inference algorithm for Relational Continuous Models (RCMs), a new relational probabilistic language for continuous domains. Our main insight is that, for some classes of potential functions (or potentials), marginalizing out a ground random variable in a RCM can yield a RCM representation that does not force other random variables to become propositional. Further, *relational pairwise models*, i.e. products of relational potentials of arity 2, remain *relational pairwise models* after elimination of ground random variables in those models. Thus, it leads to the compact representation and the efficient computation. We provide conditions for relational pairwise models, and report a Gaussian potential.

Given a RCM, our algorithm marginalizes continuous variables by analytically integrating random variables except query variables. It does so by finding potentials and variables, eliminating them by *Inversion Elimination*. If such elimination is not possible, it eliminates each pairwise

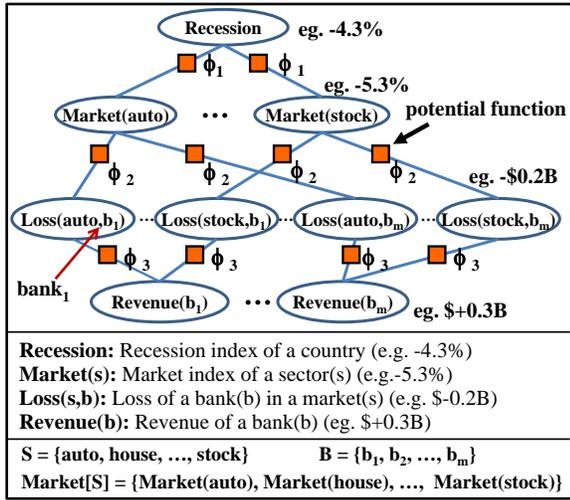


Figure 1: This figure shows a model over banks and market indices. *Recession*, *Market[S]*, *Loss[S,B]* and *Revenue[B]* are continuous variables whose range are $[-\infty, \infty]$. For example, *Market(stock)* is -5.3% , and *Loss(stock, B_m)* is $-\$0.2B$.

form in a linear time. If the marginal is not in pairwise form, it converts the marginal into a pairwise form.

This paper is organized as follows. Section 2 provides the formal definition of RCMs. Section 3 overviews our inference algorithms. Section 4 presents main intuitions and results in Gaussian potentials. Section 5 provides the generalized algorithm for arbitrary potentials. Section 6 provides experimental results followed by related works in Section 7. We conclude in Section 8.

2 Relational Continuous Models

We present a new relational model for continuous variables, Relational Continuous Models (RCMs). Relations among attributes of objects are represented by *Parfactor* models.¹ Each *parfactor* (L, C, A, ϕ) is composed of a set of objects (L), constraints on L (C), a list attributes of objects (A), and a potential on A (ϕ). Here, each attribute is a random variable with a continuous domain.

We define a *Relational Atom* to refer the set of ground attributes compactly. For example, *Revenue[B]* is a *relational atom* which refers to revenues of banks (e.g. $B = \{\text{'Pacific Bank'}, \text{'Central Bank'}, \dots\}$). To make the *parfactor* compact, a list of relational atoms is used for A . To refer to an individual random variable, substitutions are used. For example, if a substitution ($B = \text{'Pacific Bank'}$) is applied to a relational atom, then the relational atom *Revenue[B]* becomes a ground variable *Revenue('Pacific Bank')*. Formally, applying a substitution θ to a parfactor $g = (L, C, A, \phi)$ yields a new parfactor $g\theta = (L', C\theta, A\theta, \phi)$, where L' is obtained by renaming the variables in L according to θ . If θ is a

¹Part of its representation and terms are based on the previous (Poole 2003; Pfeffer et al. 1999; Milch and Russell 2006) However, our representation allows continuous random variables.

ground substitution, $g\theta$ is a factor. The set of *groundings* of a parfactor g is represented as $gr(g) = \{g\theta : \theta \in gr(L : C)\}$. We use $RV(X)$ to enumerate the random variables in the relational atom X . Formally, $RV(a) = \{\alpha[\theta] : \theta \in gr(L)\}$. $LV(g)$ refers the set of objects (L) in g .

The joint probability density over random variables is defined with *factors* in a *parfactor* model. A *factor* f is composed of A_g and ϕ . A_g is a list of ground random variables (i.e. $(X_1(\theta), \dots, X_N(\theta))$). ϕ is a *potential* on A_g : a function from $range(A_g) = \{range(X_1(\theta)) \times \dots \times range(X_N(\theta))\}$ to non-negative real numbers. The factor f defines a weighting function on a valuation ($v = (v_1, \dots, v_m)$): $w_f(v) = \phi(v_1, \dots, v_m)$. The weighting function for a *parfactor* F is the product of weighting function of all factors, $w_F(v) = \prod_{f \in F} w_f(v)$. When G is a set of *parfactors*, the density is the product of all factors in G :

$$w_G(v) = \prod_{g \in G} \prod_{f \in gr(g)} w_f(v). \quad (1)$$

For example consider the model in Figure 1. S and B in L are two objects which represent markets and banks respectively. S can be substituted by a specific market sector (e.g. $S = \text{'stock'}$). A parfactor $f_1 = (\{Market[S], Loss[S, B]\}, \phi_1)$ is defined over two relational atoms, *Market[S]* and *Loss[S, B]*. *Market(auto)* represents the quarterly market change (e.g. *Market(auto)* = -3.1%). *Loss(auto, Pacific Bank)* represents the loss of the bank in the auto market. Given two values, a potential $\phi_1(Market(auto), Loss(auto, Pacific Bank))$ provides the probability density.

3 Algorithm Overview for RCMs

One inference task with such models is to find the conditional density of query variables given observations of some variables in the model. Our inference algorithm, *First-Order Variable Elimination(FOVE)-Continuous*, recursively eliminates relational atoms as described in Figure 2.

First, it splits (terminology of (Poole 2003); *shattering* in (de Salvo Braz, Amir, and Roth 2005)) relational atoms such that groundings, $RV(X) RV(Y)$, of every pair of relational atoms, $X Y$, are disjoint. It introduces observations of groundings as separate relational variables. For example, observing $Market(auto) = 30\%$ creates two separate relational atoms: *Market(auto)*, *Market(M)_{M≠auto}*. The ' $M \neq auto$ ' then appears in parfactors relating to the latter relational atom. After *split*, *FIND-ELIMINABLE* finds a relational atom which satisfies conditions for one of the elimination algorithms: *Inversion-Elimination* (Sections 5.1) and *Relational-Atom-Elimination* (Section 5.2). The found atom is eliminated by our *ELIMINATE-CONTINUOUS* algorithm explained in Sections 4 and 5. It iterates the elimination until only query variables are remained.

Our main contributions are the algorithm **ELIMINATE-CONTINUOUS**, a lifted variable eliminations for continuous variables. We describe it in detail in Section 4 and 5.

4 Inference with Gaussian Potentials

This section presents our first main technical contribution, efficient variable elimination algorithms for relational Gaus-

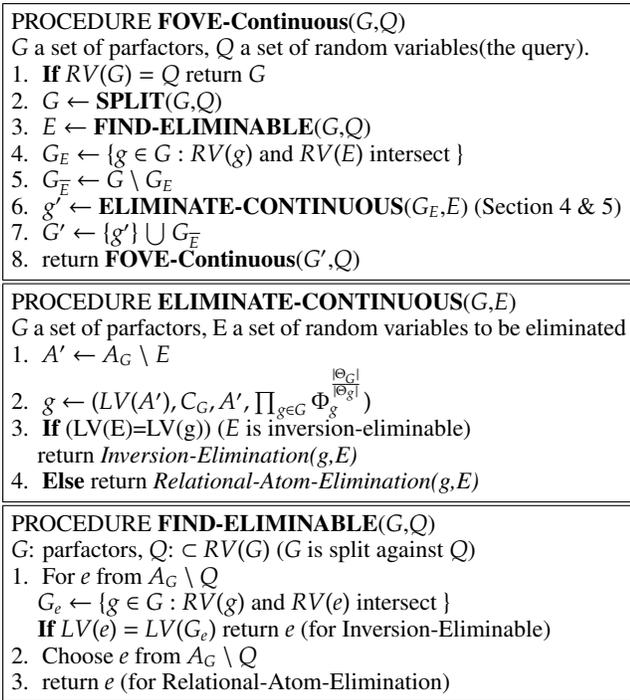


Figure 2: FOVE.Continuous (First-Order Variable Elimination with continuous variables) algorithm.

sian models. We focus on the inference problem of computing the posterior of query variables given observations for some random variables. Efficiently marginalizing out relational atoms is important for solving this inference problem.

4.1 Relational Pairwise Potentials

This section focuses on the product of potentials which we call *Relational Normals* (RNs). An RN is the following function with arity 2 (Section 5 provides a generalization for arbitrary potentials.):

$$\phi_{RN}(X, Y) = \prod_{x \in X, y \in Y} \phi_{RN}(x, y) = \prod_{x \in X, y \in Y} \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - y)^2}{2\sigma^2}\right)$$

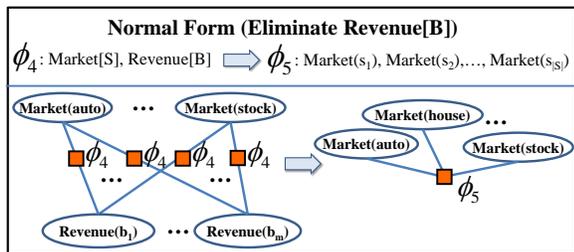


Figure 3: This figure shows a challenging problem in a RCM when eliminating a set of variables ($Revenue[B]$). Eliminating $Revenue[B]$ in ϕ_4 generates an integral ϕ_5 that makes all variables in $Market[S]$ ground. Thus, the elimination makes the RCM into a ground network.

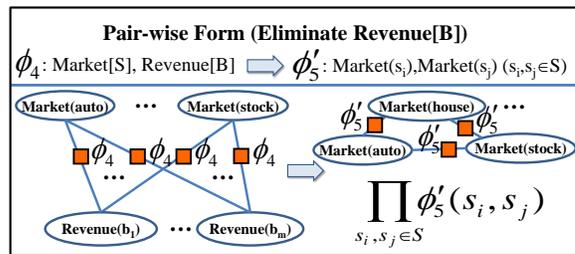


Figure 4: This figure shows our method for the problem shown in Figure 3. When eliminating $Revenue[B]$, we do not generate a ground network. Instead, we directly generate the pairwise form which allows the inference at the lifted level.

This potential indicates that the difference between two random variables follows a Gaussian distribution.

Consider the models shown in Figure 3 and 4. These random variables model the relationship between each market change and the revenue of each bank. To simplify notations, we respectively shorten $Market(s)$, $Loss(s, b)$ and $Revenue(b)$ to $M(s)$, $G(s, b)$ and $R(b)$ in equations. The potential ϕ_4 in these figures is $\phi_{RN}(M(s), R(b))$, and the complete model is $\prod_{s \in S, b \in B} \phi_{RN}(M(s), R(b))$

Figure 4 shows that marginalizing out a random variable $R(b_i)$ from the joint density results in the product of RNs again (c and c' are constants).² Formally,

$$\begin{aligned} \int_{R(b_i)} \prod_{s \in S} \phi_4(M(s), R(b_i)) &= c \cdot \exp\left(\frac{(\sum_{s \in S} M(s))^2}{2\sigma^2 \cdot |S|} - \frac{\sum_{s \in S} M(s)^2}{2\sigma^2}\right) \\ &= c \cdot \prod_{1 \leq i < j \leq |S|} \exp\left(-\frac{(M(s_i) - M(s_j))^2}{2\sigma^2 \cdot |S|}\right) = c' \cdot \prod_{1 \leq i < j \leq |S|} \phi'_5(M(s_i), M(s_j)) \end{aligned}$$

Definition 1 (Connected Relational Normal) *The product of RNs is connected, when the connectivity graph of RNs is a connected component. Each vertex of the connectivity graph is a random variable or a constant in RNs, and each edge is a potential (RN). ■*

Lemma 1 *The product of RNs is a probability density function when it is connected, and at least a RN includes a constant argument.*

Lemma 1 can be proved by that the product of connected RNs integrates to a constant given a constant argument. However, we omit the proof for lack of space.

4.2 Constant Time Relational Atom Eliminations

We provide two constant time elimination algorithms for RNs involving a single relational potential (i.e. the product of RNs over different instances of relational atoms). The algorithms eliminate variables, while maintaining the product of RNs.

Elimination of a relational atom X from $\phi_{RN}(X, Y)$ The first problem is to marginalize a relational atom (X) in the product of RNs with two relational atoms (X, Y):

²Note, $\int_{R(b_i)} \exp(-aR(b_i)^2 + 2bR(b_i) - c) = \sqrt{\frac{\pi}{a}} \exp(\frac{b^2}{a} - c)$.

$\phi_{RN}(X, Y)$. The potential is the product of $|X| \cdot |Y|$ RNs. Note that each random variable in X has a relation with each variable in Y .

Algorithm ‘Pairwise Constant₁’ It marginalizes x_i in X , and converts the potential into a pairwise form.

$$\int_{x_i} \prod_{y \in Y} \exp\left(-\frac{(x_i - y)^2}{2\sigma^2}\right) = \prod_{y_i, y_j \in Y, i < j \leq |Y|} \exp\left(-\frac{(y_i - y_j)^2}{2\sigma^2 \cdot |Y|}\right) \quad (2)$$

Note that the marginal over $x_i \in X$ and the marginal over $x_j \in X$ ($i \neq j$) are same. Thus, the following result is derived when it marginalizes all variables in X .

$$\begin{aligned} \int_{x_1} \dots \int_{x_{|X|}} \prod_{x_i \in X} \prod_{y \in Y} \exp\left(-\frac{(x_i - y)^2}{2\sigma^2}\right) &= \prod_{x_i \in X} \left(\int_{x_i} \prod_{y \in Y} \exp\left(-\frac{(x_i - y)^2}{2\sigma^2}\right) \right) \\ &= \left(\prod_{y_i, y_j \in Y, i < j \leq |Y|} \exp\left(-\frac{(y_i - y_j)^2}{2\sigma^2 \cdot |Y|}\right) \right)^{|X|} = \prod_{y_i, y_j \in Y, i < j \leq |Y|} \exp\left(-\frac{|X|(y_i - y_j)^2}{2\sigma^2 |Y|}\right) \end{aligned} \quad (3)$$

The result of integration is the product of pairwise RNs ($\phi_{RN}(Y, Y)$) with the parameter $\frac{|X|}{2\sigma^2 \cdot |Y|}$.

Theorem 2 For the product of RNs between two relational atoms, ‘Pairwise Constant₁’ eliminates all variables in a relational atom at a constant time.

Proof Eliminating a variables x_i in X takes a constant time as Equation 2. Eliminating other variables in X can be done without iterations as shown in Equation 3. Thus, the computation takes only a constant time. ■

Elimination of n random variables from $\phi_{RN}(X, X)$ The second problem is to marginalize some (n) variables in a relational atom (X) in the product of RNs within the relational atom: $\phi_{RN}(X, X)$. The potential is the product of $\frac{|X| \cdot (|X| - 1)}{2}$ pairwise RNs between two variables in X .

Algorithm ‘Pairwise Constant₂’ It updates the marginal after eliminating a random variable without iterations. When it eliminate x_m , it calculates the parameters of ϕ_{RN}' given ϕ_{RN} as the following equation.

$$\begin{aligned} \int_{x_m} \prod_{1 \leq i < j \leq m} \phi_{RN}(x_i, x_j) &= \prod_{1 \leq i < j \leq m-1} \phi_{RN}(x_i, x_j) \cdot \int_{x_m} \prod_{1 \leq i \leq m-1} \exp\left(-\frac{(x_i - x_m)^2}{2\sigma^2}\right) \\ &= \prod_{1 \leq i < j \leq m-1} \phi_{RN}(x_i, x_j) \cdot \prod_{1 \leq i < j \leq m-1} \exp\left(-\frac{(x_i - x_j)^2}{2\sigma^2 \cdot (m-1)}\right) \\ &= \prod_{1 \leq i < j \leq m-1} \phi_{RN}(x_i, x_j) \cdot \prod_{1 \leq i < j \leq m-1} \phi_{RN}'(x_i, x_j) = \prod_{1 \leq i < j \leq m-1} \phi_{RN}''(x_i, x_j) \end{aligned}$$

The coefficient of ϕ_{RN}'' is the sum of coefficient of ϕ_{RN} , σ^2 , and the coefficient of ϕ_{RN}' , $\sigma^2(m-1)$. The sum of two coefficients results in $\sigma^2 \cdot \frac{m-1}{m}$. Similarly, eliminating the next random variable α_{m-1} results in $\sigma^2 \frac{m-2}{m}$ ($=\sigma^2 \frac{m-1}{m} \frac{m-2}{m-1}$). Thus, eliminating n random variables results in $\sigma^2 \frac{m-n}{m}$ without iterations.

Theorem 3 For the product of RNs within a relational atom, ‘Pairwise Constant₂’ eliminates n variables in the relational atom at a constant time.

Proof Updating the parameter of $\phi_{RN}(X, X)$ from σ^2 to $\sigma^2 \frac{m-n}{m}$ takes only a constant time. ■

4.3 A Linear Time Relational Atom Elimination

This section provides a linear time variable elimination algorithm $O(|U|)$ which can be applied to any product of RNs when the constant time algorithms of the previous sections are not applicable.

Elimination of relational atoms from $\prod \phi_{RN}(X_i, X_j)$ This problem is to marginalize some variables in U , ($U = \{X_1, X_2, \dots, X_{|N|}\}$) in the product of RNs between two relational atoms: $\prod \phi_{RN}(X_i, X_j)$. If all relational atoms are related each other, there are $\frac{|N| \cdot (|N| - 1)}{2}$ pairwise RNs.

Lemma 4 For $|U|$ variables in $|N|$ relational atoms ($U = \{X_1, X_2, \dots, X_{|N|}\}$) and RN potentials, marginalizing n variables in a ground model takes $O(n \cdot |U|^2)$.

Proof Suppose we eliminate a variable $x \in U$. Eliminating a variable x in RN needs updates coefficients of terms ($x_i x_j$) where x_i and x_j have relations with the variable x . When x has relations with all other variables in U , the number of terms is bounded by $O(|U|^2)$. Thus, eliminating n variables takes $O(n \cdot |U|^2)$ because it needs n iterations. ■

Thus, any inference algorithm in a ground model has the order of $O(|U|^3)$ time complexity, when it eliminates all ground variables except a few query variables.

Algorithm ‘Pairwise Linear’ To reduce the time complexity, our lifted algorithm uses following notations which refer multiple variables in an atom X : $X_{[m]} = \sum_{1 \leq i \leq m} x_i$; $X_{[m]}^2 = \sum_{1 \leq i \leq m} x_i^2$; $X_{[m][m]} = \sum_{1 \leq i < j \leq m} x_i \cdot x_j$; and $(X_{[m]})^2 = X_{[m]}^2 + 2X_{[m][m]}$. The notations give the following properties:

$$\begin{aligned} \exp(2X_{[m][m]} - (m-1)X_{[m]}^2) &= \prod_{1 \leq i < j \leq m} \exp(-(x_i - x_j)^2) = \phi_{RN}'(X, X) \\ \exp(2X_{[m]}Y_{[n]} - nX_{[m]}^2 - mY_{[n]}^2) &= \prod_{1 \leq i \leq m, 1 \leq k \leq n} \exp(-(x_i - y_k)^2) = \phi_{RN}''(X, Y) \end{aligned}$$

For a potential over X, Y , and $\{x'\}$, it marginalizes x' :

$$\begin{aligned} \int_{x'} \phi_{RN}(X, x') \cdot \phi_{RN}(Y, x') &= \int_{x'} \exp(-(m+n)x'^2 + 2(X_{[m]} + Y_{[n]})x' - (X_{[m]}^2 + Y_{[n]}^2)) \\ &= c \cdot \exp\left(\frac{2X_{[m][m]} + 2X_{[m]}Y_{[n]} + 2Y_{[n][n]} - (m+n)(X_{[m]}^2 + Y_{[n]}^2)}{m+n}\right) \\ &= c \cdot \phi_{RN}'(X, X) \cdot \phi_{RN}''(X, Y) \cdot \phi_{RN}'''(Y, Y) \end{aligned} \quad (4)$$

It iterates until all n variables are eliminated.

Theorem 5 For $|U|$ variables in $|N|$ relational atoms ($U = \{X_1, X_2, \dots, X_{|N|}\}$) and potentials in RN, ‘Pairwise Linear’ eliminates n variables in $O(n \cdot |N|^2)$.

Proof WLOG, we marginalize a variable $x' \in X_1$. We make an artificial atom Y which includes all relational atoms those have relationships with X_1 . Then, $\{x'\}$ is separated from X_1 ($X'_1 = X_1 \setminus \{x'\}$). When marginalizing $\phi_{RN}(X'_1, x') \cdot \phi_{RN}(Y, x')$ over x' , the marginal is also the product of RNs as Equation 4: $\phi_{RN}'(X'_1, X'_1) \cdot \phi_{RN}''(X'_1, Y) \cdot \phi_{RN}'''(Y, Y)$.

The marginal can be represented without the artificial atom Y . We remove Y from $\phi_{RN}''(X', Y)$ and $\phi_{RN}'''(Y, Y)$. $\phi_{RN}''(X'_1, Y)$ is represented as the product of RNs between atoms in Y and X'_1 : $\prod_{X_i \in Y} \phi_{RN}''(X'_1, X_i)$. $\phi_{RN}'''(Y, Y)$ is also

represented as the product of RNs between atoms in B : $\prod_{X_i, X_j \in Y} \phi''_{RN}(X_i, X_j)$.

For each elimination, it updates parameters of all pairs $O(|N|^2)$ among $|N|$ atoms. Thus, computational complexity to eliminate n variables is the order of $O(n \cdot |N|^2)$. ■

Thus, ‘Pairwise Linear’ has the order of linear $O(|U|)$ time complexity with respect the number of ground variables.

5 Exact Lifted Inference with RCM

This section presents our algorithm, *ELIMINATE-CONTINUOUS*, which generates a new parfactor after eliminating a set of relational atoms given a set of parfactors. A potential of each parfactor is the product of *Relational Pairwise Potentials* (RPPs):

$$\phi_{RP}(X, Y) = \prod_{x \in X, y \in Y} \phi_{RP}(x, y)$$

A *relational pairwise model* is a RCM whose potentials are RPP. Here, RPPs are not limited to the RNs in Section 4.1.

Conditions for Exact Lifted Inference The lifted *ELIMINATE CONTINUOUS* algorithm provides the exact solution for potentials of parfactors when the potentials satisfy three conditions: (I) analytically integrable; (II) closed under product operation; and (III) represented as the product of *relational pairwise potentials* after marginalizations. RNs are examples which satisfy the conditions.

5.1 Inversion-Elimination

Inversion elimination is applicable when the set of objects in g is same with the set of objects in e , $LV(e) = LV(g)$. Let $\theta_1, \dots, \theta_n$ be enumeration of Θ_g .

$$\begin{aligned} \int_{RV(e)} \phi(g) &= \int_{RV(e)} \prod_{\theta \in \Theta_g} \phi_g(A_g \theta) = \int_{e[\theta_1]} \dots \int_{e[\theta_n]} \phi_g(A_g \theta_1) \dots \phi_g(A_g \theta_n) \\ &= \prod_{\theta \in \Theta_g} \int_{e[\theta]} \phi_g(A_g \theta) (\because \text{split}) = \prod_{\theta \in \Theta_g} \int_{e[\theta]} \phi_g(A_g \theta) \\ &= \prod_{\theta \in \Theta_g} \int_e \phi_g(A' \theta, e) = \prod_{\theta \in \Theta_g} \phi'(A' \theta) = \phi(g') \end{aligned}$$

Return to the financial market example, inversion elimination can eliminate $G[S, B]$. Before elimination, we get a parfactor $g = (\{S, B\}, \top, (M[S], G[S, B], R[B]), \phi_2 \cdot \phi_3)$ which combines the two parfactors, $(\{S, B\}, \top, (M[S], G[S, B]), \phi_2)$ and $(\{S, B\}, \top, (G[S, B], R[B]), \phi_3)$. Then, the elimination procedure is as follows.

$$\begin{aligned} \int_{RV(G)} \phi(g) &= \int_{RV(G)} \prod_{s \in S, b \in B} \phi_g(M(s), G(s, b), R(b)) \\ &= \prod_{s \in \{auto, \dots, stock\}, b \in \{b_1, \dots, b_m\}} \left(\int_{G(s, b)} \phi_g(M(s), G(s, b), R(b)) \right) \\ &= \prod_{s \in \{auto, \dots, stock\}, b \in \{b_1, \dots, b_m\}} \phi_{new}(M(s), R(b)) = \phi_{new}(M[S], R[B]) = \phi(g') \end{aligned}$$

Note that, the number of substitutions ($|\Theta_g|$) is the number of market sectors ($|S|$) times the number of banks ($|B|$). Regardless the number of substitutions, we can apply the same integration to eliminate $|\Theta_g|$ number of random variables ($G(s, b)$). Thus, it calculates the integral ($= \int_L \phi_g(M(s), G(s, b), R(b))$) once regardless of specific s and b . The marginal becomes the potential ($\phi_{new}(M[S], R[B])$) of the output parfactor (g').

5.2 Relational-Atom-Elimination

Relational-Atom-Elimination marginalizes atoms when *Inversion-Elimination* is not applicable. It is a generalized algorithm of those for RN shown in Section 4. It marginalizes each relational atom of a parfactor g according to three cases: (1) variables in the atom e has no direct relationship each other (i.e. ‘ $\phi(X, Y)\phi(X, Z)$ ’); (2) variables in the atom e has relationships only each other (i.e. ‘ $\phi(X, X)$ ’); and (3) other cases (i.e. ‘ $\prod \phi(X_i, X_j)$ ’).

For the case (1), a generalized ‘Pairwise Constant₁’ eliminates an atom e . In this case, marginalizing a random variable in the atom does not affect marginalizing another variable in the atom as shown in Section 4.2. That is, $\int_{RV(e)} \prod_{\theta \in \Theta_g} \phi_g(\cdot) = \prod_{\theta_e \in \Theta_e} \int_{e(\theta_e)} \prod_{\theta \in \Theta_{g \setminus |e|}} \phi_g(\cdot)$. Here, E is the set of atoms in g , and $\bar{E} = E \setminus \{e\}$.

$$\begin{aligned} \int_{RV(e)} \phi(g) &= \int_{RV(e)} \prod_{\theta \in \Theta_g} \phi_g(A_g \theta) = \int_{RV(e)} \prod_{\theta_e \in \Theta_e} \prod_{\theta \in \Theta_{g \setminus |e|}} \phi_g(A_g \theta_e, A_g \theta) \\ &= \prod_{\theta_e \in \Theta_e} \int_{e[\theta_e]} \prod_{\theta \in \Theta_{g \setminus |e|}} \phi_g(A_g \theta_e, A_g \theta) = \prod_{\theta_e \in \Theta_e} \phi'(RV(\bar{E})) (\because \text{condition(I)}) \\ &= \phi'(RV(\bar{E}))^{RV(e)} = \phi''(RV(\bar{E})) (\because \text{condition(II)}) \end{aligned}$$

The marginal $\phi''(RV(\bar{E}))$ is not a *relational pairwise potential* anymore, because all random variables in \bar{E} are arguments of the potential. When condition (III) is satisfied, the marginal can be converted into the product of *relational pairwise potentials*: $\phi''(RV(\bar{E})) = \prod_{X_i, X_j \in RV(\bar{E})} \phi_{i,j}(X_i, X_j)$.

In the econometric example, it eliminates $R[B]$ as follows.

$$\begin{aligned} \int_{RV(R)} \phi(g') &= \int_{RV(R)} \prod_{s \in S, b \in B} \phi_{new}(M(s), R(b)) \\ &= \prod_{b \in B} \int_{R(b)} \prod_{s \in S} \phi_{new}(M(s), R(b)) = \prod_{b \in B} \phi'_{new}(M(auto), \dots, M(stock)) \\ &= \phi'_{new}(M(auto), \dots, M(stock))^{RV(R)} = \phi''_{new}(M(auto), \dots, M(stock)) \end{aligned}$$

Beyond Relational Gaussian defined in Section 4.1, any potential function satisfying the third condition can convert the potential ϕ''_{new} into the pairwise form $\prod \phi'''_{new}$.

$$\phi''_{new}(M(auto), \dots, M(stock)) = \prod_{s_1, s_2 \in S} \phi'''_{new}(M(s_1), M(s_2))$$

For the cases (2) and (3), generalized algorithms of ‘Pairwise Constant₂’ and ‘Pairwise Linear’ are also applied, respectively.

6 Experiments

We report experiments for the recession model provided in the paper. For experiments, we implemented three algorithms: (A) inference with a grounded model; (B) inference with only Inversion-Elimination; and (C) inference with both Inversion-Elimination and Relational-Atom-Elimination. Our new algorithm (C) is significantly faster than the grounded model (A) and Inversion-Elimination (B). Note that Inversion-Elimination (B) is also our new algorithm for continuous variables, even though comparable elimination methods for discrete variables (de Salvo Braz, Amir, and Roth 2005; Milch et al. 2008; Pfeffer et al. 1999) existed prior to ours. Our experimental results are shown in Figures 5 and 6.

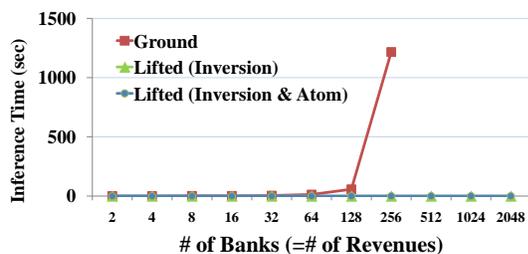


Figure 5: Inference time with different numbers of banks

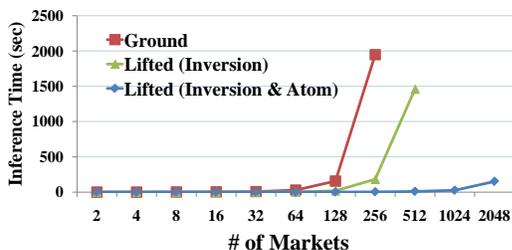


Figure 6: Inference time with different numbers of markets

In the recession model, we provided observations for one market variable (M) and one revenue variable (R). Those variables were split from relational atoms. Then, we calculated the marginal density of the Recession variable. We increased the number of markets and the number of banks from 2 to 2048 exponentially. We set an hour of cut-off time. With 512 banks, the grounded inference (A) did not complete within an hour. The Inversion Elimination (B) and our new algorithm (C) finished computations in almost a constant time for 2048 banks. With 512 markets, (A) could not finish within an hour, again. With 1024 markets, (B) did not finish in an hour. Our new algorithm (C) finished in a reasonable time (about 151 secs) even with 2048 markets.

7 Related Works

(Poole 2003) solves inference problems with the unification which dynamically splits a set of ground nodes and unifies them. With a counting formula, (de Salvo Braz, Amir, and Roth 2005; 2006) provide tractable algorithms. (Milch et al. 2008) devises an improved algorithm using the counting formula to represent conditional density tables compactly. However, such lifted inference algorithms for discrete variables are not applicable to continuous variables.

Markov Logic Network (MLNs) (Richardson and Domingos 2006) use First-order logic sentences to represent relationships over nodes in a graphical model. In this regard, MLNs also represent graphical models at the relational level. (Singla and Domingos 2008) provides an approximated lifted inference algorithm over discrete domain. (Domingos and Singla 2007) makes an analysis for infinitely many discrete variables. However, these achievements are not for continuous domains, too. Thus, they are comparable to lifted inferences (de Salvo Braz, Amir, and Roth 2005;

Milch et al. 2008; Pfeffer et al. 1999) over discrete domains.

Inference with Gaussian distributions is a traditional problem (Roweis and Ghahramani 1999). In detail, calculating conditional densities of multivariate Gaussians requires matrix inversions (Kotz, Balakrishnan, and Johnson 2000) which are intractable for high dimensions. (Paskin 2003) shows that efficient inference is possible for a linear Gaussian when the treewidth of the model is small. For models with large treewidth, however, those inference algorithms over ground models are not applicable in practice.

8 Conclusion

In this paper, we propose a new exact lifted inference algorithm for Relational Continuous Models (RCMs). This algorithm is an advancement of exact inference in RCMs, since all previous works are restricted to discrete domains. Given a query and observations, our algorithm computes the conditional density of the query efficiently.

9 Acknowledgment

This material is supported by NSF CAREER 05-46663 and UIUC/NCSA AESIS 251024 grants.

References

- de Salvo Braz, R.; Amir, E.; and Roth, D. 2005. Lifted first-order probabilistic inference. In *IJCAI*.
- de Salvo Braz, R.; Amir, E.; and Roth, D. 2006. Mpe and partial inversion in lifted probabilistic variable elimination. In *AAAI*.
- Domingos, P., and Singla, P. 2007. Markov logic in infinite domains. In *Probabilistic, Logical and Relational Learning - A Further Synthesis*.
- Friedman, N.; Getoor, L.; Koller, D.; and Pfeffer, A. 1999. Learning probabilistic relational models. In *IJCAI*.
- Getoor, L., and Taskar, B. 2007. *Introduction to Statistical Relational Learning (Adaptive Computation and Machine Learning)*. The MIT Press.
- Koller, D., and Pfeffer, A. 1997. Object-oriented Bayesian networks. In *UAI*.
- Kotz, S.; Balakrishnan, N.; and Johnson, N. 2000. *Continuous Multivariate Distributions*. New York: Wiley.
- Milch, B., and Russell, S. J. 2006. First-order probabilistic languages: Into the unknown. In *ILP*.
- Milch, B.; Zettlemoyer, L. S.; Kersting, K.; Haimes, M.; and Kaelbling, L. P. 2008. Lifted probabilistic inference with counting formulas. In *AAAI*.
- Ng, R., and Subrahmanian, V. S. 1992. Probabilistic logic programming. *Inf. Comput.* 101(2).
- Paskin, M. A. 2003. Thin junction tree filters for simultaneous localization and mapping. In *IJCAI*.
- Pfeffer, A.; Koller, D.; Milch, B.; and Takusagawa, K. T. 1999. Spook: A system for probabilistic object-oriented knowledge representation. In *UAI*.
- Poole, D. 2003. First-order probabilistic inference. In *IJCAI*.
- Richardson, M., and Domingos, P. 2006. Markov logic networks. *Machine Learning* 62(1-2).
- Roweis, S., and Ghahramani, Z. 1999. A unifying review of linear gaussian models. *Neural Comput.* 11(2).
- Singla, P., and Domingos, P. 2008. Lifted first-order belief propagation. In *AAAI*.