Analysis of the Quantum Advantages for Deep Hedging

Soham Deshpande
Srinandan Dasmahapatra

October 2024

Abstract

Parameterised Quantum Circuits (PQCs) have opened many doors, one such being the use in financial markets. Through the Quantum Circuit Born Machine (QCBM), we are able to generate synthetic data that mimics the statistical distribution of the original dataset. In this research we are generating synthetic market data; here the dynamics of a given vanilla option is learnt. The market generator is then used to simulate the option to maturity with the intent of learning an optimal hedging strategy π^*, a showcase of a data-driven approach to risk hedging.

Contents

1 Related Literature **1**

2 Markets and Derivatives **2**
2.1 Market . . . 2
2.2 Derivatives . . . 3
2.2.1 Futures . . . 3
2.2.2 Options . . . 3

3 Deep Hedging **4**

4 Parameterised Circuits **5**
4.1 Born Rule . . . 5
4.2 State Preparation . . . 5
4.2.1 Amplitude Encoding . . . 5
4.2.2 Angle Encoding . . . 6
4.3 Measurements . . . 6
4.3.1 Complete Measurements . . . 6
4.4 Weak Measurements . . . 6
4.5 Quantum Circuit Born Machine . . . 7
4.5.1 Barren Plateau . . . 7

5 Quantum Architectures **8**
5.1 Butterfly . . . 8
5.2 Brick . . . 8
5.3 Pyramid . . . 8

6 Design **8**

7 Planning **9**

8 Risk Assessment **10**

9 Bibliography **11**

1 Related Literature

To place this research within the context of existing literature, we can split the project into 2 components: the market generator for options data, and quantum parameterised circuits.

Traditional methods of pricing options has shown to be ineffective for equity markets, new information resulting in rapid changes. Much of the available literature models the market as a smooth, continuous stochastic process within a Gaussian space. Such models are sufficient for common market activity but fail when presented with discontinuous moves in price. In the context of commodities, these can be reactions to geopolitical events or natural disasters; traditional models being incapable of capturing the effects. The introduction of Jump- diffusion models aimed to solve this issue though faces similar issues. In reaction we have recently observed non-parametric models that harness neural networks and machine learning which aim to demonstrate high accuracy on out-of-sample forecasts.

Similarly, the work around deep hedging has evolved with time, moving away from greek-based and statistical hedging towards a more sturdy framework using machine learning. Here a lot of work is being done, with many papers emphasising on using neural networks for optimising delta and gamma exposure. In recent times, we have also seen research done on using quantum computing to hedge portfolios, here the authors presented a quantum reinforcement learning method based on policy-search and distributional actor-critic algorithms.

There is an immense amount of research being done on exploiting the benefits of quantum computing, recent advancements being in quantum algorithms. These claim to provide exponential speed-up over classical methods, though in reality we see great complexity in state preparation, requiring $\Theta(2^n/n)$ circuit depth with n qubits or $\Theta(n)$ circuit depth with $\Theta(2^n)$ ancillary qubits. Here we see hybrid models such as Born machines and Quantum generative adversarial networks boasting high generalisation ability.

There has also been research in harnessing quantum weak measurements to enhance the ability of quantum machine learning algorithms. In quantum reservoir computing, the ability to retain information from past inputs plays a key role in processing temporal series and producing future predictions.

This research aims to combine the needs of financial firms in hedging portfolios using realistic market models provided by utilising QCBMs for time-series data in combination with quantum neural networks for learning optimal policies.

2 Markets and Derivatives

The market, though inherently can be thought of as a completely random process, where bids and asks are fulfilled, can be modelled as a stochastic process. The aim of this chapter is to serve as a brief introduction and set up notation for later chapters.

2.1 Market

Consider a market with a finite time horizon T defined on the probability space $(\Omega, \mathcal{F}, P)$ along with a filtration $\mathbf{F} = \{\mathcal{F} | 0 \leq t \leq T\}$ This can be thought of as an adapted (n+1) dimensional Itô process $X(t) = (X_0(t), X_1(t), ..., X_n(t))$ which has the form

$$dX_0(t) = \rho(t,\omega)X_0(t)dt; \quad X_0(0) = 1 \tag{2.1}$$

and

$$dX_i = \mu_i(t,\omega)dt + \sigma_i(t,\omega)dB(t); \quad X_i(0) = x_i \tag{2.2}$$

where $X_i(t)$ as the price of asset i at a given time t.
We can define a portfolio in the market as

$$\theta(t,\omega) = (\theta_0(t,\omega), \theta_1(t,\omega), ..., \theta_n(t,\omega)) \tag{2.3}$$

where the components $\theta_n(t,\omega)$ represents the number of units of a given asset held at time t.
Following from that, we can define the value of a given portfolio to be

$$V(t,\omega) = V^{\theta}(t,\omega) = \sum_{i=0}^{n} \theta_i(t)X_i(t) \tag{2.4}$$

Lastly, it is important to state that the portfolio is self-financing, any trading strategy α requires no extra cost beyond the initial capital

$$V(t) = V(0) + \int_0^t \theta(s) \cdot dX(t) \tag{2.5}$$

We can also make the following assumptions about the market:

- The market is liquid, allowing the trade to execute instantaneously
- We do not pay transaction fees
- There is no bid-ask spread, the price to buy and sell is the same
- Trading actions taken have no impact on the price of the asset traded

It is important to address that we have assumed a frictionless market, though unrealistic, this constraint will be held for simplicity and may be explored in further detail if time permits.

2.2 Derivatives

A derivative refers to any financial instrument whose value is derived from an underlying security, the most fundamental being futures and options.

2.2.1 Futures

A futures contract is a contract that gives the right and obligation to buy a given asset i a specified time T at price K. This can be thought of as the underlying asset for an options contract.

2.2.2 Options

The two types of options going to be explored are Puts and Calls; a Call option gives the owner the right but not the obligation to buy a given asset i at a specified price K at time T. Similar to the Call, a Put option gives the owner the right but not the obligation to sell a given asset i at a price K at time T. If the owner can exercise the option any time up to T, we call this an American option.
Pricing options

3 Deep Hedging

The problem of hedging a portfolio of derivatives is an important part of risk management used widely in financial institutions. We can picture a perfect, frictionless market where transaction costs are negligible and every asset in the space has a price. Here we can price and hedge perfectly. Unfortunately in practice, we experience incomplete markets due to frictions. This generates the need for complex, realistic market models that can account for these. Hedging can be thought of as a stochastic control problem where we have 3 actors, a market state,randomness from market returns and a hedging strategy. By making observations on the market state we are required to make adjustments to one's hedging policy.
We can formally define hedging as :

$$\min_{\theta \in \Theta} \rho(H^{\theta} - P)$$

where Θ is the set of admissible hedging strategies, P is the uncertain payoff of derivative at final time, H^{θ} is the uncertain payoff of hedging portfolio if strategy θ is used at final time, and ρ is a risk measure.

4 Parameterised Circuits

A basic understanding of quantum mechanics and computing is assumed.

4.1 Born Rule

An essential part of modern quantum computing involves the existence of the Born Rule. Born's measurement rule states that:

$$p(x) = |\langle x|\psi(\theta)|\rangle|^2$$

The state $|\psi(\theta)\rangle$ is generated by evolving the vacuum state $|0\rangle$ according to a Hamiltonian H that is constructed from gates. Once combined, the gates form a parameterised quantum circuit which is parameterised by using the variables governing each gate, θ. By tuning the values of θ_i one can allow for an evolution to any state that will serve as a solution to a given problem.
By taking the distribution associated to the state, $|\psi(\theta)\rangle$ we can treat the PQC as a generative model, which upon measuring in a given basis, will generate samples of a target distribution χ. This model is parameterised by θ, which defines a quantum circuit $U(\theta)$ made up of a set of quantum gates such that:

$$|\psi(\theta)\rangle = U(\theta)|0\rangle^{\otimes n}$$

By measuring the circuit, we can obtain samples. Producing samples that emulate the target distribution involves minimising the parameters of the circuit $U(\theta)$, a process once convergence is reached, will generate accurate samples. [**bornmachine**]

4.2 State Preparation

State preparation is a core part of the hybrid quantum computing model. This concerns the most efficient methods to input classical data into a quantum system. Without the use of ancillary qubits, we can expect an exponential circuit depth to prepare an arbitrary quantum state. Using them we can reduce the depth to be sub-exponential scaling, with recent advancements reaching $\Theta(n)$ given $O(n^2)$ ancillary qubits. [**stateprep1**] [**stateprep2**] We require state preparation to transfer the classical data, bits, onto the Hilbert space. This involves a function ϕ that maps the input vector to an output label. There are many encoding schemes, each of which aim to offer high information density and low error rates. The main methods of state preparation include: basis, amplitude, angle encoding, and QRAM.

4.2.1 Amplitude Encoding

$$|\psi_x\rangle = \sum_{i-1}^{N} x_i|i\rangle$$

Amplitude encoding consists of the following transformation:

$$S_X |0\rangle = \frac{1}{||x||} \sum_{i=1}^{2n} x_i |i\rangle$$

where each x_i is a feature of the data point x, and $|i\rangle$ is the basis of the n-qubit space. This does boast the advantage of being able to store 2^n features using only n qubits but does create a resultant circuit with depth $O(2^n)$.

4.2.2 Angle Encoding

The transformation used is:

$$S_x |0\rangle = \otimes_{i=1}^{n} cos(x_i) |0\rangle + sin(x_i) |1\rangle$$

It can be constructed using a single rotation with angle x_i which is normalised to be within $[-\pi, \pi]$. This allows us to store n features with n qubits.
A possible encoding method is dense angle encoding which allows us to encode $2n$ features in n qubits but hasn't been chosen in this project for simplicity; the same can be said for amplitude encoding.

4.3 Measurements

4.3.1 Complete Measurements

In quantum mechanics, we can define measurement to be any process that probes a given quantum system to obtain information, we may also refer to this process as a measurement in the computational basis when focussing on quantum information science. Let's consider a quantum state $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ which gives us $|0\rangle$ with probability $|\alpha_0|^2$ and $|1\rangle$ with probability $|\alpha_1|^2$. If measured in the standard basis we would expect the outcome to be $|k\rangle$ (for k = 0,1)with a given probability. This outcome would result in the output state of the measurement gate to also be $|k\rangle$, resulting in the original state $|\psi\rangle$ to be irreversibly lost. We can refer to the process the state undergoes as a collapse of state.

4.4 Weak Measurements

Consider an arbitrary system observable A. Assume a probe $[\hat{q}.\hat{p}] = i$ in a minimum uncertainty state, defined by σ_p^{in},$\bar{p}^{in}$, and $\bar{q}^{in} = 0$. We can then assume a Von Neumann type interaction between the observable of interest and the position of the particle, consequence being the probe receiving a momentum kick when it interacts with the system and that the change in the momentum is exactly equal to the observable to be measured, $\hat{A}$. Formally, we can assume $\hat{H} = \delta(t)\hat{A} \otimes \hat{q}$, so that $\hat{p}^t - \hat{p}^{in} = \hat{A}$ Mathematically, we can define $\hat{A}(p^t) = p^t - p^{in}$ to be the estimate we get for A from measuring probe $\bar{p}^t$. Then for initial system state $|\psi^{in}\rangle$, $E[\psi_{in}|A(p^t)|\psi^{in}] = \langle\psi_{in}|\hat{A}|\psi^{in}\rangle$.
Now consider a final projective measurement on the system too, considering the

sub-ensemble where the system is found in state $|\phi^t\rangle$. Then we can consider the post-selected average $E[A(p^t)|\psi^{in}, \phi^t]$. In the weak measurement limit, $\sigma_p \to \infty$ we get the expectation value for the estimate of the observable A given by:

$$E[A(p^t)|\psi^{in}, \phi^t] \to Re\frac{\langle\phi^y|\hat{A}|\psi^{in}\rangle}{\langle\phi^t|\psi^t\rangle}$$

The amount of disturbance to any system operator due to the coupling of the probe is very small.

4.5 Quantum Circuit Born Machine

Given a dataset $D = \{x_1, x_2..x_n\}$ consisting of n samples and obeys a given distribution χ_d, we would like the QCBM to learn the distribution and generate synthetic data points that are of the distribution χ_s such that χ_s approximates χ_d. The Quantum circuit Born machine is a subclass of parameterised quantum circuits, here the quantum circuit contains parameters which are updated during a training process. The QCBM takes the product state $|0\rangle$ as an input and through an evolution, transforms into a final state $|\phi_0\rangle$ by a sequence of unitary gates. This can then be measured to obtain a sample of bits $x \sim p_\theta(x_s) = |\langle x| |\phi_\theta\rangle|^2$. By training the model we are aiming to let p_θ approach χ_d.
The ansatz for the quantum circuit used by the Born machine consists of 7 layers of 1-qubit gates with entangling layers in between them. These are entangled using the CNOT gates as shown in figure x. The number of wires(qubits) needed depends on the precision required for the generated data. The estimated precision is 12-bit, so the data being able to take 2^{12} different values in the range of $(v_{min} - \epsilon, v_{max} + \epsilon)$, where $\epsilon > 0$ allows data to be generated that lie outside the range (v_{min}, v_{max}) of the original data.
The QCBM takes a $n \times m$ matrix of parameters in the range $(-\pi, \pi)$ as input, in the form of a dictionary. Each angle takes one of 2^k discrete values, where k is a model parameter. The resulting space therefore spans to: $(2^m)^{n \cdot m}$

4.5.1 Barren Plateau

A point of concern when searching for the optimal set of θs is the exploration of the large space, here we may observe issues such as barren plateau(BP). BP insists that the gradient of the parameters of a given PQC will vanish exponentially w.r.t the search space. A formal proof can be found in [barren plateau reference].
I will aim to explore using gradient-based methods as well as alternatives such as genetic algorithms

5 Quantum Architectures

5.1 Butterfly

5.2 Brick

5.3 Pyramid

6 Design

7 Planning

8 Risk Assessment

9 Bibliography

To sort out later:
Born machine: https://link.springer.com/article/10.1007/s42484-022-00063-3
Stateprep1: https://arxiv.org/pdf/2201.11495
Stateprep2: https://arxiv.org/pdf/2108.06150