Bayesian Searches and Quantum Oscillators

3rd Workshop on Microwave Cavities and Detectors
Lawrence Livermore National Laboratory

G. Chapline and M. Schneider
Lawrence Livermore National Laboratory

Aug 23, 2018

LLNL-PRES-XXXXXX
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Background

- A new LLNL SI is focused on developing Gaussian process inference techniques for Bayesian searches

  - Part of this project involves investigating the idea of using an array of quantum oscillators (*viz* microwave cavities) as an analog computer for implementing Bayesian data analysis

- Practical motivations:

  1) Analyzing sparse/ambiguous data problems
  2) Detection of signals below the threshold for single quantum excitation

Bayesian Pattern Recognition

- Best classification procedure is to choose class such that posterior probability $P(\alpha \mid x)$ for an explanation α is largest, where

$$P(\alpha \mid x) = \frac{p(\alpha)\, p(x \mid \alpha)}{\sum_{\beta} p(\beta)\, p(x \mid \beta)}$$

"Physics" Interpretation for Bayesian Data Analysis

"Energy cost" for explanation α :

$$E_\alpha \equiv -\log[p(\alpha)p(\alpha \mid x)]$$

Bayes' formula for the posteriori probabilities for various explanations

$$P(\alpha \mid x) = \frac{e^{-E_\alpha}}{\sum_\alpha e^{-E_\alpha}}$$

Which minimizes the information theory "free energy"

$$F(x) = \sum_\alpha [E_\alpha P(\alpha) - (-P(\alpha)\log P(\alpha))]$$

Often referred to as the minimum description length

Bayesian networks

In Bayesian networks the world model for input data is represented by conditional probabilities for the activations of the nodes in a layered network given a set of inputs to the first layer:

$$P(d \mid \alpha) = \sum_{\{w_j\}} p(d \mid \{w_j\}, \alpha) p(\{w_j\} \mid \alpha)$$

The conditional probabilities $p(\{w_j\}|\alpha)$ represent "connection strengths" between nodes

Unfortunately the usefulness of Bayesian networks is limited because summing over model parameters requires Monte Carlo sampling which can be impractical

Gaussian Process (GP) Neural Networks

(Google Brain 2017)

What underlies classical neural networks is the Arnold-Kolmogorov theorem: any function of N-variables can be constructed from compositions of a nonlinear function of one variable. When the errors in the model - data residuals are Gaussian, this construction can be carried out analytically; in particular the covariance of the output GP is obtained by inverting a nonlinear transform of the data covariance matrix.

Minimizing the quadratic form that describes the mean errors in the GPNN predictions corresponds to minimizing the ”free energy”.

Remarkably this is related to inverse scattering theory for multi-channel quantum mechanics (Dyson 1975).

Quantum oscillator dynamics in the presence of a noisy signal

- The propagator for the density matrix for a quantum oscillator driven by a classical noise signal *f*(*t*) is

$$J = \iint e^{i\{S[x(t)]-S[x'(t)]\}/\hbar}\,\Phi[x(t)-x'(t)]Dx(t)Dx'(t)$$

where $\Phi[k(t)] = \int e^{i\int k(t)f(t)dt}\, P[f(t)]Df(t)$

is the generating function for signal correlations

- For Gaussian noise

$$P[f(t)] = exp\left\{-\frac{1}{2}\iint (f(t)-\bar{f}(t))A^{-1}(t,t')(f(t')-\bar{f}(t'))dtdt'\right\}$$

where A is the autocorrelation function for the signal *f*(*t*)

Where is the the non-linearity required for a quantum universal calculator?

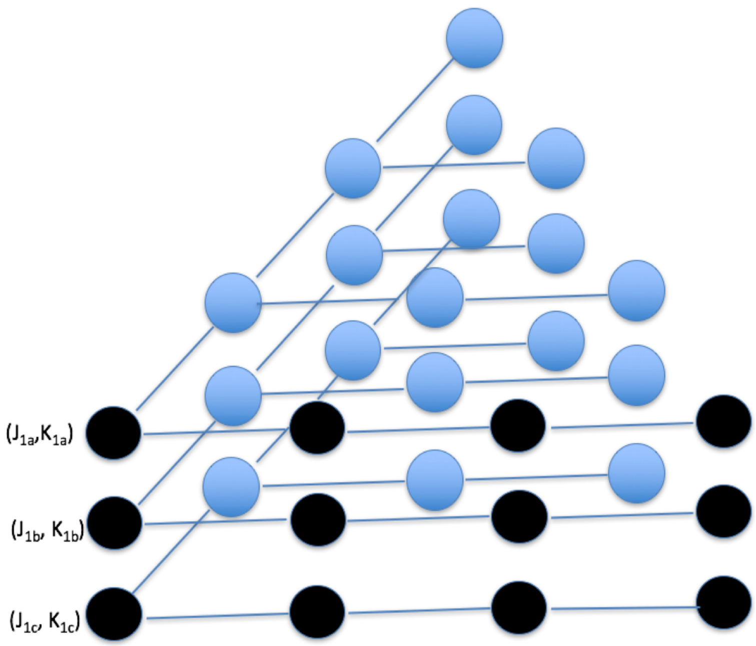
- Pattern recognition using neural networks always requires a nonlinearity in order to map input data to a arbitrary functions in feature space.

- The quantum dynamics of density matrices is non-linear if one allows coupling to a bath (eg TLs) where TL wave functions are reset at specified intervals to the ground state:

$$\rho(x,y,x_0,y_0,t) = \iint \exp\frac{im}{2\hbar}\left(\dot{x}^2 - \omega_0^2 x^2 - \dot{y}^2 + \omega_0^2 y^2\right) F(x,y) DxDy$$

$$F(x,y) \cong \exp - \frac{g^2}{2\Delta}\left(\frac{m\omega_0^2}{2\hbar}\right)\left(\int_0^T \int_0^t \left[\left(e^{-i\Delta(t-t')} x(t') - e^{i\Delta(t-t')} y(t')\right)\left(x(t) - y(t)\right)\right] dt'\, dt\right)$$

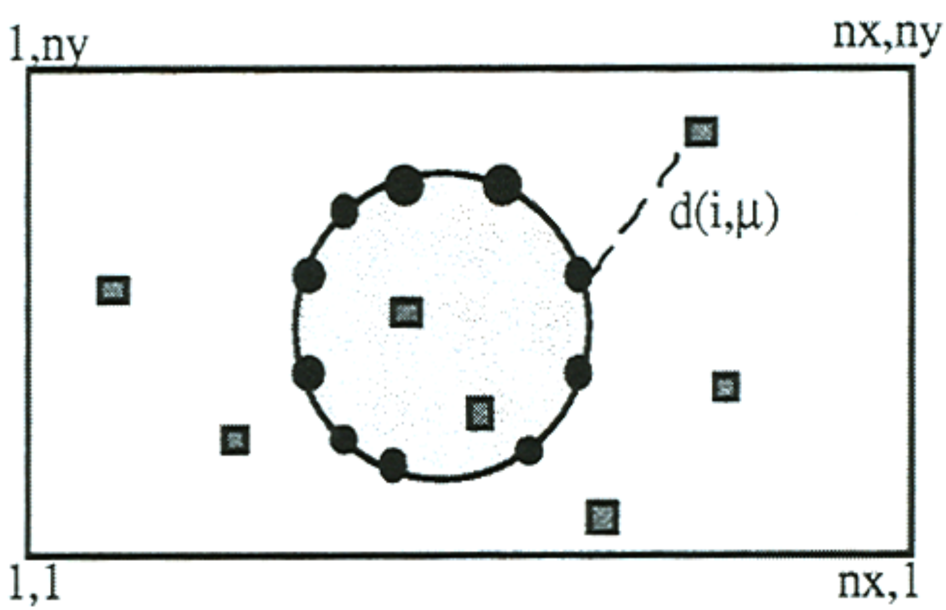
Quantum Oscillator Boltzmann Machine



If the oscillators in a layered array of quantum oscillators are coupled to each other, and also a bath that is periodically reset, the density matrix of the array will relax in such a way to minimize the "free energy"; i.e. the quadratic form that describes the differences in the predicted and observed excitations of the black nodes.

Visualization of an optimum Bayesian search

- The information theoretic free energy optimization of a Bayesian search using a harmonic oscillator array coupled to reset noise can be visualized as minimization of the area swept out by a string; which is reminiscent of the solution of the traveling salesman problem.



$$d(I,\mu) = |\xi_\mu - w_\mu|^2$$

How all this can be simulated – not to mention hardware realization – is a work in progress

- Expanding the end point of the exact path integral density matrix propagator from *t* to *t* + ε yields a non-Markovian master equation:

$$\frac{\partial\rho}{\partial t}=\left(\frac{i\hbar}{2L}\left(\frac{\partial^2\rho}{\partial Q^2}+\frac{\partial^2\rho}{\partial Q'^2}\right)+\frac{iL\omega_0^2}{2\hbar}\left(Q^2-Q'^2\right)-\frac{g^2}{\hbar^2}\right.$$

$$*\exp\left(-\frac{g^2}{\hbar^2}\left(\int_{t_0}^{t}Q(t')\int_{t_0}^{t'}Q(s)e^{-i\Delta(t'-s)/\hbar}\,ds\,dt'+\int_{t_0}^{t}Q'(t')\int_{t_0}^{t'}Q'(s)e^{i\Delta(t'-s)/\hbar}\,ds\,dt'\right)\right)$$

$$\left.*\left(\int_{t_0}^{t}\left(Q(t)Q(s)e^{-i\Delta(t-s)/\hbar}+Q'(t)Q'(s)e^{i\Delta(t-s)/\hbar}\right)ds\right)\right)\rho$$

Summary

- Entangled coherent sates for an array of multi-dimensional quantum oscillators can be used to to encode the statistical properties of input signals and organize input data so as provide the “minimum information free energy” explanation for the data.

Longer term it appears that large arrays of quantum oscillators could be used to detect weak signals with low signal to noise and below the threshold for singe quantum detection

Classical Bayesian Searches

- Classical Bayesian searches: a sequence of choices "locations" x_n to be searched and observations Y_n

$$p_n\left(x^* = x | Y_n, \boldsymbol{U}_n, p_{n-1}\right) = \frac{P(Y_n | U_n, X^* = x, p_{n-1}) p_{n-1}(x^* = x | Y_{n-1}, \boldsymbol{U}_{n-1})}{P(Y_n | U_n)}$$

where $U_n = \Delta x_n$ defines the search strategy.

- Finding the optimal search strategy for many problems of practical interest is a "work in progress"

  - When the sources of noise and measurement errors are Gaussian, then the strategy can be obtained by solving a nonlinear matrix equation; which also solves the inverse scattering problem in quantum mechanics (Dyson 1975)