A Path Integral-based Relationship between Meta-Heuristic Synchronization, Quantum Transitions and Emergent Condensates in Networks

John Campbell

Introduction - Abstract Models and their importance in physics:

Abstract models are of utmost importance to understand aspects of nature from first principles that would otherwise appear impossibly complex by observation alone. There is a clear basis in the theory of reasoning itself, as well as the history of science, for why abstract models are important for our understanding which is sometimes taken for granted.

The original basis of the forces of motion Galileo himself derived by setting up abstract and in many ways unnatural models such as describing perfect, frictionless inclines with a perfect frictionless spheres rolling down on it to describe inertial forces. This was a very strange idea at the time given the fact that most of his contemporaries were studying the natural forces of motion as they really appeared, i.e. rocks falling, leaves or apples falling which are in fact subject to chaos, i.e. small deviations in initial conditions which propagate outwards in irreducibly complex fractal forms.

It was thus impossible for Galileo's contemporaries to deduce physics from first principles and so Galileo's abstractions were eventually accepted as being the only way to deduce physics. Furthermore, Newton showed that, indeed, calculus could only be done on such abstract models and generated predictions on objects operating within those abstract models of nature that contained certain constraints and boundary conditions, not all necessarily imposed by nature itself.

In quantum physics we have a different story but with a similar outcome, not only leading to chaos but often complete collapse of contemporary modelling. For example, in quantum physics we continue with Galileo's modality and set up an abstract theoretical model, the wave function, based on the position and momentum variables of the hydrogen atom say. We can only ever hope to set up an equally abstract circumstance to match prediction with theory – (i.e. by isolating hydrogen atoms in magnetic traps in a uniform spin orientation.)

This is true in this case because the observations made of the subject we want to study, the atom, is so vulnerable to externalities that under the parameters of quantum mechanics the observation itself changes the state that we would want to observe. When we experiment on or even "observe" the atom (i.e. using Heisenberg's gamma ray microscope on a free electron), physicists are influencing it and force the wavefunction of the atom, or any quanta for that matter, to collapse. The only way to describe any further what was going on was to create a yet more abstract model, the path-integral interpretation of quantum mechanics, to by-pass the use of wavefunctions which resulted in a more concise, fundamental model.

The approach of using the path integral formalism from quantum mechanics for meta-heuristic algorithms and systems of entanglement is an abstract one, in a similar sense as the examples above, and is an attempt at finding the most fundamental method to describe self-organising systems of coupling, that exist both in the natural and artificial world.

In the field of insect communication for example, such as with squids, glow-worms, fireflies and bees, scientists study with observations external swarm behaviours which are in fact very complex and causes incredible debate and disagreement. Simulation and study of the apparent metaheuristic nature of coherent quantum systems, which indeed have enormous practical applications, have not however created much headway in developing a mechanism for how the coherent systems behave as they do, which is taking their genuinely fascinating mathematical behavioural properties for granted. Constant empirical observation and modelling based on those observations alone has not revealed the fundamentals how self-organising coupling systems work for the same reasons perhaps that Galileo's peers could not have deduced classical physics from first principles – in that it seems we really only can do modelling on first principles using abstract models.

Power-Law Based Signalling - The Archetypal Signalling Meta-Heuristic Algorithm:

The most basic meta-heuristic signalling algorithm we can construct is one based on some sort of power law for signalling. Light signalling can of course be represented by a power law. Light signal transmission, like light in general, obeys the **inverse-square law** in its propagation through space. That is to say, the **intensity** of the light, **I**, is **inversely proportional** to the **square of the distance** from the light source.

$$I(r) = \frac{I_0}{r^2}$$

Therefore the light signal gets weaker and weaker as the distance increases.

The meta-heuristic algorithm therefore must contain a function which monotonically decreases its signalling power between each transceiver node with **distance**, **r**, under a **discrete light absorption coefficient for the physical medium γ.**

Under a given physical medium, the Gaussian form of the light intensity is then determined by:

$$I(r) = I_0 e^{-\gamma r^2}$$

We can use this equation of course to find the light absorption coefficient itself:

$$\frac{I(r)}{I_0} = e^{-\gamma r^2} => ln(\frac{I(r)}{I_0}) = -\gamma r^2 =>$$

$$\gamma = \frac{-ln(\frac{I(r)}{I_0})}{r^2}$$

The generic **power law signalling function** of is then written to be monotonically decreasing:

$$\beta = \beta_0 e^{-\gamma r^m}$$

Where r is the distance between the different nodes.

m is the power law parameter for the signal, for light signals it is m=2.(for signals that use audio, chemical, etc, power laws it will be different)
Where β_0 is the emitted signal at r=0

The **distance** between any two nodes, **i** and **j** at **Xi** and **Xj** respectively, is the Cartesian distance as follows:

$$r_{ij} = \|X_i - X_j\| = \sqrt{\sum_{k=1}^{d} (x_{i,k} - x_{j,k})^2}$$

Where $x_{i,k}$ is the **(k)**th component of the spatial coordinate X_i of the **(i)**th node.

Where **d** is the number of dimensions.

The time evolution of each individual node, i, over its signalling cycle is the governed by the fact each typical node, i, is coupled to the brightest node it sees, j, by the following equation:

$$X_j = X_i + (\beta_0 e^{-\gamma r^2})(X_i - X_j) + \alpha$$

The second term is the signalling term, for most cases in implementation $\beta_0 = 1$

The third term is the randomization term for a possible discrete coupling path which uses a randomization parameter, **α**, which for most cases in implementation is distributed in the domain of [0,1].

An interesting way to write this equation is:

$$X_j(1 + \beta_0 e^{-\gamma r^2}) = (1 + \beta_0 e^{-\gamma r^2})(X_i) + \alpha$$

Which when tidied up further reads:

$$(X_i - X_j) = \frac{\alpha}{(1+\beta_0 e^{-\gamma r^2})}$$

We can then define the signalling **Activation Function, A:**

$$A(X_{i,j}) = (1 + \beta_0 e^{-\gamma r^2})^{-1}$$

Which reduces our equation to a more fundamental **signal field representation**:

$$\Phi(X_{i,j}) = A(X_{i,j})\cdot\alpha$$

Based on these formulae, we have the corresponding rules:

- Each node in the field will emit and absorb discrete light signals equally.
- The signals are proportional to the intensity of the light emitted, which both decrease in proportion to increasing distance between the nodes.
- For any 2 flashing nodes in the field, the signal absorbed with the highest intensity will induce the strongest coupling
- If there is no signal detected with higher intensity than any one particular node, it will designate itself as being isolated and signal randomly
- The light insanity of any signalling node is determined by the landscape of the field, itself determined by the nature of the activation function itself.

Defining the criteria of the meta-heuristic behaviour:

Heuristic algorithms typically intend to find a good solution to an optimization problem by ‘trial-and-error’ in a reasonable amount of computing time. Here ‘heuristic’ means to ‘find’ or ‘search’ by trials and tallying the hits and misses. There is no guarantee to find the best or optimal solution, though it might be a better or improved solution than an educated guess.

Any reasonably good solution, often suboptimal or near optimal, would be good enough for such problems. Broadly speaking, local search methods are heuristic methods because their parameter search is focused on the local variations, and the optimal or best solution can be well outside this local region. However, a high-quality feasible solution in the local region of interest is usually accepted as a good solution in many optimization problems in practice if time is the major constraint.

Metaheuristic algorithms are higher-level heuristic algorithms. Here, ‘meta-’ means ‘higher-level’ or ‘beyond’, so metaheuristic means literally to find the solution using higher-level techniques, though certain trial-and-error processes are still used. Broadly speaking, metaheuristics are considered as higher-level techniques or strategies which intend to combine lower-level techniques and tactics for exploration and exploitation of the huge space for parameter search.

From the rules derived from the physical parameters used to construct the pre-programmed signalling mechanism we can define that the metaheuristic nature of the signalling algorithm is based the behaviours of exploration and exploitation

Exploration in the metaheuristic algorithm is achieved, in the case of the above signalling algorithm, by the use of randomization which enables the algorithm to have the ability to jump out of any local state of coupling between 2 or more signalling nodes so as to explore the search for more couplings in the network over a potentially larger area.

Randomization is also used for local search around the current state if the signalling are limited to a local region. When the signalling frequency is long, randomization then allows for exploration of the the search space on a larger scale. Fine-tuning the right amount of randomness and balancing local search and global search are crucially important in controlling the performance of the synchronizing metaheuristic algorithm.

Exploitation is the use of local knowledge of the search and solutions found so far so that new search moves can concentrate on the local regions or neighbourhood where the optimality may be close; however, this local optimum may not be the global optimality. Exploitation tends to use strong local information such as gradients, the shape of the mode such as convexity, and the history of the search process.

Observations, using simulations, of the convergence behaviour of common optimisation algorithms suggests that exploitation tends to increase the speed of convergence, while exploration tends to decrease the convergence rate of the algorithm. On the other hand, too much exploration increases the probability of finding the global optimality, while strong exploitation tends to make the algorithm being trapped in a local optimum.

The relationship between exploration and exploitation both can be seen with a simple implementation of the metaheuristic derived above in the context ofsensory exploration and contrastive learning exploitation based on the activation signalling function

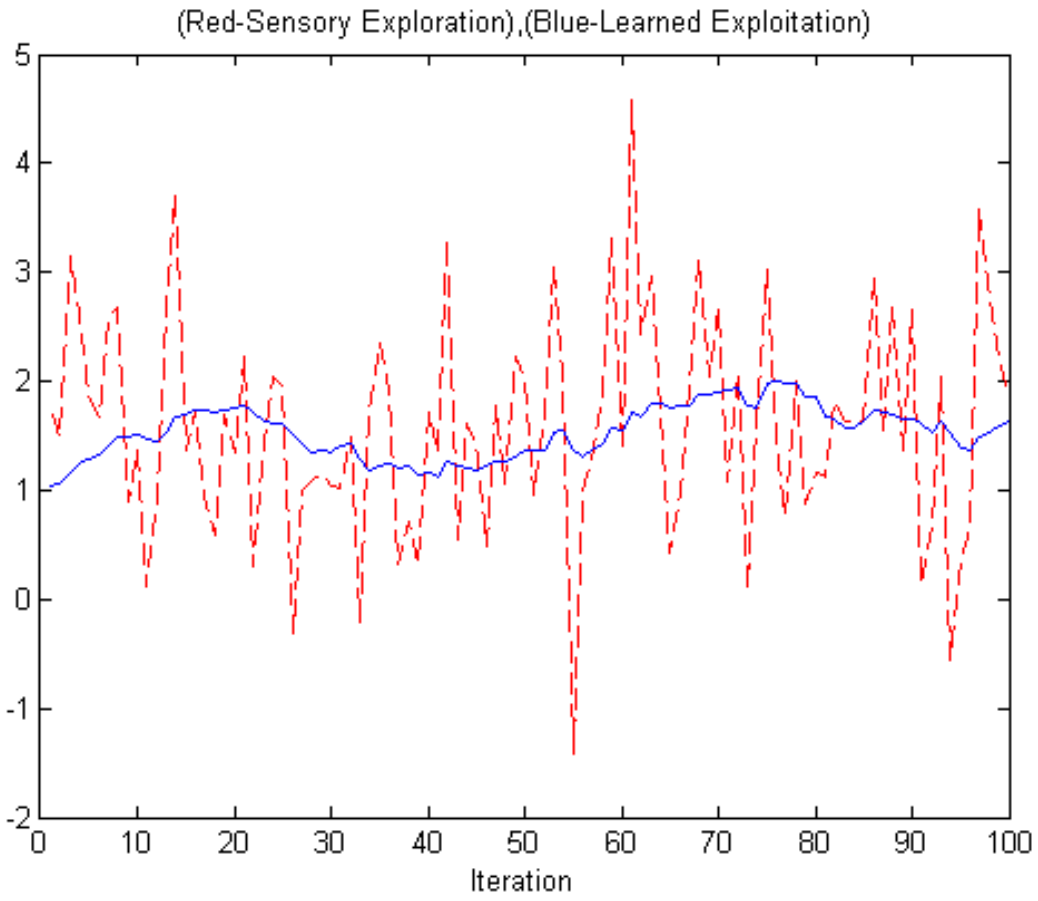
The sample code is implemented as:

```
    for k=1:N  % Generate sensed signals, given that the initial values are
instincts to exploit

        z(k,1)=sigmaz2(k,1)*(0.5-randn); % Generate internal signalling
system based on activation function A(r)

        y(k,1)=x(k,1)+z(k,1); % Generate input – i.e. an external signal that
is sensed - exploration
```

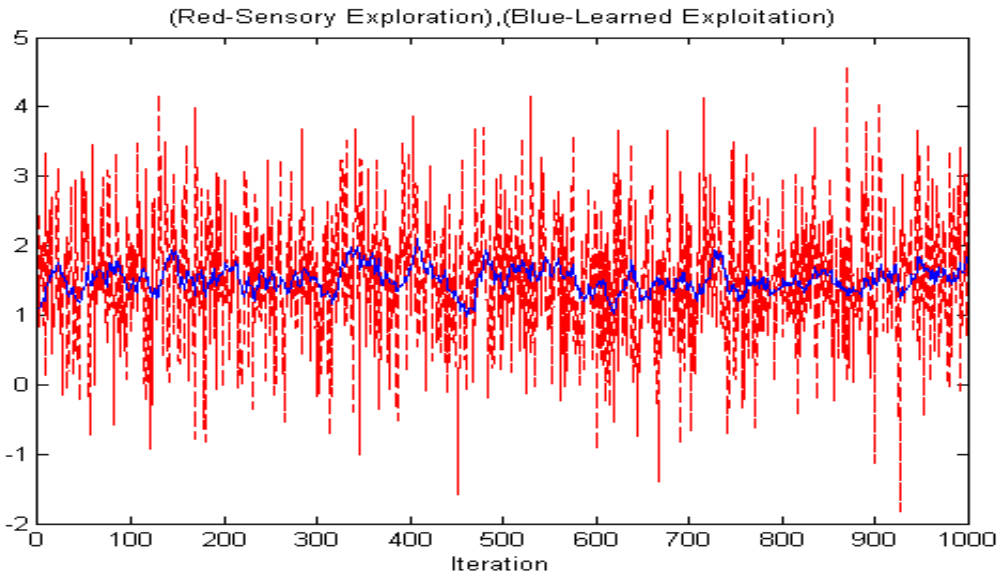
```
        Y=[y(k,1); Y(1:(n-1),1)]; % Learn by contrast - i.e. shift regression
vector and load in new value - exploitation (or endowment)

        xhat(k,1)=mean(Y); %Get the mean of the contrastive learning
procedure
    end
```

We plot the exploration and exploitation convergence plots for 100, 1000 and 10,000 iterations:

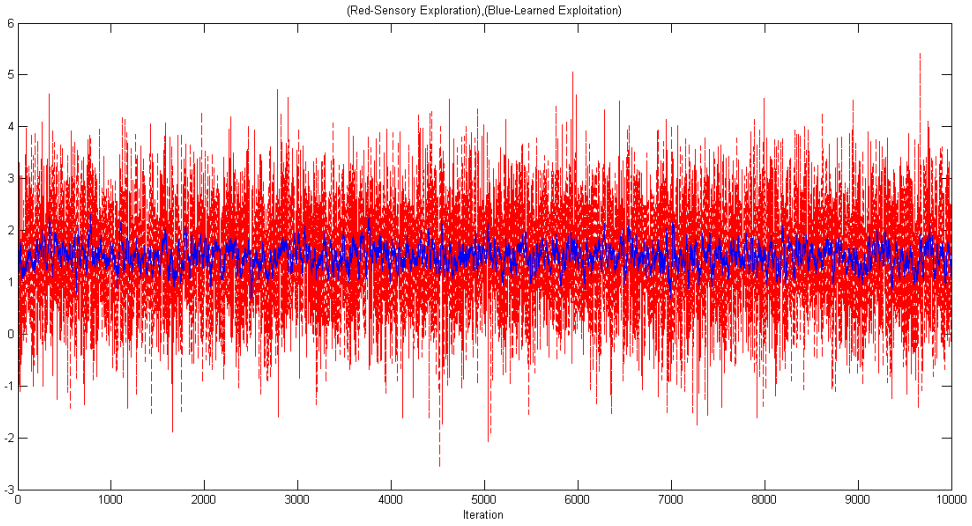
100 iterations:



1,000 iterations:



10,000 iterations:



As seen and discussed, there is a fine balance between the right amount of exploration and the right degree of exploitation. Despite its importance, there has never been any known practical guideline for this balance as regards to learning behaviour.

As we will find out, it is actually not practical at all to find a guideline based on the balances between exploration and exploitation by simply reading convergence plots, which are simply data, which is at least partially recalcitrant as the overall self-organising behaviour of the system that follows the algorithm has not itself been pre-programmed but nor is the system based entirely on probability.

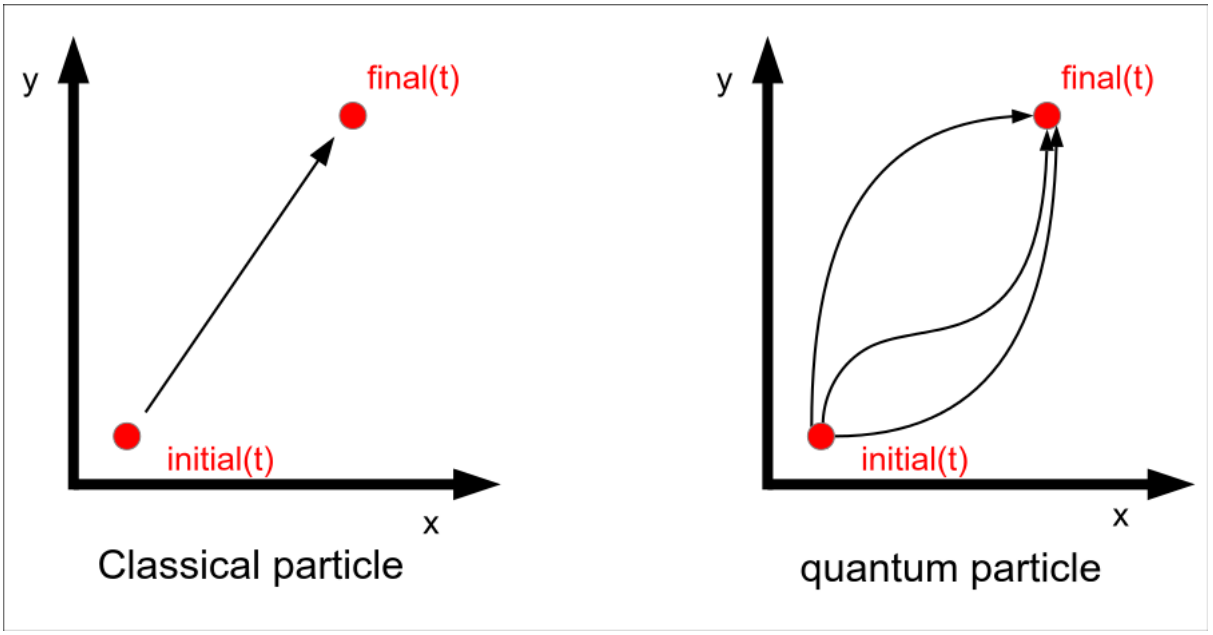
Therefore, it makes more sense to think about the behaviour and data we empirically observe using an abstract model of a physical system that reduces the

behaviour, otherwise taken for granted, to a simple signalling action principle based on the environmental parameters of the system. This then leads to very fascinating consequences and applications of metaheuristics in the realm of quantum information.

Abstraction – The Path Integral Interpretation of Quantum Mechanics:

Path-integral theory is a rigorous formulation of quantum theory that uses a particles path across space-time instead of wave functions.

Imagining a particle at some initial position in the (x,y)-plane, and we want to know what path it will take to some final position. By the classical least action principle, the particle will take a path between the two positions that costs the least energy. But, if the particle is a quantum particle, it's not really localized at a point. Instead, the particle is a wave… and it doesn't take one path from the initial position to the final position, it takes all possible paths.



The basic idea then is to construct matrix elements of the time-evolution operator by summing all possible paths between two points and weighting the paths by the classical principle of least action. Remarkably, this gives identical results to solving the Schrödinger wave equation, but is much more concise a model.

The Hamiltonian generates the time evolution of quantum states. If $|\psi(t)\rangle$ is the state of the system at time *t*, then

$$H\,|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle\,.$$

This equation is the Schrödinger wave equation

From the Schrödinger wave equation we of course get the time evolution operator to work with.

Given the state at some initial time (t = 0), we can solve it to obtain the state at any subsequent time. In particular, if H is independent of time, then

$$|\psi(t)\rangle = e^{-iHt/\hbar}\,|\psi(0)\rangle\,.$$

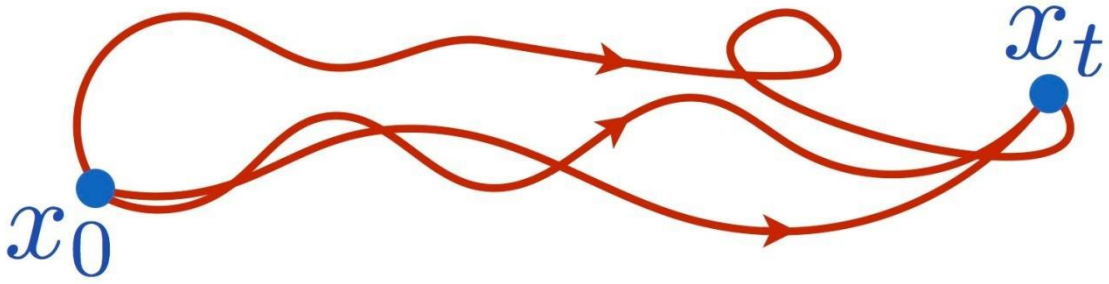
By the homomorphic property of the functional calculus, the operator

$$U = e^{-iHt/\hbar}$$

This is the time evolution operator, or propagator, of a closed quantum system.

Given the time evolution operator we can construct a matrix element which is a probability amplitude for the transition of a particle from state X0 to Xt

$$\left\langle x_t \middle| e^{-i\hat{H}t/\hbar} \middle| x_0 \right\rangle =$$

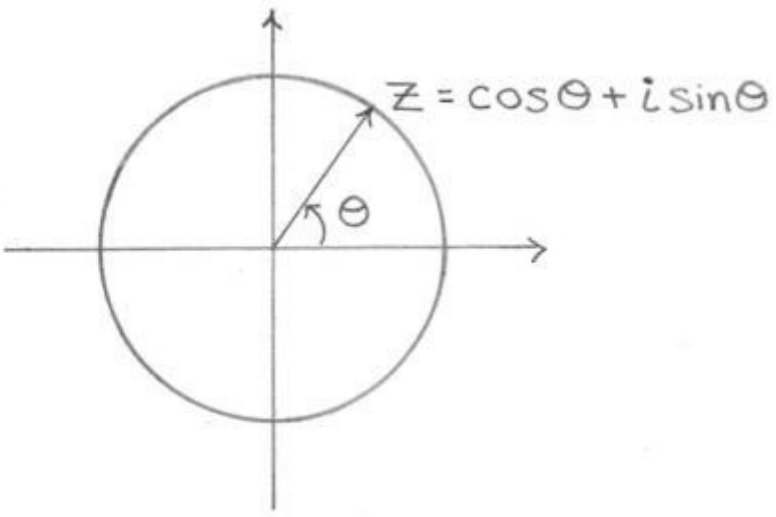


The quantum action, **S**, is represented here as being the phase of each path being determined by ∫ ***L*** **dt**, for that trajectory, where L is the Lagrangian. Lagrangian mechanics is used for discrete particles each with a finite number of degrees of freedom. Furthermore, the action, S, is a relativistically invariant operator. **[2].**

Quantum mechanics is a probabilistic theory, hence the principle of least action has to be written to make a probabilistic statement. Suppose we have a quantum particle with some wavelength or frequency (either will do—since they're inverses of each other) which we just measured at some initial point. We want to know the probability of finding it at some final point.

Each path the particle takes is equivalent to a probability wave as described by the Schrödinger wave equation. Waves are in fact represented here in the complex plane as being circles traced out by the wave-form.

This phenomenon is inherently related to the nature of the complex plane, where the complex number **z=exp(iθ)=cos(θ)+isin(θ)** acts like a unit vector with a phase angle θ in the complex plane:



Euler's Equation can then be used to establish the fundamental relationship between the trigonometric functions and the complex exponential function. Euler's formula states that, for any real number x:

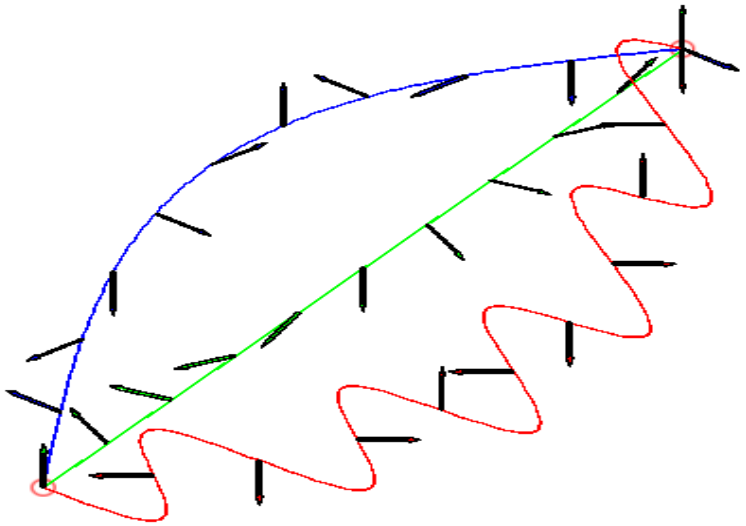
$$e^{ix} = \cos x + i \sin x$$

Where e is the base of the natural logarithm, i is the imaginary unit, and cos and sin are the trigonometric functions cosine and sine respectively, with the argument x given in radians.

The path of least action arises since the unit vector, z, tend to cancel out when the phase angles are out of phase with each other, and the vectors tend to add together when the phase angles are in phase with each other:

By taking each unit vector along probability wave represented by the path the particle takes, we form a series of arrows and make them oscillate around the complex unit circle at a frequency based on how difficult it is for a particle to travel through a given point—i.e. the harder the path to travel, the slower the arrow spins. (oscillating with frequency equal to the action.)

As discussed before that the particle takes all paths between the initial state and the final state. So we make our arrows follow each path and rotate them as we go along, as shown below.



Hence, in the path integral interpretation, the clockwork of unit vector arrows are along every possible path along the initial and final state (in this case 3 possible paths, red, green and blue). The paths are different only insofar as the arrows are rotated at different frequencies equal to the deviation from the path of least action, S. In other words the arrows are rotated with a frequency equal to that of the action imparted to deviate the particle from that of the least action possible. Therefore the path of least action itself should have the lowest frequency of unit vector rotation, where an increase in frequency corresponds to an increase in action.

Ref (3)

Using path-integrals has revolutionized quantum dynamics by by-passing the task to solve the Schrödinger equation itself (which is possible only for few-atom systems).

Avoiding going down the path of irreducible complexity, it is important to clarify that the least action principle is itself an extension of the principle of conservation of energy.

Viewing Meta-Heuristic Synchronization using the Path Integral :

The resemblance of the unitary time-evolution operator to the monotonically decreasing signalling function is obvious in their respective mathematical forms:

For one thing, the signalling algorithm is readily compatible with Lagrangian mechanics which are inherently used for discrete particles each with a finite number of degrees of freedom. Our Signalling Lagrangian, **Lf**, for each signalling node are in-effect integrated across the time-domain in the path it takes to achieve synchronization, hence leading to the total Signalling Action **Sf**, represented here as being the phase of each path being determined by $\int \boldsymbol{Lf}\,\mathbf{dr}$, for that trajectory, where Lf is the Lagrangian.

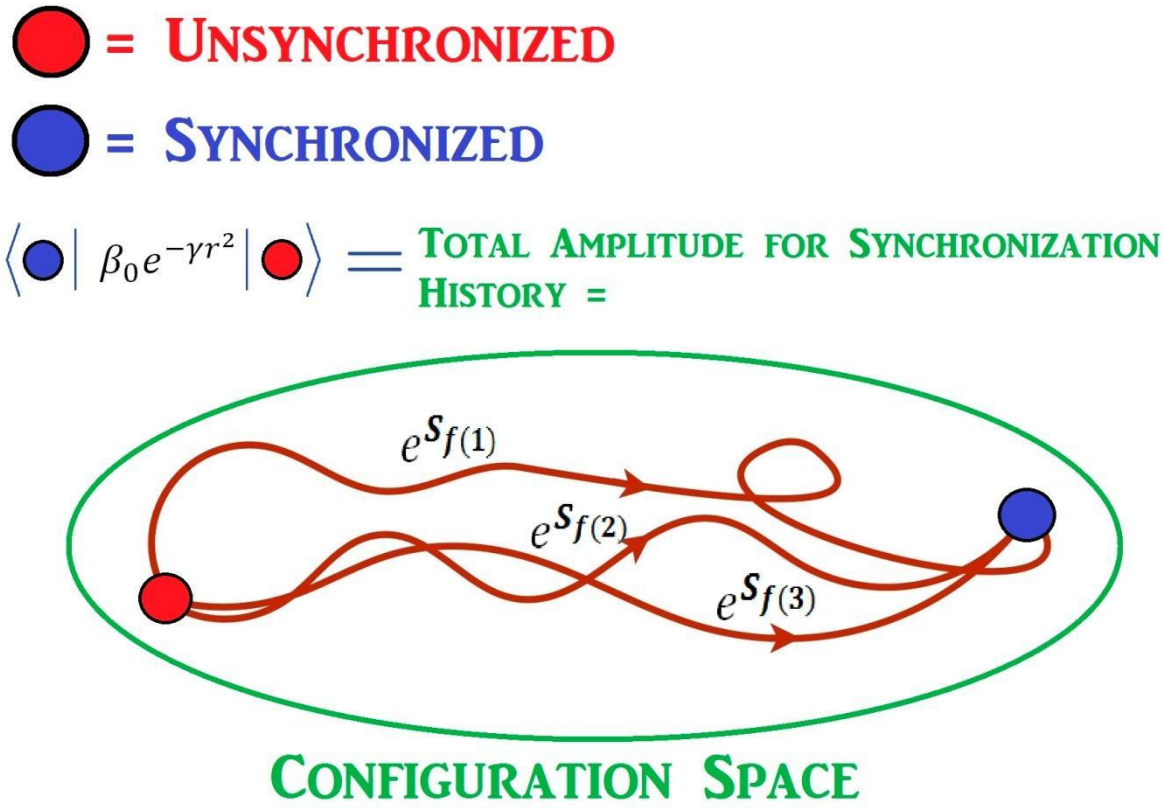
$$S_f = \int L_f \, dr = ln(\frac{I(r)}{I_0}) = -\gamma r^2$$

The frequency of signalling is then, as an oscillation in the complex plane:

$$e^{S_f}$$

However, we can investigate go further and compare the probabilistic representation of the path integral theory with the description of the algorithm as a system of discontinuous pas-coupling

Using this in the path-integral view we can represent the synchronization procedure as:



In this convention, the signalling action, **Sf**, which is a real operator is defined as being essentially characteristic of the physical and environmental parameters of the system – namely the light absorption coefficient and the power law over the distance by which it is subject to

Since each deviation from the path of least action is, just like in the case with a quantum particle, proportional to the action imparted on the system, the signal activation function will activate with a frequency proportional to the action. Therefore, the path of least possible action, which occurs during complete internal synchronization, should have the lowest frequency of signal function activation.

An increase in frequency of signal activation indicates a deviation from the path of least possible signalling action.

Therefore, when using any metaheuristic algorithm, when looking for any deviations, i.e. faults, between the coupled oscillators when internal synchronization is achieved we should just have to look for any changes in the frequency of the signalling between the different nodes.

This also explains how the self-organising system restores itself – as a fault is equivalent to a deviation from the principle of least action, which is energetically unfavourable.

What do I mean by this?

I mean that when an interaction occurs with our signalling model, the deviation leads to a deviation from the path of least signalling action to achieve synchronization, as we have imparted our own equivalent signalling action, S, onto the system. This causes an initial synchronization collapse. However, the action, S, imparted towards the path will lead to an increase in the rate of cycles, i.e. an increase in the signalling action away from the least signalling action, and thus an increase thee probability of achieving synchronization with neighbours again. In effect, internal processes are stimulated under any signalling action, increasing their state of energy, which is "unfavoured" - unfavoured as the increasing energy state will itself drive the quantised nodes, which have a discrete threshold, that make up the system to spend more energy and thus increase cycling which overall increases the probability that the area under interaction will be synchronised with its neighbours again. Therefore the system is able to restore itself back to synchronization.

The fact that the system regulates itself to the most energy efficient configuration possible, which is to the state of uniform synchronisation is a key conclusion of studying the metaheuristic algorithm in the context of the sum over histories of the path integral representation.

A Metaheuristic View Of Quantum Dynamics : Boltzmann Machines

An interesting consequence of this new point of view we are having is the fact that the representation works both ways and it makes sense to think of the quantum transitions themselves, i.e. an evolution from state x0 to xt, can be thought as being

in-effect, characteristic of a meta-heuristic function. This has interesting ramifications in the field of quantum computing, in which it has already been proven experimentally that meta-heuristic algorithms can be in fact solved using certain transitions on certain machinery, with the problem represented as matrix elements, An example in quantum computing is representing a computational procedure as a series of quantum tunnelling on a Josephson Junction network.

Josephson Junction circuits which in fact perform meta-heuristic calculations, such as simulated annealing, have already been developed by D-Wave Systems in Vancouver, Canada. Our model at the very least provides a way, using the path-integral interpretation, an abstract model of what the physical mechanism is in the ability of these systems to solve these meta-heuristic algorithms, such as simulated annealing, using quantum transitions rather than simply take it for granted that the devices just happen to perform these tasks.

Simulated annealing is probably the best example of modern metaheuristic algorithms, and it was developed by Kirkpatrick, Gelatt and Vecchi in 1983 [7], inspired by the annealing process of metals during heat treatment and Metropolis algorithms for Monte Carlo simulations.

The basic idea of the simulated annealing algorithm is similar to dropping some bouncing balls over a landscape, and as the balls bounce and loose energy, they will settle down at some local minima, sometimes referred to as a *spurious attractor*. If the balls are allowed to bounce enough times and loose energy slowly enough, some of the balls will eventually fall into the globally lowest locations, and hence the minimum will be reached. Of course, we can use many balls (parallel simulated annealing), or use a single ball to trace its trajectory (standard simulated annealing). The optimization process typically starts from an initial guess with higher energy. It then moves to other locations randomly with slightly reduced energy. The move is accepted if the new state has lower energy and the solution improves with a better objective or lower value of the objective function for minimization.

With non-stochastic networks, the appearance of local energy minima as a result of the network internally signalling along the path of the least signalling action possible can cause the signalling between the different nodes to get stuck in a region which do not satisfy the real constraints of the network very well.

This can be seen in the following network, simulated in Matlab.

In the simulation, we have a grid network of 10*10 nodes with input signal in the form of an input vector which is weighted by the size of the width 2D grid:

Weights = 2Dweights(GridWidth);

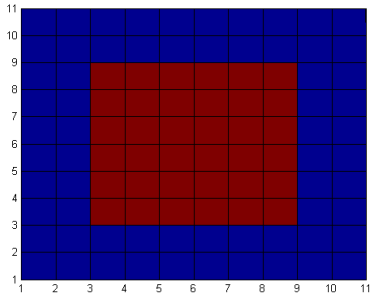
This allows the plotting of the input vector as an inputWidth * inputWidth grid

ActivationGrid = zeros(GridWidth);

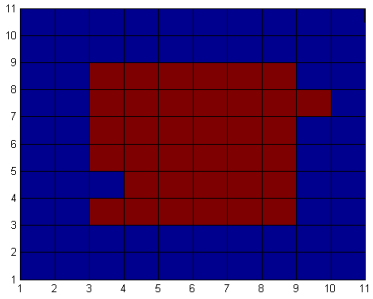
we use our signalling activation function as

```
%% as a 2D image, and think of the network as de-noising the image.
activations = ones(nInputs,1) * (-1);
for col = 3:(GridWidth-2),
  for row = 3:(GridWidth-2),
     activations((col - 1)*GridWidth + row) = 1;
  end;
end
for i = 1:nInputs,
  if rand(1,1) < SignalActionLevel
     activations(i,1) = activations(i,1) * (-1);
  end
end
```

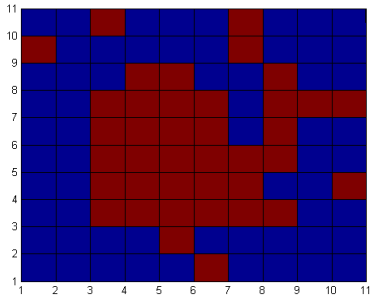
Signal Action Level (0->0.001):



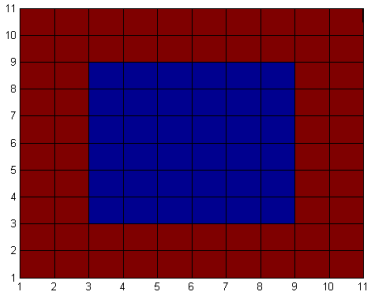
Signal Action Level (0.01):



Signal Action Level (0.1):



Signal Action Level (1->∞):



The signalling activations are attracted to the center of the network, which is where the local energy minima – i.e. the path of least signalling, is located. This can only disappear in the network by introducing more of an external signalling action.

Another way to think about it is if we think of the network as achieving synchronization in order to remove noise from an externally imposed signal, passed as a input vector through the network, then when the network gets itself trapped in a local energy minimum by the principle of least signalling action the only way to get it out of this minimum is to impose action, i.e. noise, into the network.

The main failing of this is the fact that we have not properly defined what the network constraints are in terms of network structures. The constraints themselves are really imposed by the weights of the inputs, $W_{i,j}$.

For a given input vector:

$$X^{0}_{i=1,2\ldots N}$$

In a neural network structure we weight the inputs and sum them through k-number of hidden layers in the following summation:

$$X_{j}^{(k)} = A(W_{0j}^{(k)} + \sum_{i=1}^{M_{k-1}} W_{ij}^{(k)} \cdot X_{i}^{(k-1)})$$

Our activation function which passes the weighted values through the k-number of layers, **A(x)**, is then in the same form as the signalling basis function for a metaheuristic algorithm, which is based on the Gaussian form of the signalling intensity which is monotonically decreasing.

The activation then takes the form:.

$$A(x) = (1 + e^{-x})^{-1}$$

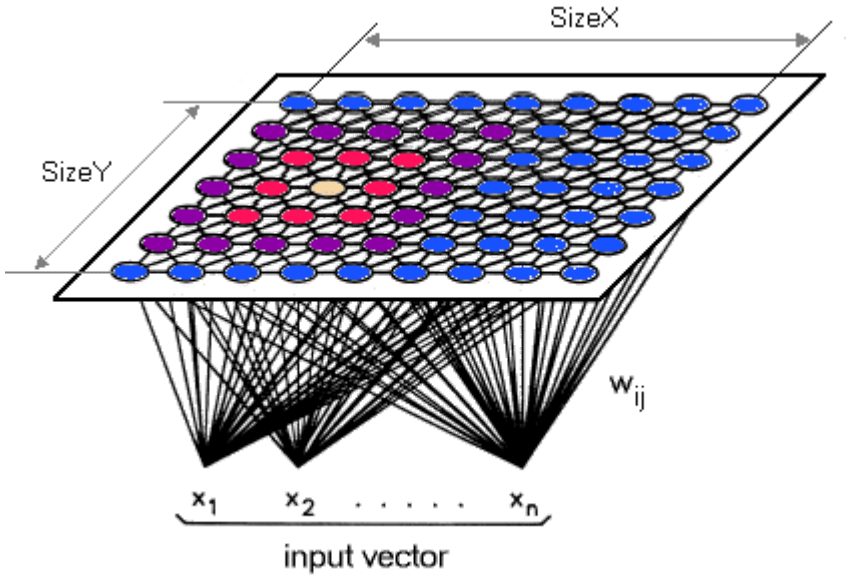
Which is a **sigmoidal activation function**. This is exactly the same form of equation as our **meta-heuristic activation function**.

Hence the “hidden” layers of the neural network will be a sum over histories of all possible weights the inputs take across the sigmoidal activation function in the neural network structure.

The output is then a finite discrete integer response, which could be represented on a digital number line, characteristic of the initial input vector, that has been effectively broadened by the k-layered neural network:

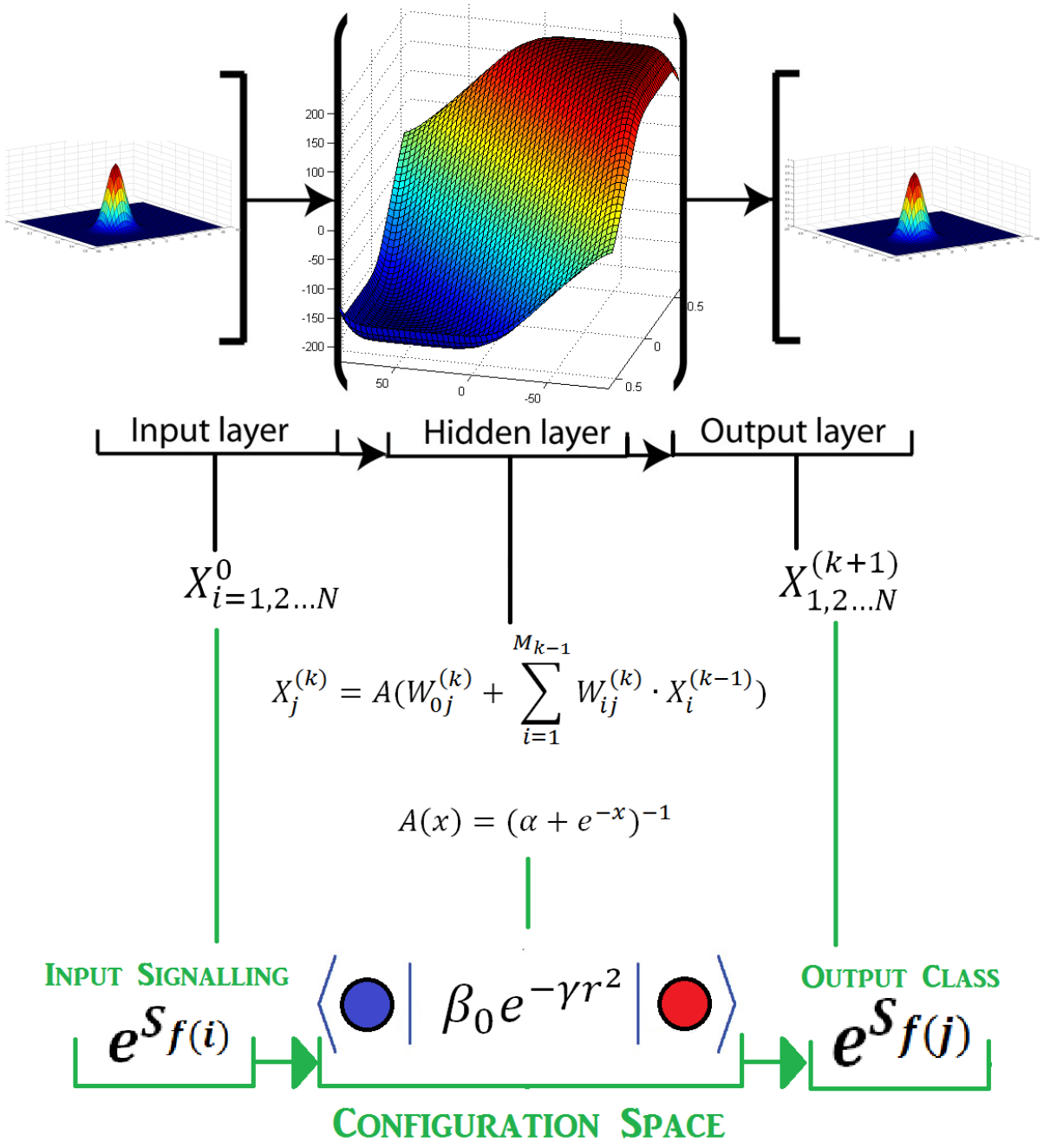
$$X^{(k+1)}_{1,2\ldots N}$$

This framework forms the classic and familiar neural network structure.



Using this information in our view of meta-heuristics under the path integral interpretation, we can then reduce the input vectors to an input signalling action, $S_{f(i)}$, oscillating at a certain input frequency in the complex plane, $e^{S_{f(i)}}$, which passes through the network, for example under sigmoidal or other Gaussian analog activation functions, in our path integral interpretation. This then gives an output signalling action, $S_{f(j)}$, of certain nodes in the network which will now carry a frequency, $e^{S_{f(j)}}$, characteristic of the input signal as interpreted by the neural network. The weights the input vector acts on in the neural network picture are then replaced by the paths in the sum over all possible taken in the time-evolution to achieve synchronization within the network.

I represent this construct in the diagram below, in which I have created a Radial Basis Function synchronised neural network in Matlab which receives a signal and creates an output response characteristic of this input, broadened by the k-number of paths of sigmoidal activation function signalling within the network:

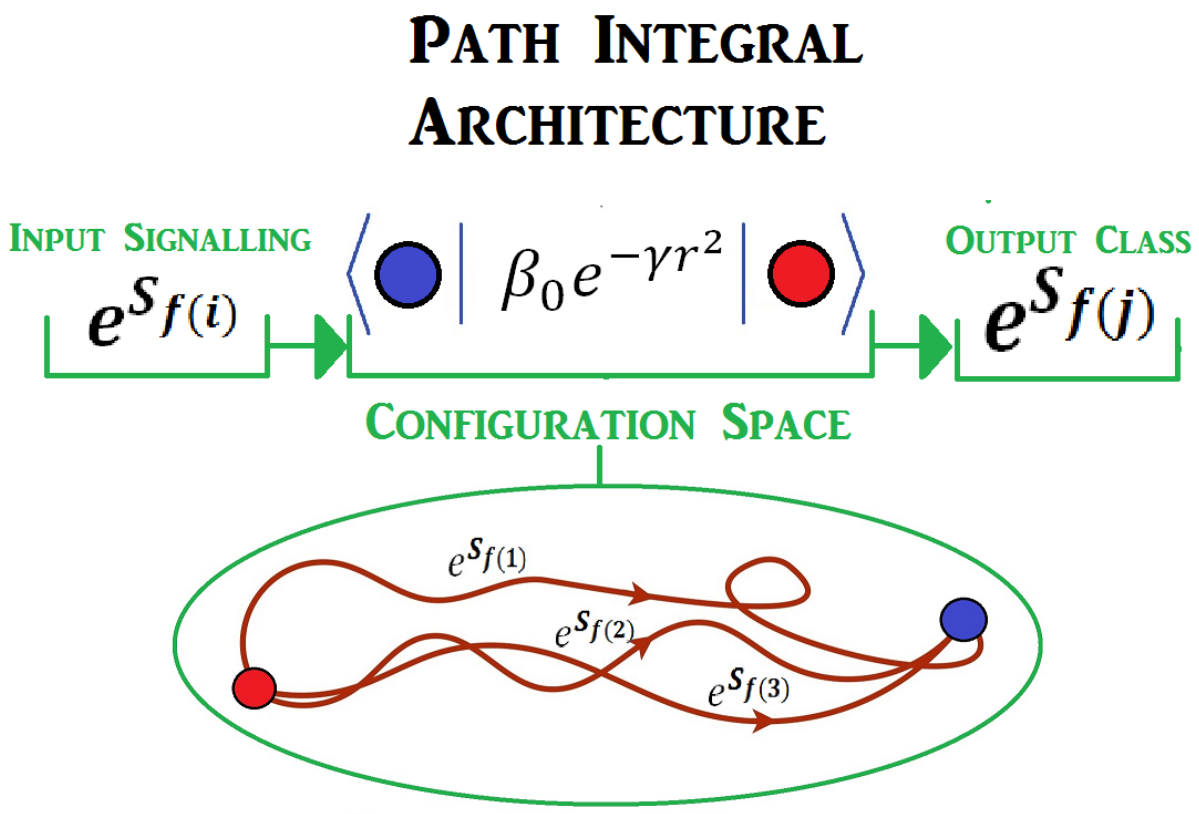


We then say there is an output “class” of internal node signalling action in the network from being stimulated by an external signalling action, in which sense our input signalling action is interpreted as a “class” by the nodes after passing through the neural network.

It is perhaps important to note that the signalling action towards the generation of nodal output classes is clearly not the core function of the neural network and is merely ancillary to it. By this I mean that the vector inputs and their associated weights cause action on the network but in this case are not in fact forming the

network structure itself. We are reminded that the network structure is formed, independently, by the time-evolution of the discontinuous pas-coupling generated by the meta-heuristic nature of the internal signalling along the path of least signalling, which is an extension of the principle of conservation of energy is itself as discussed earlier.

Hence we can now think of the established meta-heuristically synchronized system as an independent computational architecture, based on the path integral interpretation:



In the classical picture of such a network, which have local energy minima, we have a network made up of binary nodes (activations either 0 or 1), where the probability of achieving a coupling between the nodes would be set to 1 if and only if the connected nodes have the same activation, i.e A1=A2, i.e. the 2 activation levels have an allowed crossing. This has the effect of greatly limiting the nodal capacity of such systems.

To solve this problem we must relate our quantum mechanical system to equilibrium dynamics which will allow us to create a meta-heuristic representation of quantum diabatic behaviour which will induce the necessary condition of avoided crossing between nodes in the network.

Diabatic transitions can in fact be represented as a meta-heuristic algorithm therefore one need not go through the trouble of constructing a quantum computer

to simulate annealing in the optimum way. For example, one thing that a metaheuristic, under the lens of the path integral interpretation of quantum mechanics, can do is a form of self-generating quantum noise correction, just in the same way as true adiabatic quantum computer would.

Using quantum dynamics, we can say that the probability of a Diabatic Transition between 2 discrete energy levels E1 and E2 is given by the solution to the Landau-Zener Formula, solved by contour integration with the contour chosen to be in accord with the perturbative limit [7]

The Landau-Zener equation is:

$$P_{(E2)} = P_{(E1)} e^{-\Gamma 2\pi}$$

or

$$P_{D(E2-E1)} = e^{-\Gamma 2\pi}$$

Where:

$$\Gamma = \frac{g^2}{\hbar|\frac{dq}{dt}\frac{\partial}{\partial q}(E_2 - E_1)|} = \frac{g^2}{\hbar|\alpha|}$$

Where $\frac{dq}{dt}$ is the Landau-Zener Velocity for the perturbation variable q (where q is either a quantisation, i.e. a charge, of an electric or magnetic field, the length scale of a molecular bond, or any other perturbation to the system), and E1 and E2 are the energies of the two diabatic (crossing) states.

The quantity g is the off-diagonal coupling element of the two-level system's Hamiltonian that couples the crossing states, and as such it is half the distance between the two unperturbed eigenenergies at the avoided crossing, when E1 = E2.

We can write the transition probability as:

$$P_D = e^{-\frac{g^2 2\pi}{\hbar|\alpha|}}$$

The adiabatic theorem in quantum mechanics therefore tells us, that as the difference between the energy states, α, Goes to 0, then the transition probability P_D goes to 1 (unity).

This means that the quantised transitions between the energy states themselves are supressed at the point of avoided crossing where 2 or more eigenvalues of the Hamiltonian cannot become equal in value.

This implies reversibility – i.e. no net work done, in a cyclic process.

In terms of thermal dynamics then we examine the transition probability in terms of the Energy Hamiltonian

Thermal Equilibrium Dynamics:

$$z = Tre^{-\beta H} = Tre^{-\beta e(n)}$$

Since the occupation space can only be |0> or |1> for fermions,

$$z =< 0|e^{-\beta e(0)}|0 >+< 1|e^{-\beta e(1)}|1 >$$

$$= 1 + e^{-\beta e}$$

An imaginary time path integral approach to thermal equilibrium dynamics can be performed by the fact that our canonical density operator, $e^{-\beta H}$**,**
Is related to the time evolution operator in the Schrödinger scheme, $e^{\frac{-iHt}{\hbar}}$

By the analytic continuation of time, we define:

$$t =- i\hbar\beta$$

This is known in field theory as a Wick rotation from a Minkowskian to a Euclidean field theory **[9].**
.
However, as we have shown, we can write the coordinate matrix elements of either $e^{-\beta H}, e^{\frac{-iHt}{\hbar}}, e^{-\frac{g^2 2\pi}{\hbar|\alpha|}}$ as a sum over histories in the path integral interpretation and furthermore can represent the path integral as a meta-heuristic signalling scheme in a neural network.

We can then unite the approach to achieve meta-heuristic synchronization with the adiabatic theorem of equilibrium dynamics which is where the longer the thermal

system is allowed to transition to reach its final state in the time evolution from x0 to xt, the more likely it will have stayed in the ground state (or, said another way, the longer the system is allowed to run, the less likely there is to be random excitations, which basically translate to noise).

Therefore, in order to reduce noise or error in such a system, all you need to do with a metaheuristic algorithm is to simply run the algorithm in the adiabatic fashion for a longer time. i.e. if you ran it for T = infinity, you'd be 100% accurate.

This is exactly the criterion for the standard metaheuristic signalling algorithm, as the signalling actions move towards the least signalling, and thus most energetically conservative, action possible.

This leads us to thinking of the neural network as a Boltzmann Machine Framework which can overcome the problem of systems of binary nodes getting stuck in local energy minima by the addition of three key features:

- **Stochastic activation function:** the state a unit is in is probabilistically related to its Energy gap. The bigger the energy gap between its current state and the opposite state, the more likely the unit will flip states.
- **Temperature and simulated annealing:** the probability that a node is on is computed according to a activation function of its total weighted summed input divided by T. If T is large, the network behaves very randomly. T is gradually reduced and at each value of T, all the units' states are updated. Eventually, at the lowest T, units are behaving less randomly and more like binary threshold units.
- **Contrastive Learning:** A Boltzmann machine is trained in two phases, "clamped" and "unclamped". It can be trained either in supervised or unsupervised mode; this type of training proceeds as follows, for each training pattern:
    1. *Clamped Phase:* The input units' states are clamped to (set and not permitted to change from) the training pattern, and the output units' states are clamped to the target vector. All other units' states are initialized randomly, and are then permitted to update until they reach "equilibrium" (simulated annealing). Then learning is applied.
    2. *Unclamped Phase:* The input units' states are clamped to the training pattern. All other units' states (both hidden and output) are initialized randomly, and are then permitted to update until they reach "equilibrium". Then *anti- learning* (learning with a negative sign) is applied.

The above two-phase learning rule must be applied for each training pattern we feed into the network and for a great many iterations through the whole training set. Eventually, the output units' states should become identical in the clamped and unclamped phases, and so the two learning rules exactly cancel one another. Thus,

at the point when the network is always producing the correct responses, the learning procedure naturally converges and all weight updates approach zero.

The stochasticity enables the Boltzmann machine to overcome the problem of getting stuck in local energy minima, while the contrastive learning rule allows the network to be trained with hidden features and thus overcomes the capacity limitations.

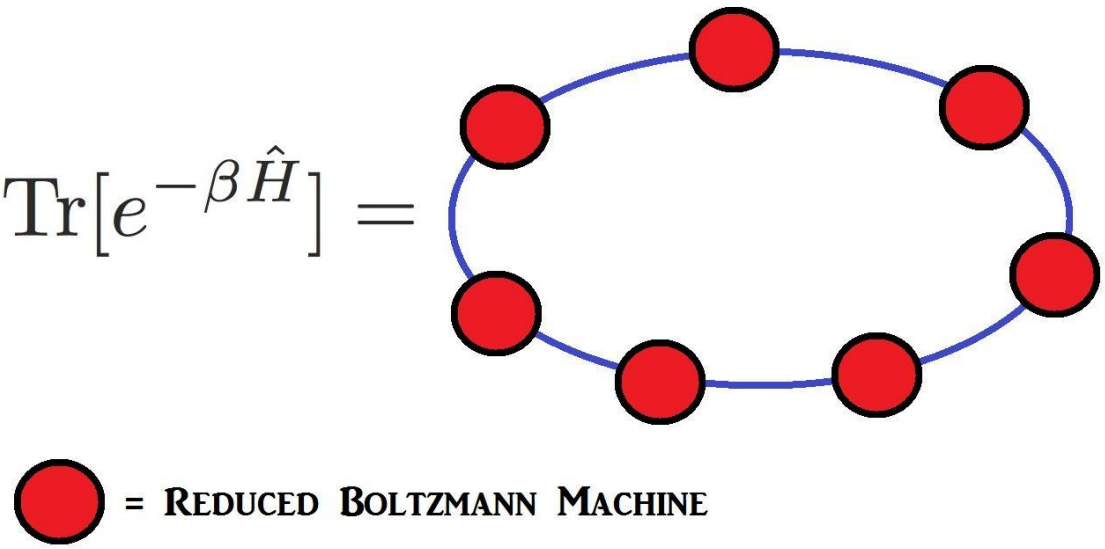
Boltzmann learning is very powerful, but the complexity of the algorithm increases exponentially as more neurons are added to the network. To reduce this effect, a Reduced Boltzmann machine (RBM) can be used. The hidden nodes in an RBM are not interconnected as they are in regular Boltzmann networks. Once trained on a particular feature set, these RBM can be combined together into larger, more diverse machines, such as polymer rings of different nodes.

Because Boltzmann machine weight updates only require looking at the expected distributions of surrounding neurons, it is a plausible model for how actual biological neural networks learn.

Boltzmann learning is statistical in nature, and is derived from the field of thermodynamics. It is similar to error-correction learning and is used during supervised training. In this algorithm, the state of each individual neuron, in addition to the system output, are taken into account. In this respect, the Boltzmann learning rule is significantly slower than the error-correction learning rule.

Reduced Boltzmann Machine learning is particular interest when considering so-called polymer rings and the means used to simulate them.

Polymer rings are essentially compact Gaussian chains of individual reduced Boltzmann machines subject to strong interactions with neighbouring chains due to the presence of topological constraints.



Therefore, one might assume that ring polymer simulations that use a Reduced Boltzmann Machine metaheuristic are potentially much more accurate and consistent with nature than using the current artificial dynamical Monte Carlo simulations. We also need not even think about in the sense of performing quantum evolution in terms of path integrals on polymer chain simulations, as doing a metaheuristic simulation is functionally equivalent as a quantum path integral. In effect, every time we do a metaheuristic simulation, we are in fact performing a quantum path integral.

Entanglement is the core feature of quantum phenomena, as it displays how quantum behaviour is irreducibly random with non-locality being a core feature of the probabilistic wavefunction.

The Von-Neuman entropy gives us a comparable way to measure the degree of entanglement of a multipartite system, such as the Boltzmann machines.

The entanglement for a pure bipartite state $|\psi_{AB}\rangle$ is given by the von Neumann entropy of the subsystem state ρ_A or ρ_B:

$$S(|\psi_{AB}\rangle) = -\mathrm{Tr}\rho_A \log \rho_A = -\mathrm{Tr}\rho_B \log \rho_B.$$

It is possible to use metaheuristics to develop Quantum computing of a kind using complex networks, although interesting in and of itself, has very broad applications in any field involving a networked topology, which covers topics such as polymer interactions under certain input actions and doubtlessly much more.

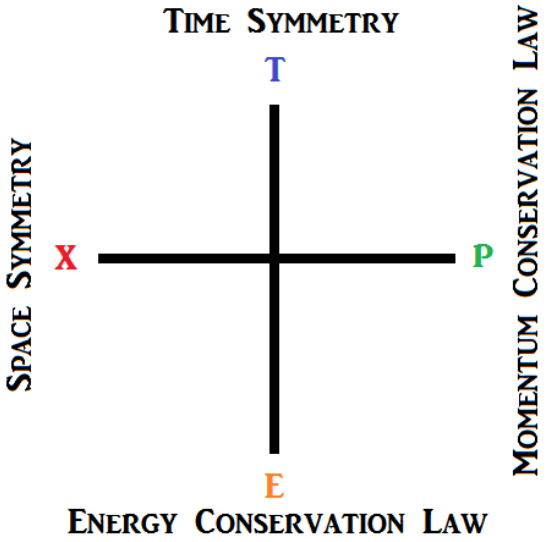
Moreover, by the path-integral interpretation of machine signalling, any system with sufficient discontinuous pas-coupling should therefore have a self-generating property of meta-heuristic behaviour which spans a huge realm of empirically observed behaviour in both organic and inorganic signalling systems.

To study this behaviour analogously in hardware and software is possible in real time using relatively simple technology in which non-preprogrammed quantum-like behaviour is observable such as the establishment of a ground state energy in multipartite systems by means of metaheuristic synchronization in which local excitations can propagate into non-local behaviour as showcased here in a schematata of “quantum fireflies”

Feynman's Fireflies Simulation - Quantum Synchronization in Varying Topology

Applying The Path-Integral Interpretation of Signalling to Systems of Quantum Entanglement and Condensates.

More interesting still is the principle of the nature of quantum entanglement and how the entangled states are not merely energy Hamiltonians evolving in a sum-over-histories across a unitary time operator, as quantum tunnelling is, of two particle states in time but operating over a symmetric equivalent, namely a sum-over-space across a unitary momentum operator of two particles entangled in a topology.



In the path integral interpretation of quantum mechanics the sum-over-histories approach can be used in the context of 2 energy states that undergo a transition from one to another under a unitary time operator.

Time and energy variables have symmetric equivalents, with space and momentum.

Following from this, it is possible to then think about using the path integral interpretation involving 2 entangled particles over a given space (i.e. a topology) that have a sum over histories across a momentum operator.

Moreover, if entanglement is equivalent to some topological action, **St**, in this view then this action must itself be the integral product of the Euler-Lagrange equations that is integrated over the topological domain for a scalar field.

For entangled particles, this action is for a free massless scalar field under calculus of variations.

The transformation rule for a scalar field Φ under the scaling transformation:

$$x^{\mu} \rightarrow x'^{\mu} = e^{\lambda} x^{\mu} \quad \textbf{Ref (6)}$$

Where λ is a positive environmental parameter

The transformation rule for the scalar field then reads:

$$\Phi'(x') = \mathrm{e}^{-\lambda d_\Phi}\,\Phi(x)$$

Where d_Φ is the canonical scaling dimension of the field Φ, i.e.

$$\mathbf{DIM}(\Phi) = meters^{-d_\Phi}$$

The action is the space–time integral of the Lagrangian density $L(\Phi(x),\partial\mu\Phi(x))$ which depends on the field and its derivatives.

The Lagrangian itself can be written as the difference of the kinetic and potential energy density of the field.

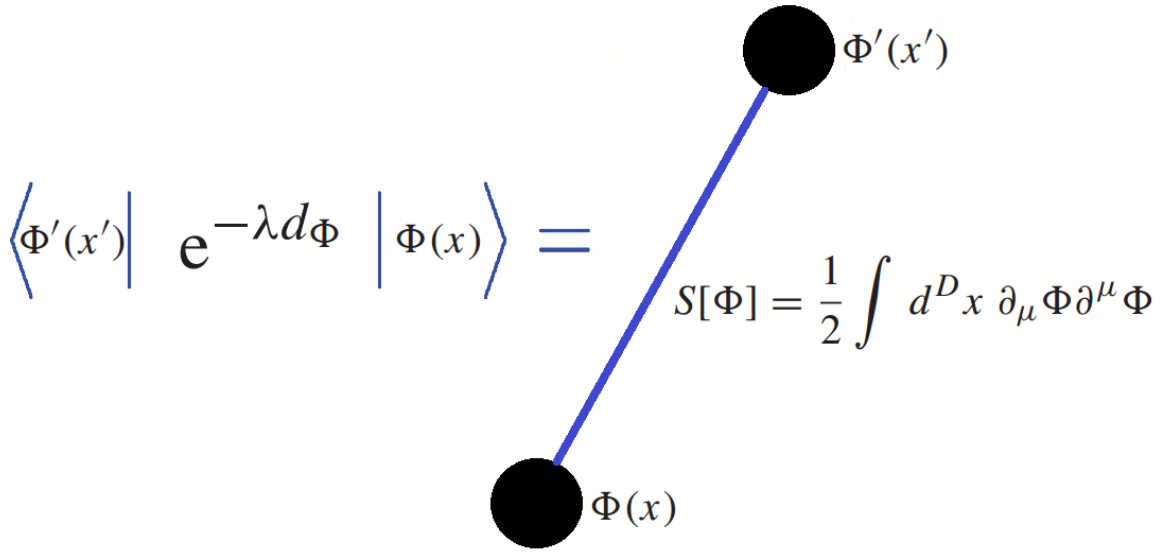
$$L = K(\partial_\mu \Phi \partial^\mu \Phi) - V(\Phi)$$

The action functional for a free massless scalar field, propagating in D-dimensional space-time, therefore has the form

$$S[\Phi] = \frac{1}{2}\int d^D x\ \partial_\mu \Phi \partial^\mu \Phi$$

In the system of units where $h^- = 1$ the action should be dimensionless.

Representing the transformation as a matrix element gives us a novel path integral interpretation of quantum entanglement:



The action functional for the massless free scalar field, $S[\Phi]$, is therefore the least action that a quantum entangled state between 2 particles in space allows and any deviation from this, as in the tripartite system case, will be unfavoured as the action functional, which depends on space, will literally be topologically deflected from the path of least action, i.e. connected by a curved topology of the field rather than a straight topology.

The nature of the action is interesting in that unlike the quantum action in the standard path integral, which depends on time evolution, the entanglement action function has no time dependence whatsoever, only topological dependence. This explains quite nicely how entanglement operates instantaneously – there simply is no time-dependence in the action functional in the first place in this model.

Within the matrix element structure of entanglement, we have a probability amplitude that has a Gaussian structure. This amplitude, which we shall call the Entanglement Amplitude $A_{Ent}(\Phi)$has the form:

$$A_{Ent}(\Phi_j) = A_{Ent}(\Phi_i) e^{-\lambda d_\Phi}$$

$$A_{Ent}(\Phi_{j-i}) = e^{-\lambda d_\Phi}$$

A new variable to consider here is the length scale, l, which is proportional to our choice of λ.

For many entangled particle systems we would still have the task of fine-tuning the length scale, however the in the simplest case – i.e. 2 particle entanglement, where the probability of achieving a coupling is unity, we simply say that:

For:

$$A_{Ent}(\Phi_{j-i}) = 1,$$

$$d_{\Phi} = \iota = \frac{1}{\lambda}$$

The scalar field itself at this point must then take the form:

$$d\Phi(x) = \iota = \frac{g^2}{\hbar|\frac{d\lambda}{dx}\frac{\partial}{\partial\lambda}(I_{Ent}(\Phi_{j-i}))|}$$

$$\Phi(x) = \int_0^x \iota = \frac{g^2}{\hbar}\int_0^x \frac{1}{|\frac{d\lambda}{dx}\frac{\partial}{\partial\lambda}(I_{Ent}(\Phi_{j-i}))|}dx$$

$$\Phi(x) = \frac{g^2}{\hbar\lambda(1+ln(I_{Ent}(\Phi_{j-i}))}$$

Making a quantisation, an oscillation around the field say, then costs some discrete unit of angular momentum, **p**.

The transformation between functions in the position space x and the momentum space k is effected via the Fourier integral:

$$\Phi(\mathrm{k}) = [1/2\pi]^{1/2}\int_{-\infty}^{\infty} e^{-ikx}\psi(x)dx$$

where the momentum

p=ħk

$$p = i\hbar\frac{\partial\Phi}{\partial x}$$

$$p = i\hbar\frac{\partial\Phi(x)}{\partial x} = \frac{g^2}{|\frac{d\lambda}{dx}\frac{\partial}{\partial\lambda}(I_{Ent}(\Phi_{j-i}))|}$$

In effect then, λ is the quantisation variable of the field – of whatever form it may take (electric charge, spin, etc.) and can be represented purely as a unit of momentum, p, of the field or as <u>oscillation</u>, of a frequency 2π around the field space, in a unit of angular momentum where we then represent the entanglement intensity function as:

$$A_{Ent}(\Phi_{j-i}) = e^{-\frac{i2\pi p_{(j-i)}}{\hbar}} = e^{-\frac{ip_{(j-i)}}{\hbar}} = e^{S_{\Phi_{j-i}}}$$

So our action, S_{Φ}, on the field Φ is now quantised as complex momentum, i**p**, in units of Planck's Constant, **h**.

Ee can then say that the momentum transition rules for the massless scalar field is:

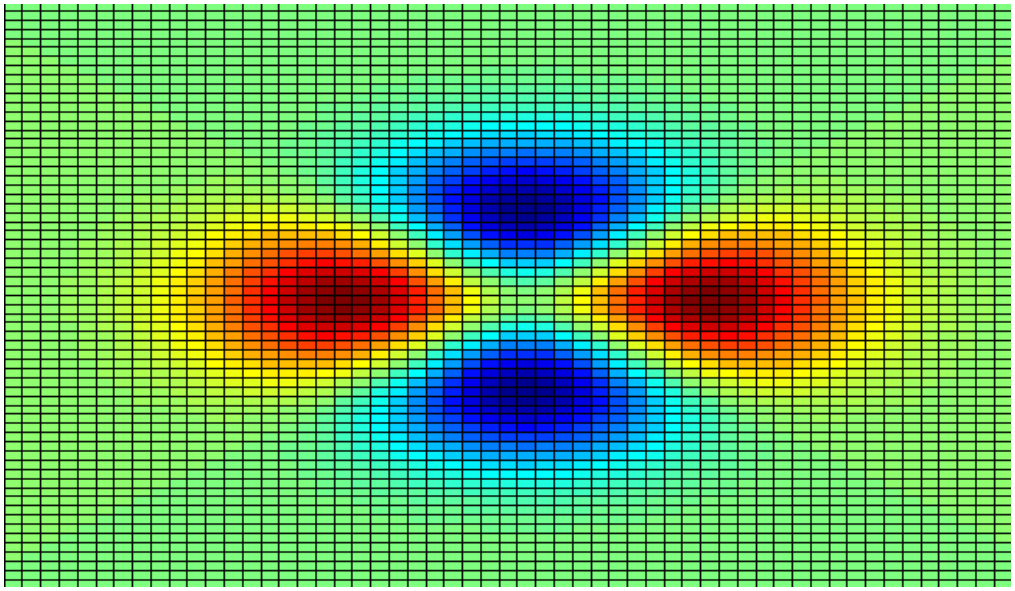
$$(|\Phi_i|^2 + |\Phi_j|^2)\Phi_i = i\hbar\frac{\partial\Phi_i}{\partial x} = p_i$$

$$(|\Phi_i|^2 + |\Phi_j|^2)\Phi_j = i\hbar\frac{\partial\Phi_j}{\partial x} = p_j$$

We can plot the single scalar field as a function of x:

$$A_{Ent}(\Phi_{j-i}) = (|\Phi_j|^2 - |\Phi_i|^2)e^{((-1)\cdot(|\Phi_i|^2+|\Phi_j|^2)}$$

in Matlab:

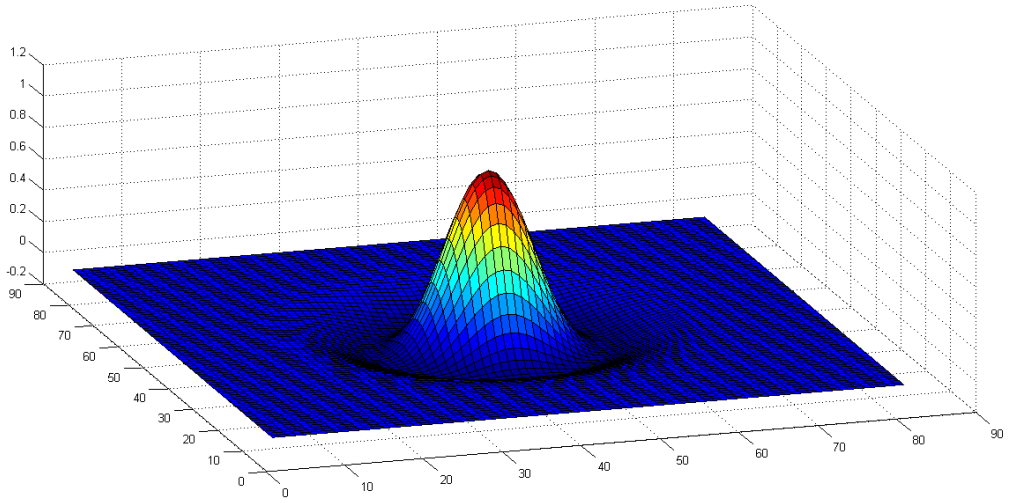


Note the appearance of avoidance crossing, as expected since we set the coupling probability to 1.

If, we were to introduce a breaking of symmetry with the scalar field:

$$(1 - |\Phi_j|^2 - |\Phi_i|^2)e^{((-1)\cdot(|\Phi_i|^2+|\Phi_j|^2)}$$

Plotting this we then get a Mexican hat shaped potential energy surface.



The quantisation energy state is then free to be at any point of lowest energy, which is within the rim of this field amplitude. Quantisation, represented as angular momentum, p, free to move at any point, x, within the field as a function of x is then subject to the Heisenberg Uncertainty Principle:

$$\Delta p \Delta x \geq \hbar$$

This causes a new problem to emerge; if we know

This is effectively the nature of spontaneous symmetry breaking.

The true lowest energy state is anywhere along the rim of the Mexican hat potential.

However, because of the uncertainty principle

In other words you can have a nonzero quantisation in the field space as some kind of condensate. A condensate then can be defined more fundamentally as a configuration of space that, with respect to whatever quantisation (charge, spin, etc.) we talk about, it is not affected by the addition or removal of the quantisation.

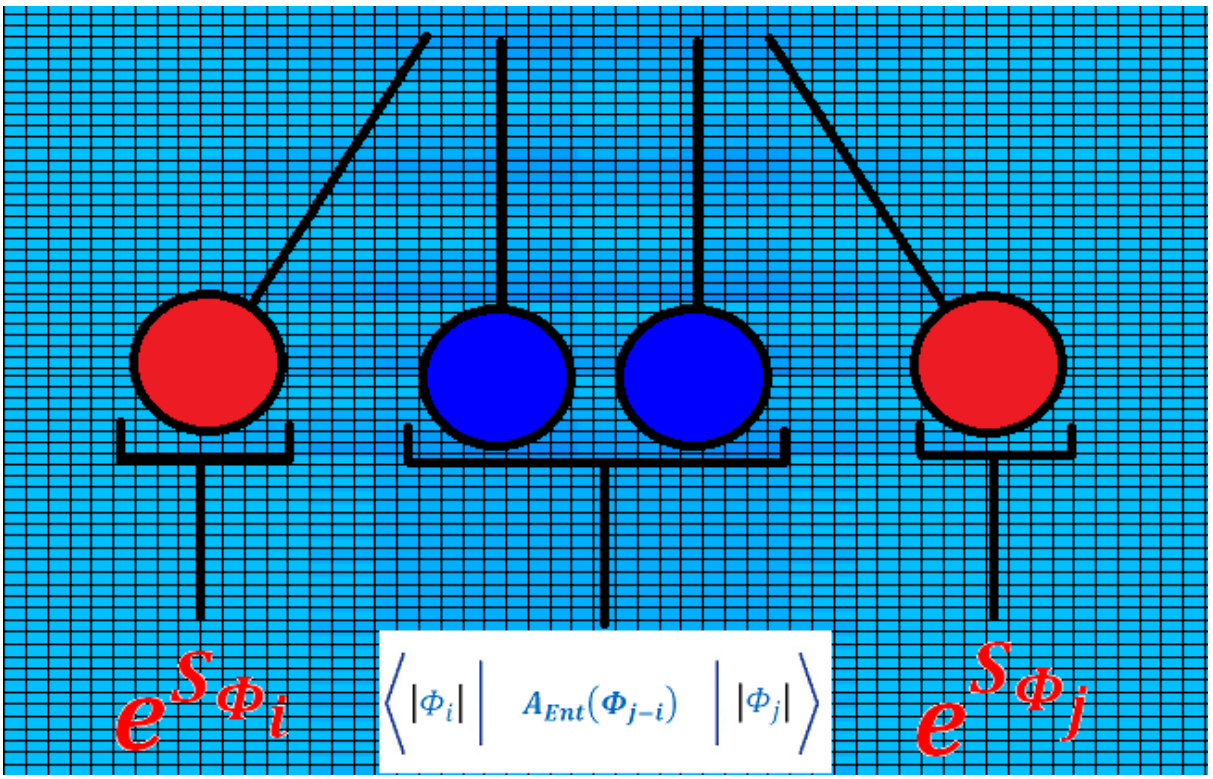
(Ref 8)

There are countless materials and phenomena that display macroscopic quantum behaviour and as such many can be treated as Bose-Einstein Condensates operating under a standard or modified Quantum Newtons Cradle. Lasers generated by electron-hole pairs in the form of a quantum dot will absorb a photon from a condensate, under an external field bias placed across the dot, which then re-emits the photon back into the condensate.

The same phenomena also happens in superconductors by which the nature of the charges in the field go under a spontaneous symmetry breaking where we then have the moving charges reduced to scalar fields, in the form of the so-called Cooper pairs, which are coupled together by a phonon absorbed from a condensate and re-emitted back into then condensate.

We can then construct a path integral-based model for scalar field actions on such condensates as a kind of “Quantum Newton’s Cradle”:

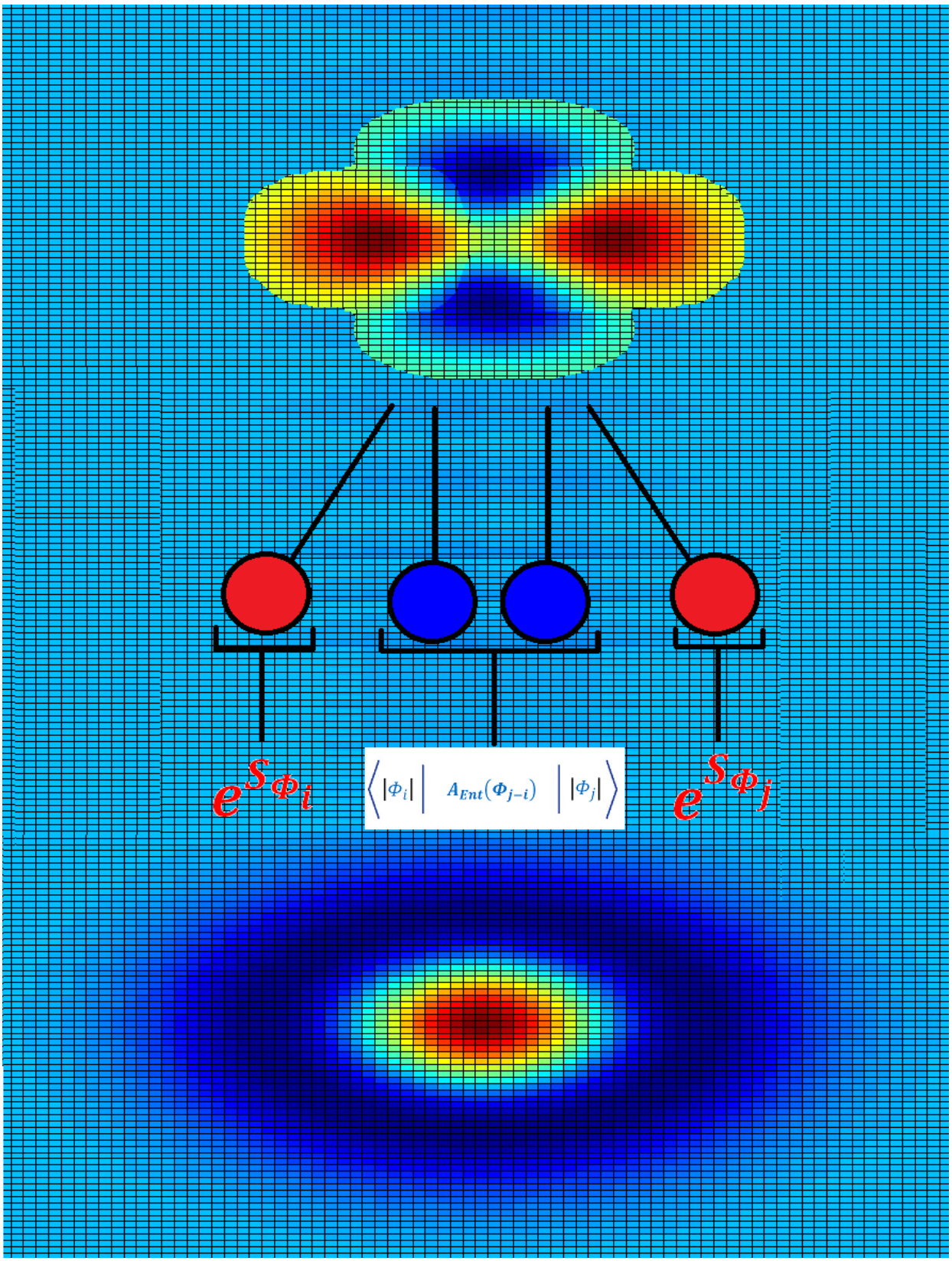
$$e^{S_{\Phi_i}} \rightarrow A_{Ent}(\Phi_{j-i}) = e^{S_{\Phi_{j-i}}} \rightarrow e^{S_{\Phi_j}}$$



The “hidden layer” is then what we refer to as the condensate itself.

The “energy gap” then at the point of avoidance crossing between the 2 field quanta of discrete momentum can then be thought as being caused by the quanta forming a condensate at the crossing.

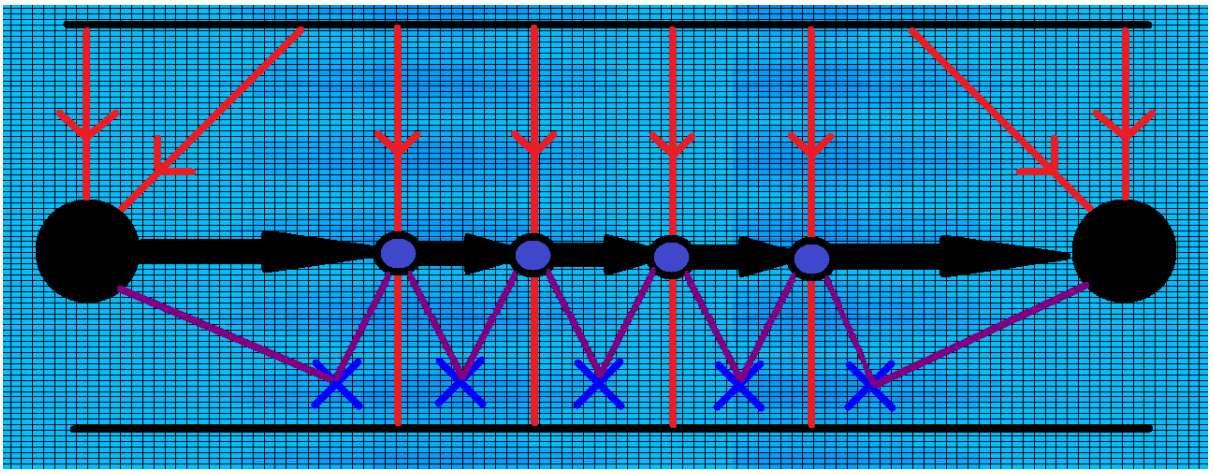
The following diagram summarised this:


$e^{S_{\Phi_i}}$
$\langle |\Phi_i| \; A_{Ent}(\Phi_{j-i}) \; |\Phi_j| \rangle$
$e^{S_{\Phi_j}}$

Going further in this view we can say the movement of any particle through what is classically considered to be a field is in fact the interaction of quantised units of action momentum on a condensate of field quanta.

The initial field perturbation is then our action on the particle through the field where it absorbs field quanta from the condensate and emits the quanta back into the condensate with no net effect.

So in any field (be it electic, magnetic, gravitational, ect) we have a new way to picture how a particle may move through a uniform classical field:



KEY:

= CLASSICAL PARTICLE

= CLASSICAL FIELD

= QUANTISED FIELD MOMENTUM

= PARTICLE CONDENSATE

= FIELD CONDENSATE

So instead of viewing the particle as perturbing the field, we should think about it as imparting a quantised unit of momentum on a condensate of field quanta that operates under a quantum Newton's cradle model.

We can extend this into a more easy to grasp form using a model based on representing the coupling between scalar fields as braids.

Future Proposals

If systems of quantum coherence, such as Bose-Einstein Condensates, and meta-heuristic coupling and synchronization display such similar mathematical properties, it is clear that meta-heuristic coupling and synchronization can be in principle implemented as the proper framework for the means of creating simulations of quantum behaviour.

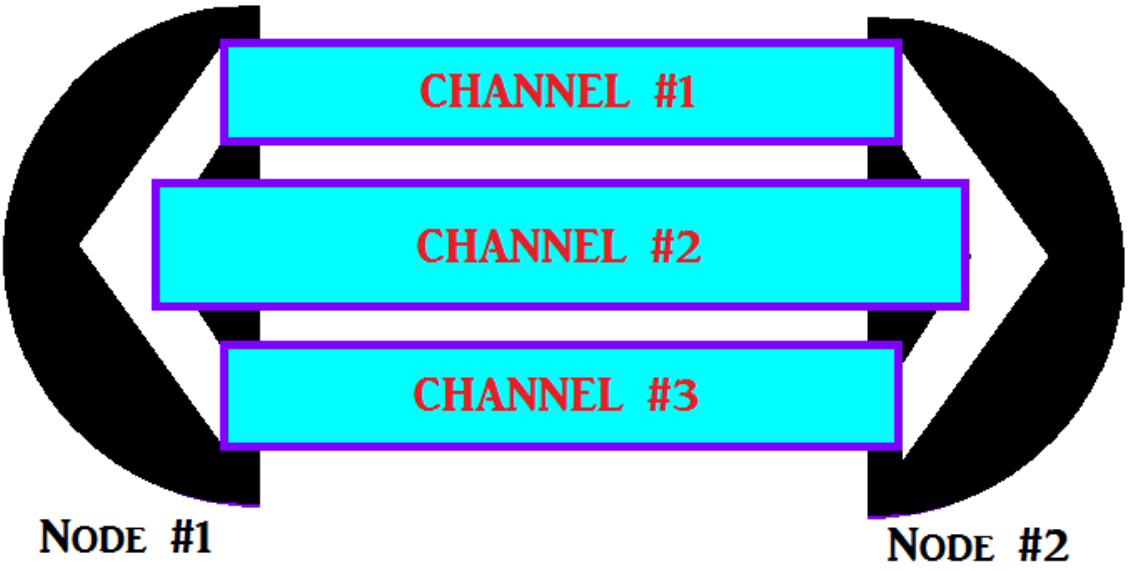
For example, if we have 2 devices that we want to share a discrete quantum signal between one another, perhaps for the purposes of transferring a quantum key for encryption, we might then tailor the design of the protocol that transmits the key to be one that employs a meta-heuristic algorithm that will adhere to the “Principle of Least Signalling” in order to best match the establishment of a “Least Entanglement”, which by the monogamy inequality has established is always the strongest form entanglement takes.

Given the interpretation, there may also be potential in the design and implementation of a secure system for transferring classical information based on this principle. As described, pure quantum entanglement is established over a topology from a single source, for example a pair of entangled photons generated by means of spontaneous parametric down conversion, SPDC.

As shown, meta-heuristic synchronization is a time-dependant evolution and this is the same case with quantum phase transitions and Boltzmann Machines under the path integral interpretation. Therefore a series of meta-heuristic oscillators, establishing synchronization over a time evolution, can be treated in effect as a technological alternative to pure quantum entanglement, which is much more difficult to establish and requires very high constraints on coherence, whereby any outside action imposed on the system will have the equivalent effect by disrupting any signalling procedure that occurs across the system when it is synchronized.

A “signalling entanglement” can in principle be made to exist then between 2 nodes in a network where they have been allowed to evolve under the metaheuristic framework to the point where they signalling the least amount possible to achieve synchronization. Hence the signal with the strongest entanglement shared between

any 2 nodes will be the one with the longest frequency based on the theory of the path integral interpretation of signalling.



From this we can create the basis framework of a secure optical communication signalling protocol, whereby we have 2 nodes we want to couple together, in synchronisation, and we have 3 meta-heuristic generated signal channels (#1,#2,#3), each with different starting frequencies, by which the signal with the lowest frequency will be, in effect, establish our signalling entanglement as the system evolves over time.

Such a system could in principle be operated using incoherent or coherent (laser) emitting diodes each operating at different signalling frequencies in superposition of one another under the initially unsynchronised state but then steadily evolving towards synchronization with the system evolving towards using only one of the LED or laser diodes being used as the longest carrier frequency for the entangled signal.

Implementing this on the simplest experimental circuitry would be the obvious next phase in development of this technology in tandem with computer simulation using neural networks for real world applications in similar vein to how we have applied the toy-model of metaheuristic Boltzmann Machines to the problem of ring reconstruction, which could be used for polymer ring simulations.

It would be interesting to see if we could combine the idea of signalling within Boltzmann condensates to problems in artificial intelligence and insect communications, for example creating a microcontroller simulation of a honeycomb of bees in a 2D surface, achieving synchronisation in a uniform lattice, which would then be stimulated by an outside bee which carries a noisy, repeating signal of a frequency proportional to the distance the bee had taken.

These are concepts which could be implemented in the context of networks and the modelling of the energy and momentum states within them that can be fundamentally modelled using signalling metaheuristics.

(1)- Xin-She Yang

(2)- Schwinger, J. (1951) Phys. Rev. **82,** 914–927

(3)- QED: The Strange Theory of Light and Matter, Chapter 2.

(4)- Monogamy of non-local quantum correlations Ben Toner Published 8 January 2009 Proceedings of the Royal Society A – Mathematical, Physical and Engineering Sciences

(5) - Masanes L, Acín A, Gisin N 2006 General properties of nonsignaling theories. Phys. Rev. A. 73, 012112

(6)- Lectures on Classical and Quantum Theory of Fields, Chapter 2 The Euler–Lagrange Equations and Noether's Theorem, Authors: Arodz, Henryk, Hadasz, Leszek – Springer 2010

[7] Optimization by Simulated Annealing
S. Kirkpatrick1, C. D. Gelatt Jr.1, M. P. Vecchi2
+ Author Affiliations
Science 13 May 1983:
Vol. 220, Issue 4598, pp. 671-680
DOI: 10.1126/science.220.4598.671

[8] The Landau−Zener Formula
Curt Wittig
Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482

[9] Quantum Dissipative Systems – Chapter 4 page 68 *Imaginary-Time Approach and Equilibrium Dynamics*
By Ulrich Weiss

[10] - Periodic waves in two-component Bose-Einstein condensates with repulsive interactions between atoms
A. M. Kamchatnov
Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow, 142190, Russia