Quantum Computing for Partition Function Estimation of a Markov Random Field in a Radar Anomaly Detection Problem

Timothé PRESLES

Cyrille ENDERLI

Gilles BUREL

El-Houssaïn BAGHIOUS

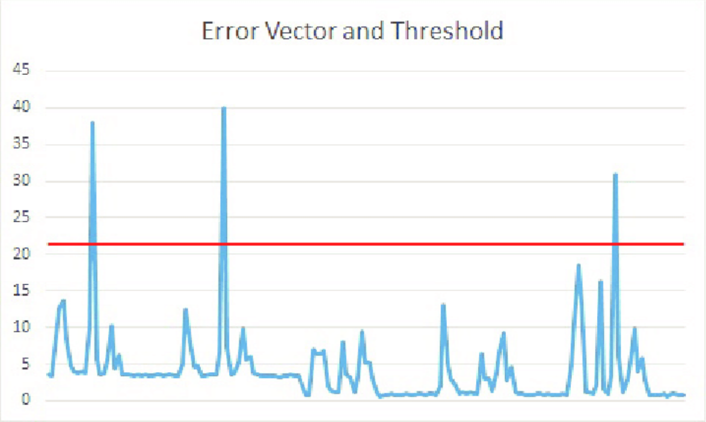
THALES

Building a future we can all trust

2. Problem presentation

> Airborne radar built-in test devices

- **Data collection**
  - Radar systems are equipped with **built-in test devices** that collect functionning data
  - Modern radar system functionning is characterized by up to **100.000 binary, discrete and continuous variables**
  - Once the radar is embedded, **functionning data <u>cannot be efficiently processed</u> to detect anomalies**



> Anomaly detection at the end of the production chain

- Enables to detect radar teething problems **<u>before its embedding</u>**
  - **Advanced analysis of the functionning data** collected by the built-in test devices
- **Probabilistic approach for anomaly detection in production chain**
  - Modelization of the functioning of the radar by a **mixed graphical model**
  - **<u>Learning of a model</u> corresponding to the good functionning of the system**
  - Enables the localization of the components source of anomalies by **<u>computing the likelihood of new acquisition files</u> for the learned model**

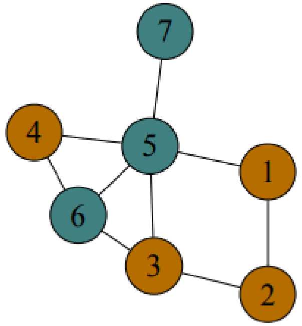


Figure 2.2: Structure of a mixed graphical model. The model has four binary variables x_1, x_2, x_3 and x_4, represented by the brawn nodes with numbers 1 to 4, and three quantitative variables x_5, x_6 and x_7, represented by the grey nodes with number 5, 6 and 7.

3. Problem formulation

- **Gibbs distribution associated to the MRF:**

$$p_\Omega(x_C, x_Q) = \frac{g_\Omega(x_C, x_Q)}{Z_\Omega}$$

$$with\ \ g_\Omega(x_C, x_Q) = \exp\left(x_C^T \Theta x_C + \mu^T x_Q - \frac{1}{2} x_Q^T \Delta x_Q + x_c^T \Phi x_Q\right)$$

$$and\ Z_\Omega = \sum_{x_C \in \{0,1\}^N} \exp(x_C^T \Theta x_C) \int_{\mathbb{R}^M} \exp(\mu^T x_Q - \frac{1}{2} x_Q^T \Delta x_Q + x_c^T \Phi x_Q)\, dx_Q$$

with $x_C \in \{0,1\}^N$ the categorical variables (binary and discrete binarized)
with $x_Q \in \mathbb{R}^M$ the quantitative variables (continuous)
with $\Omega = \{\Theta, \mu, \Delta, \Phi\}$ the model parameters

- **Objective :**
  - **Learning the values of $\boldsymbol{\Omega = \{\Theta, \mu, \Delta, \Phi\}}$** corresponding to a good functionning radar through a process of gradient descent

- **Limitation encountered in previous works:**
  - Each learning step requires to **update the log-likelihood gradient**: $\frac{\partial \ln(p_\Omega)}{\partial \Omega} = \frac{\partial \ln(g_\Omega)}{\partial \Omega} - \frac{\partial \ln(Z_\Omega)}{\partial \Omega}$
  - $\frac{\partial \ln(Z_\Omega)}{\partial \Omega} = \frac{1}{Z_\Omega} \frac{\partial Z_\Omega}{\partial \Omega}$ non-trivial *(requires to compute $2^N + (2^{64})^M$ values for each parameters of Ω)*
  - **Approximations** employed in previous work *(stochastic gradient approximation, etc…)* **considerably reduce the accuracy of the model learning**.

Idea

Take advantage of quantum computing to improve and speed-up the computation of $\frac{\partial \ln(Z_\Omega)}{\partial \Omega}$

3. Problem formulation

- Considering that **quantitative variables x_Q follow a Gaussian distribution**, straightforward calculations lead to :

$$Z_\Omega = (2\pi)^{\frac{M}{2}} |\Sigma|^{\frac{1}{2}} \exp(\frac{1}{2}\mu^T \Sigma \mu) \sum_{x \in \{0,1\}^N} \exp(x_C^T \Psi x_C)$$

$$with \;\; \Psi = \Theta + \frac{1}{2}\Phi\Sigma\Phi^T + diag(\Phi\Sigma\mu) \; symetric \; and \; \Sigma = \Delta^{-1}$$

- **Objective :**
  - Develop a quantum approach to compute $\sum_{x \in \{0,1\}^N} \exp(x_C^T A x_C)$ for any matrix A
  - **Can be used to speed up the computation of each $\frac{\partial Z_\Omega}{\partial \Omega_i}$ for $\Omega_i \in \Omega = \{\Theta, \mu, \Delta, \Phi\}$**

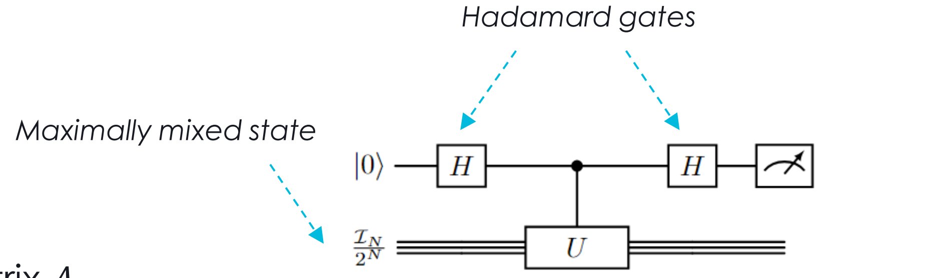


Fig. 1. Circuit for estimating $\mathrm{Re}(\mathrm{Tr}(U))/2^N$ in the one clean qubit model.

- **One-clean qubit model**
  - **Quantum algorithm designed to estimate the trace of a unitary operator**
  - For U the $2^N \times 2^N$ matrix associated to a unitary operator *(a quantum gate)* acting on N qubits, $\mathbf{Re}(\mathbf{Tr}(U))$ **can be deduced from the probability of measuring 0 on an ancillary qubit**

$$\mathrm{Re}(\mathrm{Tr}(U)) = \left(p_0 - \frac{1}{2}\right) 2^{N+1}$$

Idea

Define a unitary U such that $\mathrm{Re}(Tr(U)) = \sum_{x \in \{0,1\}^N} \exp(x_C^T A x_C)$

3. Quantum approach for partition function estimation

> Linear Combination of Unitaries (LCU) method:

- **Let's consider a Gibbs distribution** $p(x) = \exp(g(x))/Z$ with $x_C \in \{0,1\}^N$ and $g(x_C) = \exp(x_C^T A x_C)$
  - H is defined from the coefficients $A_{i,j}$ of $g(x)$ such that

$$H = \frac{-1}{\beta} \sum_{i,j=1,1}^{N} A_{i,j} U_{(i,j)} \ \ with \ \ \beta = \sum_{i,j=1,1}^{N} |A_{i,j}|$$

  - with $U_{(i,j)}$ diagonal unitary matrices of size $2^N \times 2^N$ such that:

$$U_{(i,j)} = I^{(1)} \otimes \cdots \otimes B^{(i)} \otimes \cdots \otimes B^{(j)} \otimes \cdots \otimes I^{(N)}$$

  - with $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \ I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B^{(i)}$ corresponding to the **application of the operator B to qubit i**

- **With this construction,** we have $\mathrm{Tr}\left(e^{-\beta H}\right) = \sum_{x\in\{0,1\}^N} \exp\left(x_C^T A x_C\right) = \sum_{x\in\{0,1\}^N} g(x_C) = Z$

  Intuition : Compute straightforwardly $Tr(e^{-\beta H})$ with the one-clean qubit model

  Problem : It is not straightforward to define $\mathrm{U} = e^{-\beta H}$

- **Chebyshev approximation of $e^{-\beta H}$:** $S_K = I_0(\beta) I_N e^{\beta} + 2\sum_{k=1}^{K} (-1)^k I_k(\beta) T_k(H)$ *with $I_k(\beta)$ the modified Bessel function of the 1^{st} kind*
with $T_k(H)$ the k^{th} Chebyshev polynomial of the 1^{st} order

3. Quantum approach for partition function estimation

$$S_K = I_0(\beta) I_N e^{\beta} + 2\sum_{k=1}^{K} (-1)^k I_k(\beta) T_k(H)$$

with $I_k(\beta)$ the modified Bessel function of the 1^{st} kind
with $T_k(H)$ the k^{th} Chebyshev polynomial of the 1^{st} order

> Estimation of $\mathrm{Tr}(S_K)$

- **We have $\boldsymbol{Tr}(\boldsymbol{S_K}) = I_0(\beta)2^n e^{\beta} + 2\sum_{k=1}^{K}(-1)^k I_k(\beta)\boldsymbol{Tr}(\boldsymbol{T_k}(\boldsymbol{H}))$** → **There exists a general method for defining $T_k(H)$ as a unitary operator !**
  - **Goal :** Compute $Tr(T_k(H))$ with the one-clean qubit model
  - Requires to define a unitary operator encoding $T_k(H)$
- **Quantum walk operator $\boldsymbol{W_H}$:**

Gibbs distribution parameterized by A
N variables problem

→ Definition of the corresponding matrix H associated to a Hamiltonian
$H = -\frac{1}{\beta}\sum_{i,j=1,1}^{N} A_{i,j} U_{(i,j)}$
(Linear combination of unitary operators, which is generally not unitary)
Encoded on N qubits with $U_{(i,j)}$ of size $2^n \times 2^n$

→ Block encoding of H
$U_H = \begin{pmatrix} H & * \\ * & * \end{pmatrix}$
(Definition of a unitary operator encoding H with the use of ancillary qubits)
Requires $M = \log_2(L)$ ancillary qubits with L the number of non-zero terms of A

→ Quantum walk operator W_H defined from U_H such that :
$(W_H)^k \Leftrightarrow T_k(H)$
Unitary operator acting on $N + M + 1$ qubits

- **Computation of an estimation $\boldsymbol{\phi_k}$ of $\boldsymbol{Tr}((\boldsymbol{W_H})^{\boldsymbol{k}})$** with the one clean qubit model

$$Z = \sum_{x\in\{0,1\}^N} \exp\left(x_C^T A x_C\right) = \mathrm{Tr}\left(e^{-\beta H}\right) \approx \mathrm{Tr}(S_K) = I_0(\beta)2^N e^{\beta} + 2\sum_{k=1}^{K}(-1)^k I_k(\beta) Tr(T_k(H)) \approx I_0(\beta)2^N e^{\beta} + 2\sum_{k=1}^{K}(-1)^k I_k(\beta)\phi_k$$

THALES
Building a future we can all trust

3. Quantum approach for partition function estimation

> Precision metrics

- **Accuracy of the estimation depending on two parameters :**

$$Z = \mathrm{Tr}(e^{-\beta H}) \approx \mathrm{Tr}(S_K) = I_0(\beta)2^N e^{\beta} + 2\sum_{k=1}^{K}(-1)^k I_k(\beta) Tr(T_k(H)) \approx I_0(\beta)2^N e^{\beta} + 2\sum_{k=1}^{K}(-1)^k I_k(\beta)\phi_k$$

Precision proportional to $\boldsymbol{K}$ *i.e. the degree of Chebyshev approximation*

Precision proportional to the ***number of samples*** $\boldsymbol{Q}$

$$\mathrm{Re}(\mathrm{Tr}(U)) = \left(p_0 - \frac{1}{2}\right)2^{N+1}$$

> First results

- **Simple graphs,** due to the **limited power of computation** available in current quantum hardwares
- **Results analysis :** *(simulated results)*
  - The number of samples $\boldsymbol{Q}$ **significantly impacts the accuracy of the estimation**
  - As $I_k(\beta)$ decreases exponentially for k increasing, $\boldsymbol{K}$ **can be set at low values without significantly impacting the accuracy of the estimation**
  - The high current **error rate** of quantum hardwares requires to **significantly increases the number of executions** required to obtain a satisfying precision.

Variation of the estimation error for different values of Q					
Nb. var	Q				
N	10^3	10^4	10^5	10^6	10^7
2	48.90%	5.82%	1.49%	0.80%	0.47%
3	68.56%	7.34%	2.48%	1.16%	0.72%
4	97.85%	9.17%	3.66%	1.59%	1.39%

Average precision for $K = 3$ *and* $\delta = 0{,}1$

Variation of the estimation error for different values of K					
Nb. var	K				
N	1	2	3	4	5
2	9.98%	3.41%	1.49%	1.49%	1.49%
3	17.91%	4.64%	2.48%	2.47%	2.47%
4	33.57%	8.16%	3.66%	3.65%	3.65%

Average precision for $Q = 10^5$ *et* $\delta = 0{,}1$

4. Conclusion

> Contributions

- We propose a quantum approach for **speeding up and improving the learning of** a mixed graphical model
  - We introduce a **general method for encoding a Gibbs distribution** as a linear combination of unitary operators H
  - From this linear combination of unitary operators, we define a **quantum walk operator W_H that is evaluated with the one-clean qubit model** algorithm to obtain an **estimation of the partition function of any Gibbs distribution.**
  - Our approach **enables the computation of all the partial derivates** $\frac{\partial\, ln(Z_\Omega)}{\partial \Omega}$ **required to update the gradient** of the log-likelihood required at each step of the model learning.

> Further work:

- Result analysis for the computation of $\frac{\partial\, ln(Z_\Omega)}{\partial \Omega}$
- Study of the **current limitations of state-of-the-art methods**
  - Estimation of the conditions for quantum advantage