Application of Deep Reinforcement Learning in Asset Liability Management

By Takura Wekwete, FIA, FASSA

2024 Actuarial Research Conference (ARC)
Middle Tennessee State University (MTSU)

19 July 2024

AGENDA

1. AI in Actuarial Science
2. The Asset Liability Management problem
3. Reinforcement Learning implementation for risk management
4. Results
5. Conclusion

AI IN FINANCIAL RISK APPLICATIONS

1 Supervised Learning

2 Unsupervised Learning

3 Reinforcement Learning

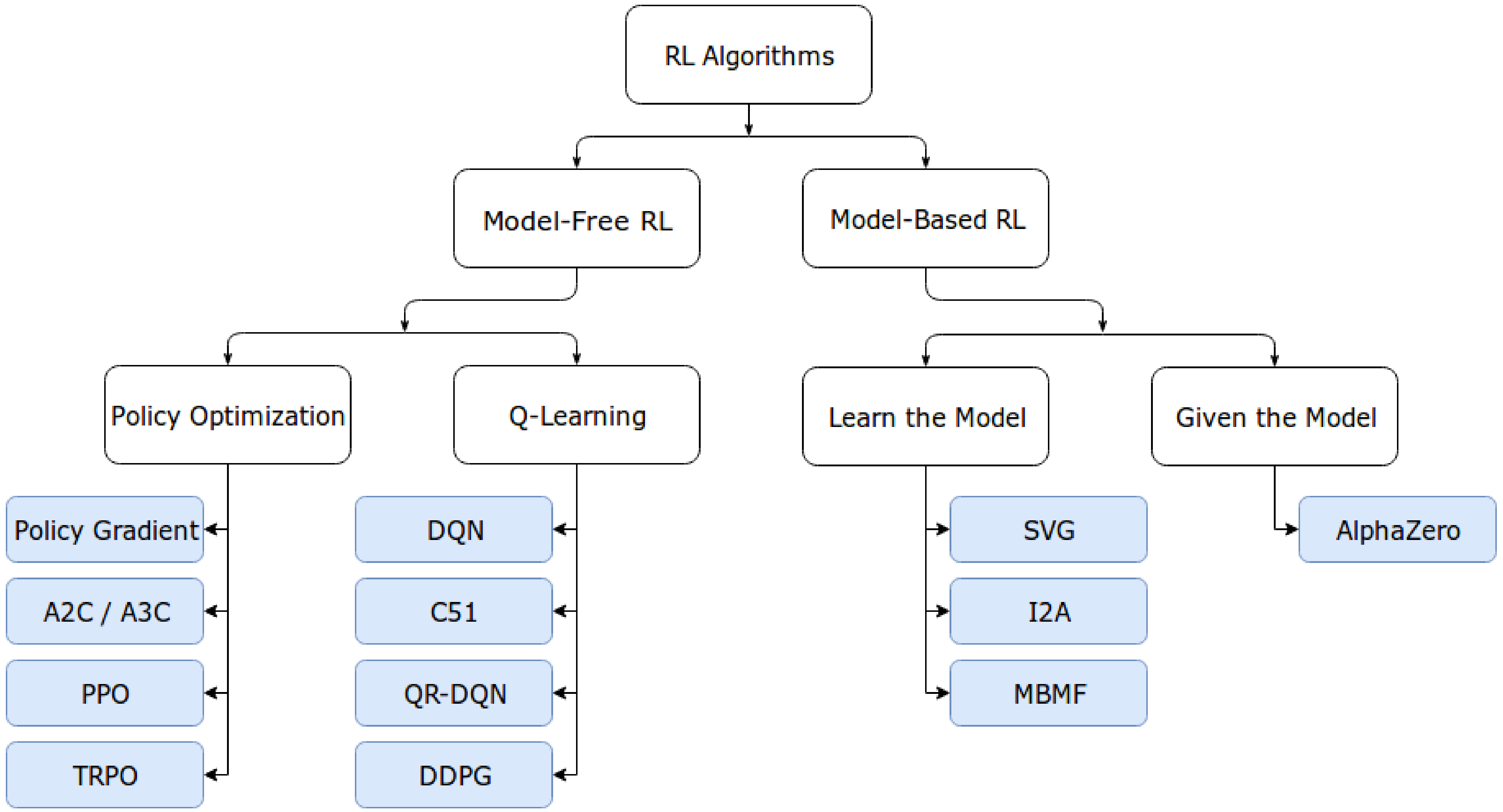
Some common applications in Finance

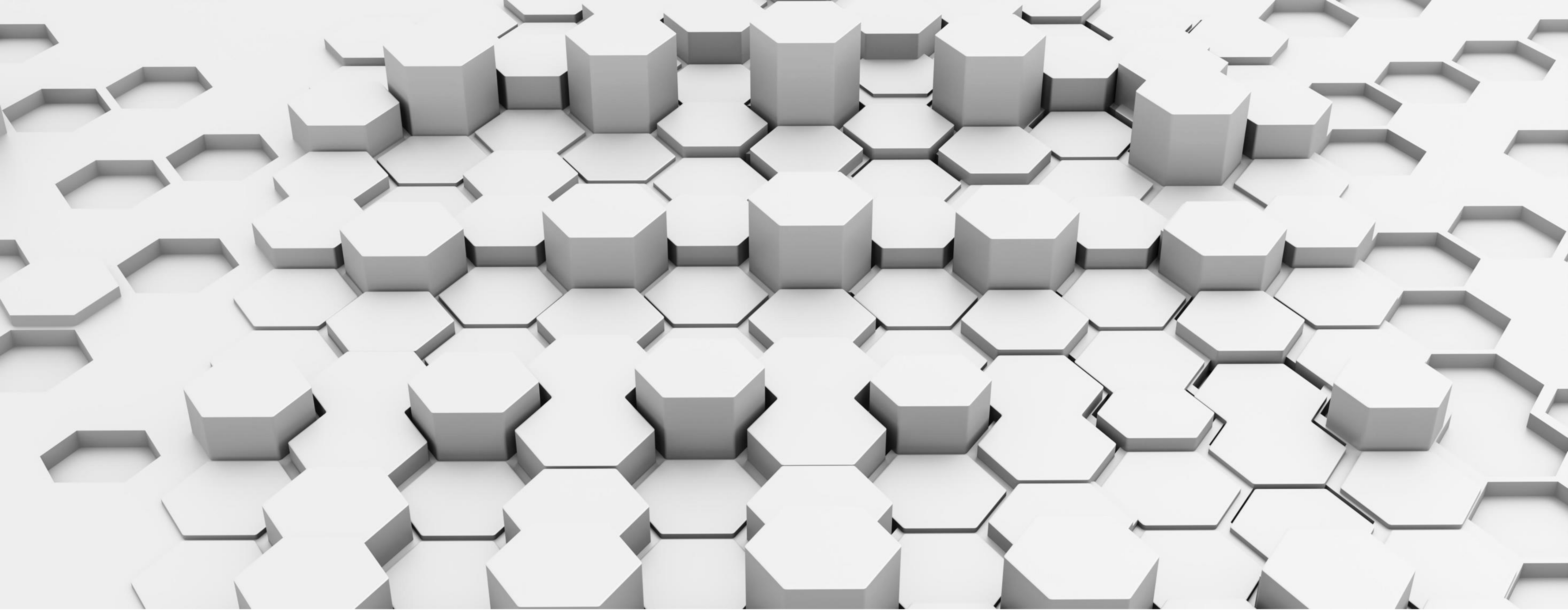
- Risk pricing
- Liability reserving
- Use of telematics data
- Lapse (“churn”) predictions
- Many other applications

Relative strengths of RL

- Problems with little/no data
- Highly dynamic environments
- Problems requiring decision automation
- Factor in user preferences
- Factor in professional expertise

VARIOUS TYPES OF REINFORCEMENT LEARNING





THE ASSET LIABILITY MANAGEMENT PROBLEM

Scan to access paper

Intelligent Systems with Applications

Volume 20, November 2023, 200286

Application of deep reinforcement learning in asset liability management

Takura Asael Wekwete [a 1], Rodwell Kufakunesu [b], Gusti van Zyl [b]

https://www.sciencedirect.com/science/article/pii/S2667305323001114?via%3Dihub

OBJECTIVES OF ASSET LIABILITY MANAGEMENT

Asset Liability Management (ALM) ≈ Liability Driven Investing (LDI)

Primary objectives = allocate assets such that:

1. Asset portfolio value sufficient for obligations
2. Timing of asset cashflows appropriate for obligations
3. Conditions 1) & 2) are maintained

Secondary objectives:

- ❑ Optimising for investment returns
- ❑ Reducing other risks
- ❑ Regulatory compliance
- ❑ Minimising costs

Cash

Bonds/ T-bills

Property

Equities/Shares

Alternatives

CONVENTIONAL APPROACH - REDINGTON IMMUNISATION

Conditions for interest rate risk management:

1. $$A = L$$ where $A = \int_0^\infty A_t e^{-rt} dt$ and $L = \int_0^\infty L_t e^{-rt} dt.$

Sufficient asset value

2. $$\frac{\partial A}{\partial r} = \frac{\partial L}{\partial r} .$$

$$\text{Macaulay Duration} = \frac{\sum_{t=1}^{n} (\text{PV} \times \text{CF}) \times \text{t}}{\text{Market Price of Bond}}$$

$$\text{Modified Duration} = \frac{\text{Macaulay Duration}}{1 + \frac{\text{YTM}}{n}}$$

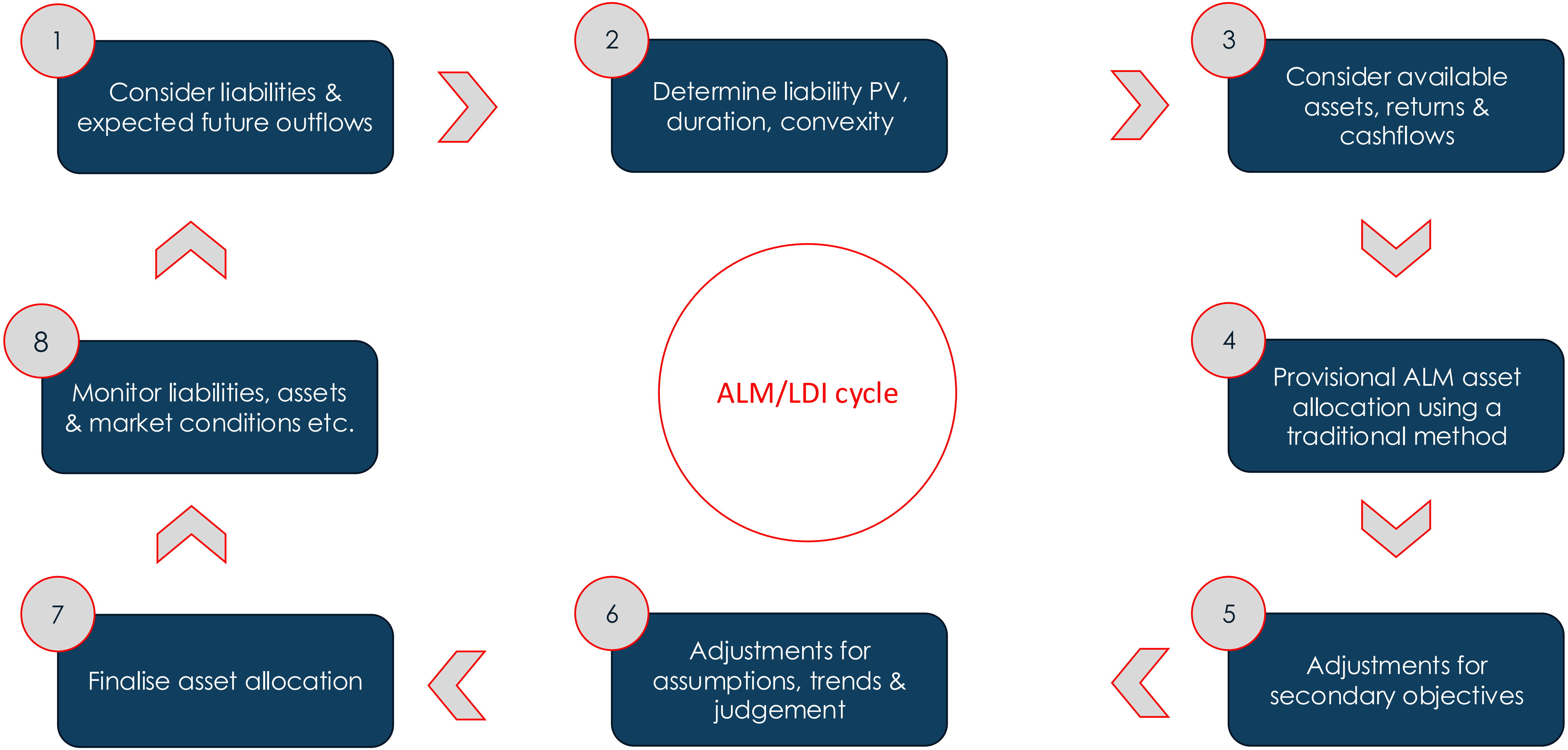
Sufficient asset timing

3. $$\frac{\partial^2 A}{\partial r^2} \geq \frac{\partial^2 L}{\partial r^2} .$$

Convexity

Stability in 1 & 2

TYPICAL CONVENTIONAL ALM IMPLEMENTATION



CONVENTIONAL ALM APPROACHES LIMITATIONS

1 PROCESS LIMITATIONS

- Frequent rebalancing
- Secondary objectives
- Time-consuming

2 THEORETICAL LIMITATIONS

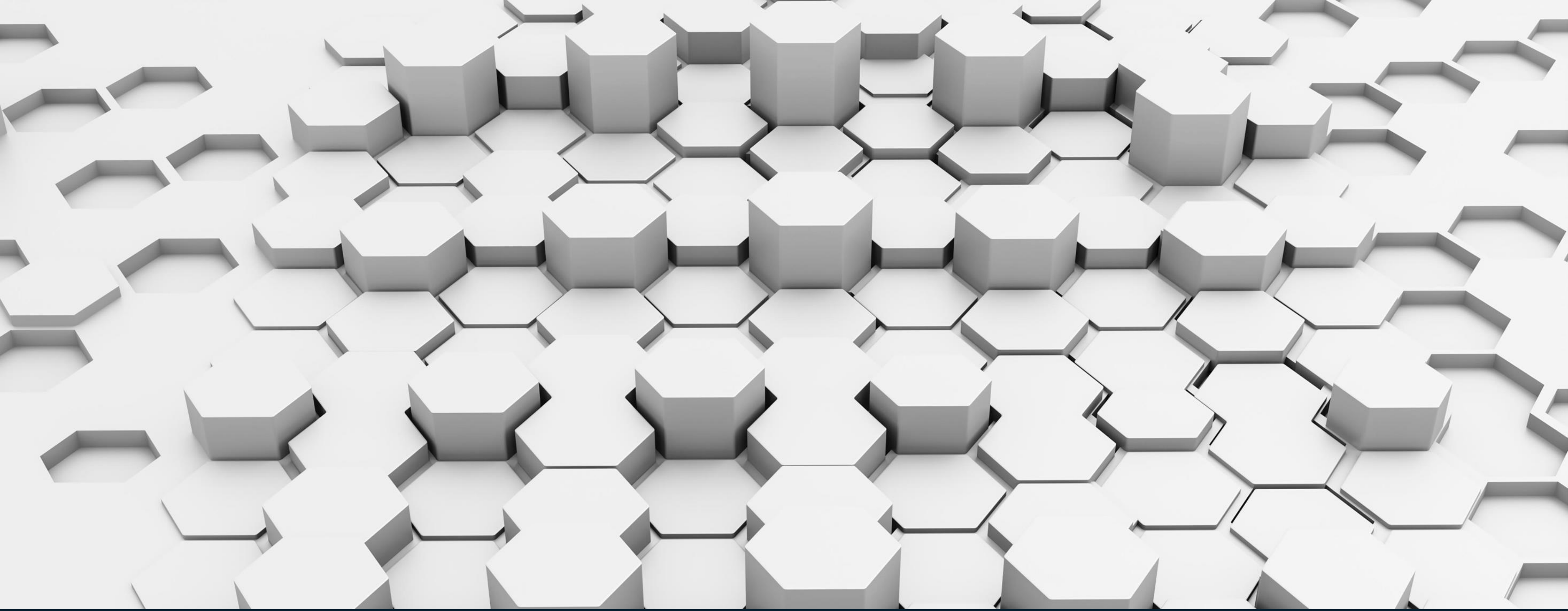
- Assumes interest rate structure
- Assumes parallel shifts
- Unavailable assets ambiguity

3 EXCESSIVE HUMAN DEPENDENCY

- Human error
- Human irrationality
- Biases & emotions

4 GOVERNANCE ISSUES

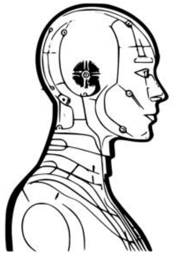
- Governance & incentives
- US Regional banking crisis
- UK LDI crisis



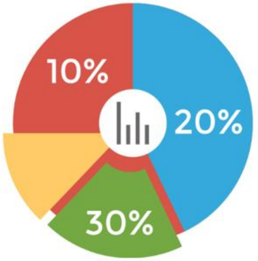
REINFORCEMENT LEARNING SOLUTION TO FINANCIAL RISK MANAGEMENT

REINFORCEMENT LEARNING COMPONENTS

1. **Agent** - The RL decision-making agent



2. **Environment** - Financial institution & investment market



3. **Actions** - Asset allocations

4. **States** – Liability duration, asset duration, PVs, history

5. **Reward function:** Minimise difference btwn timing of asset & liability portfolio

$$e_{it} = \underbrace{\omega_{1it}T(Z_1)_{it} + \omega_{2it}T(Z_2)_{it}}_{\text{Asset portfolio duration}} - D_{it} \ .$$

Liability portfolio duration

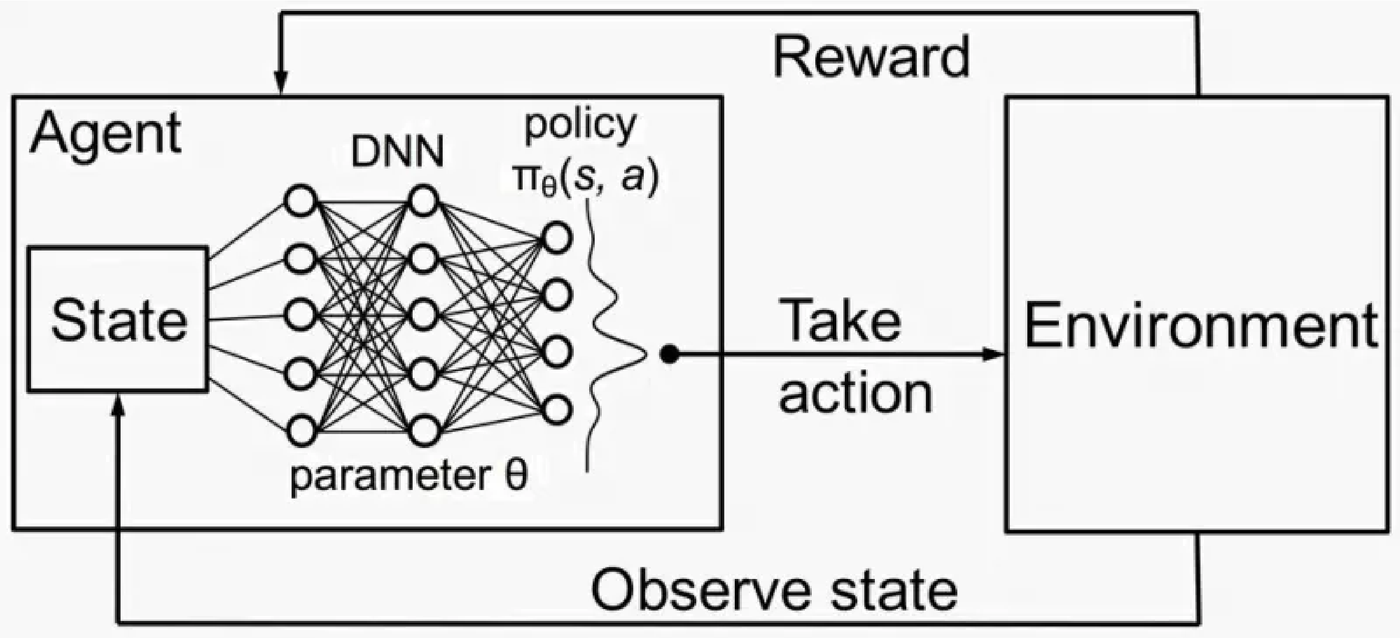
DEEP REINFORCEMENT LEARNING COMPONENTS

AGENT EQUIPPED WITH DEEP NEURAL NETWORK

Experiment & exploit ..

+ … depth of perception

+ … long-term strategy



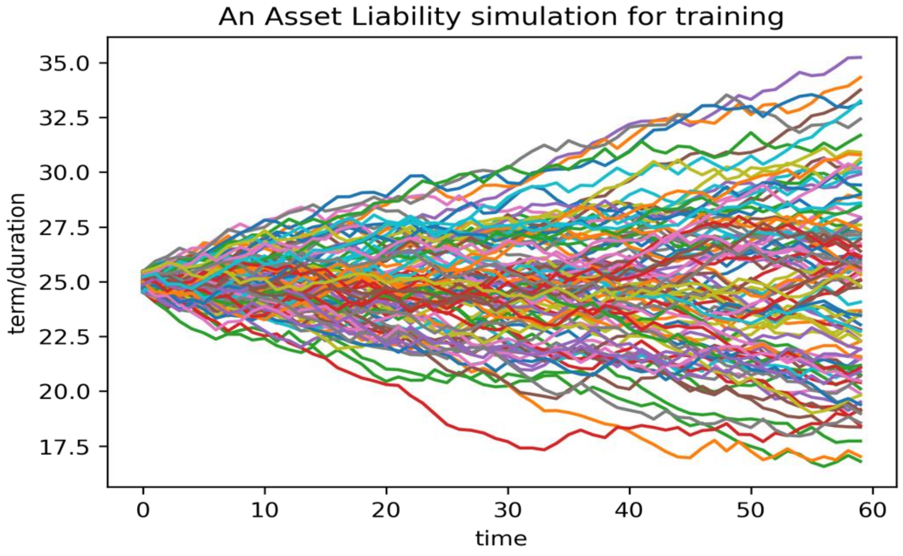
Required because of:

- ❑ Highly dynamic environments
- ❑ Large state spaces
- ❑ Large action spaces
- ❑ Non-linear states-action mapping

In OOP Framework
+ TensorFlow

SIMULATED ENVIRONMENT FOR TRAINING

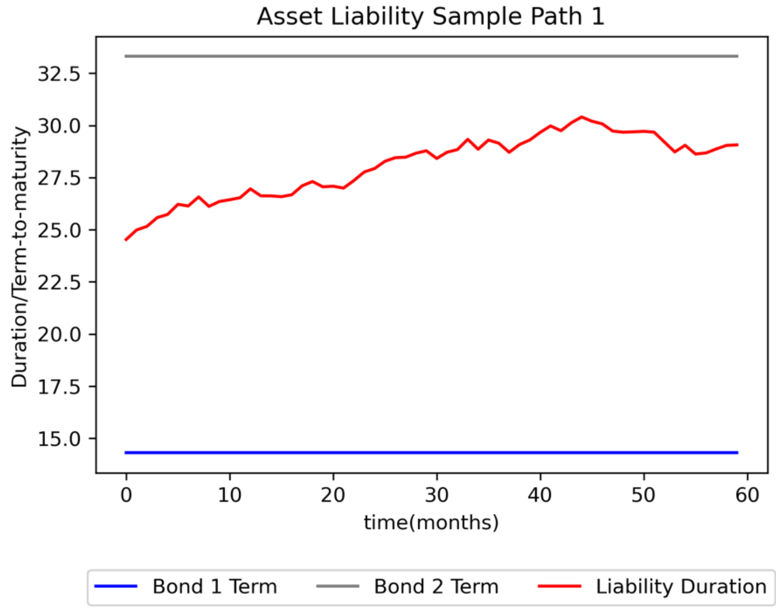
N = 10 000



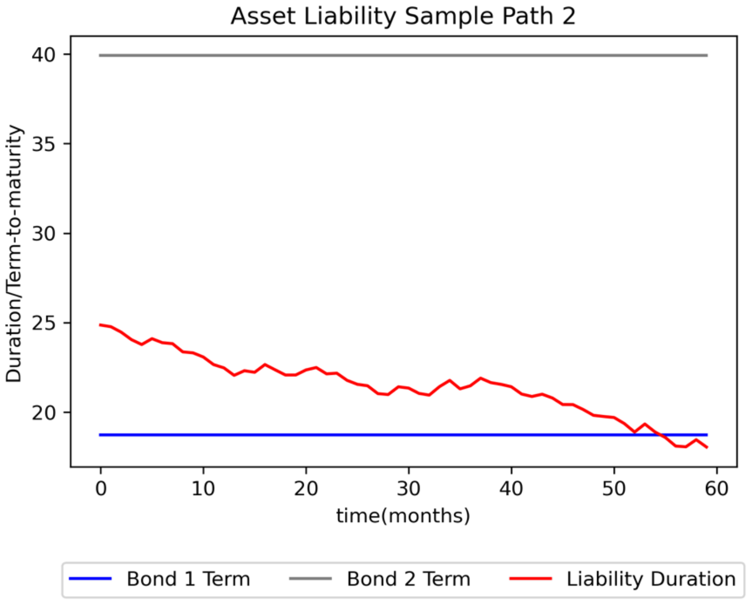
$$D_{it} = D_{i,t-1} + \Gamma_{it} \times \delta_{it}$$

$$T(Z_1)_i \sim U(10, 20)$$

$$T(Z_2)_i \sim U(30, 40)$$



- ❑ Younger/healtheir policyholders
- ❑ Longer-term bank deposits



- ❑ Older/sicker policyholders
- ❑ Shorter-term bank deposits

REINFORCEMENT LEARNING TRAINING PROCESS

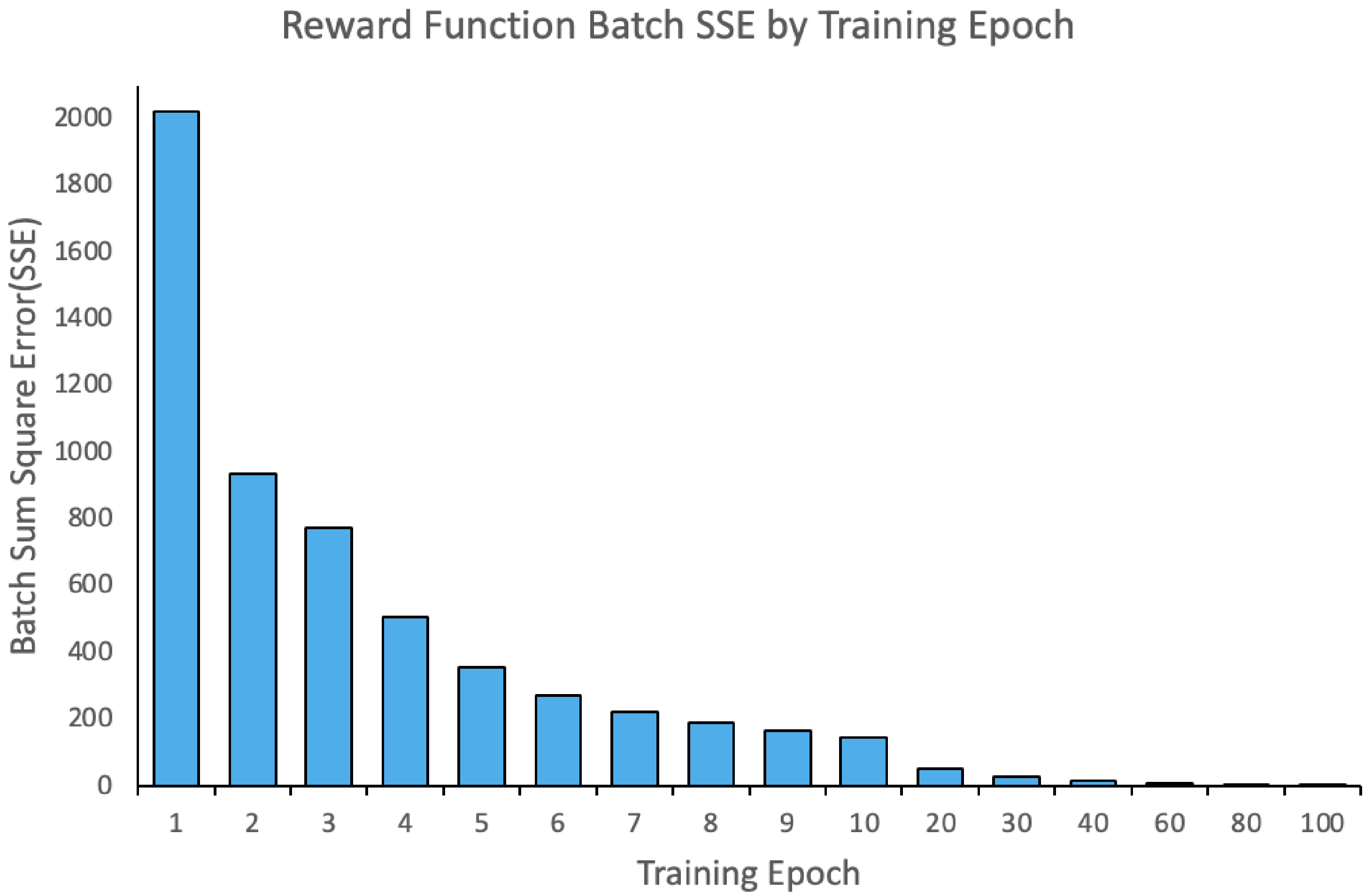
Algorithm 1 Reinforcement Learning for Asset Liability Management

Define the Agent class along with its attributes:

- TensorFlow computational graph
- Neural Network (LSTM-RNN)
- Reward Function

for $epoch = 1, 2, \ldots, k, \ldots, K$ **do**
    for $batch = 1, 2, \ldots, b, \ldots, B$ **do**
        Launch TensorFlow computational graph with data for b
        Apply policy $\pi_{\theta_{old}}$ from previous batch, $b-1$
        Evaluate the rewards at each time t and scenario, e_{it}
        Aggregate batch rewards, $\sum_{i \in Batch} \sum_{t=1}^{T} e_{it}^2$
        Update Agent policy $\pi_{\theta_{new}}$
    end for
end for

REINFORCEMENT LEARNING TRAINING PROCESS



$$e_{it} = \omega_{1it} T(Z_1)_{it} + \omega_{2it} T(Z_2)_{it} - D_{it} \ .$$

RESEARCH METHODOLOGY

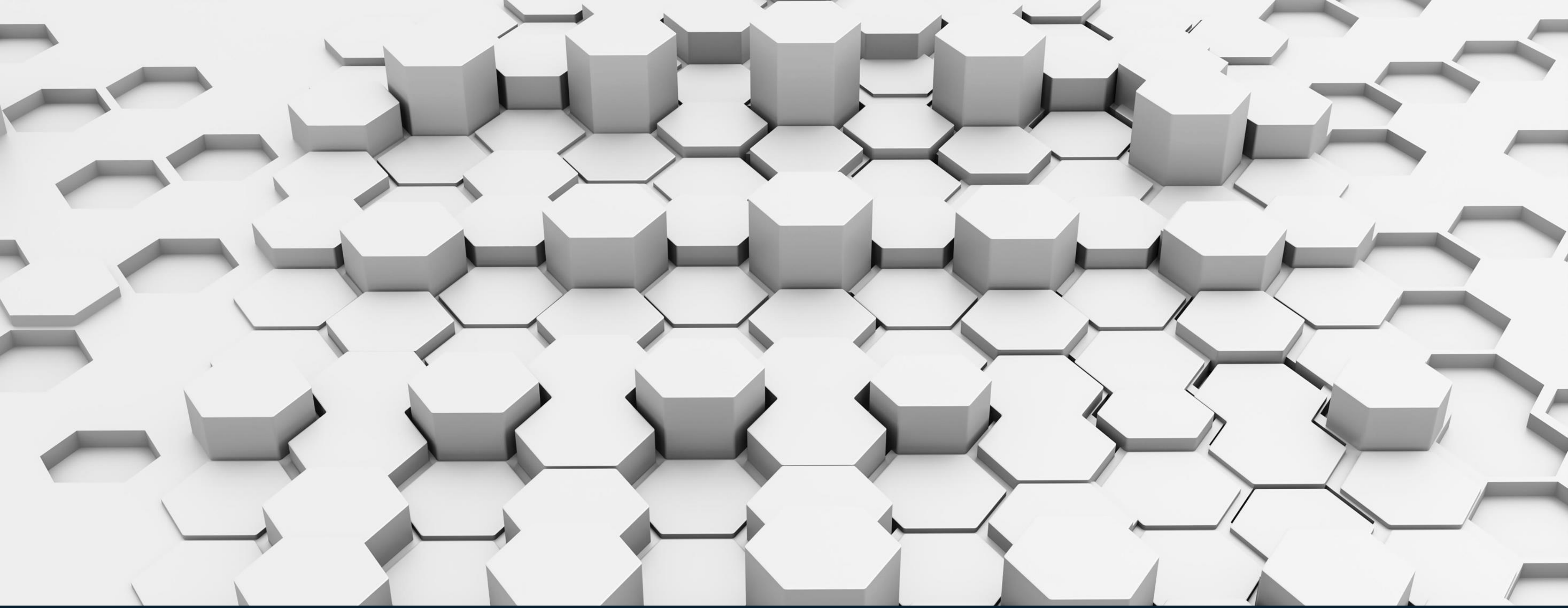
1. Simulate an environment typical of a risk-taking financial institution

2. Define a solution based on conventional methods

3. Define and train the reinforcement learning framework

4. Apply 2. and 3. to new unseen test data

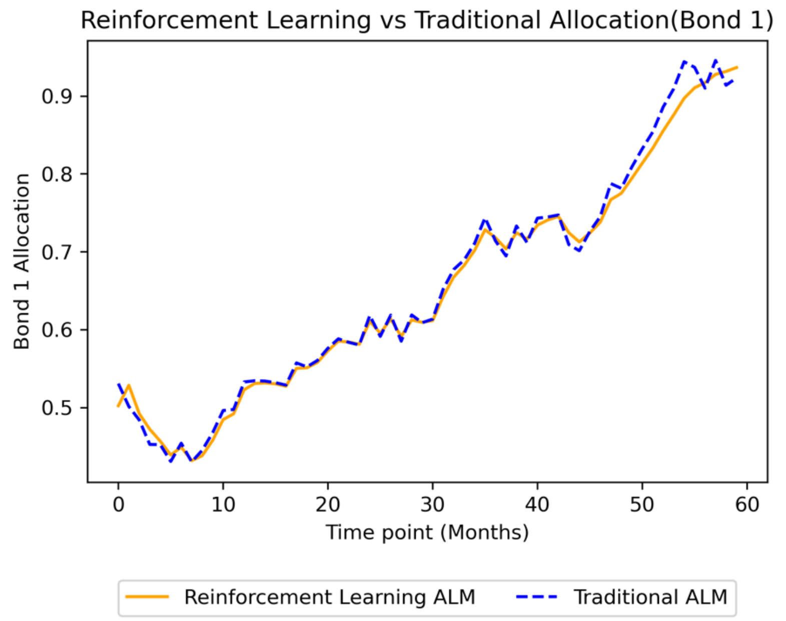
5. Compare results

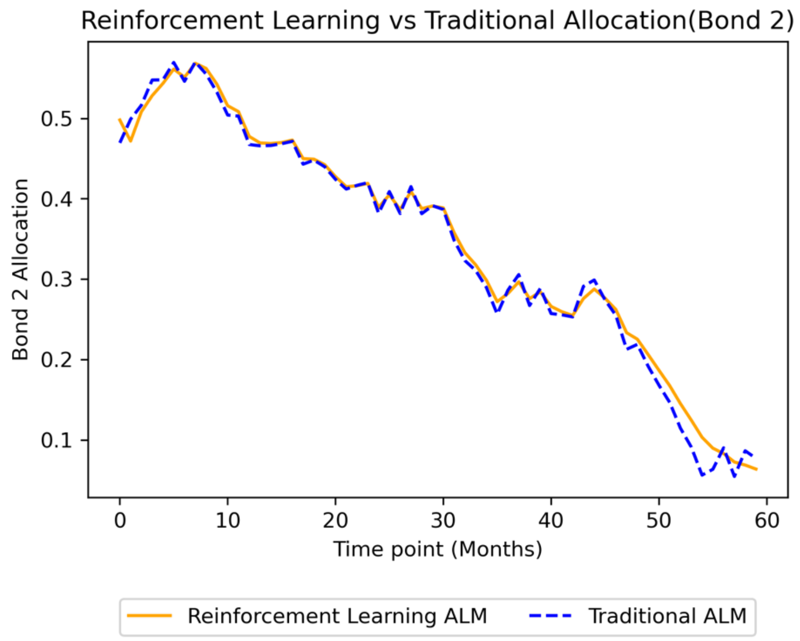


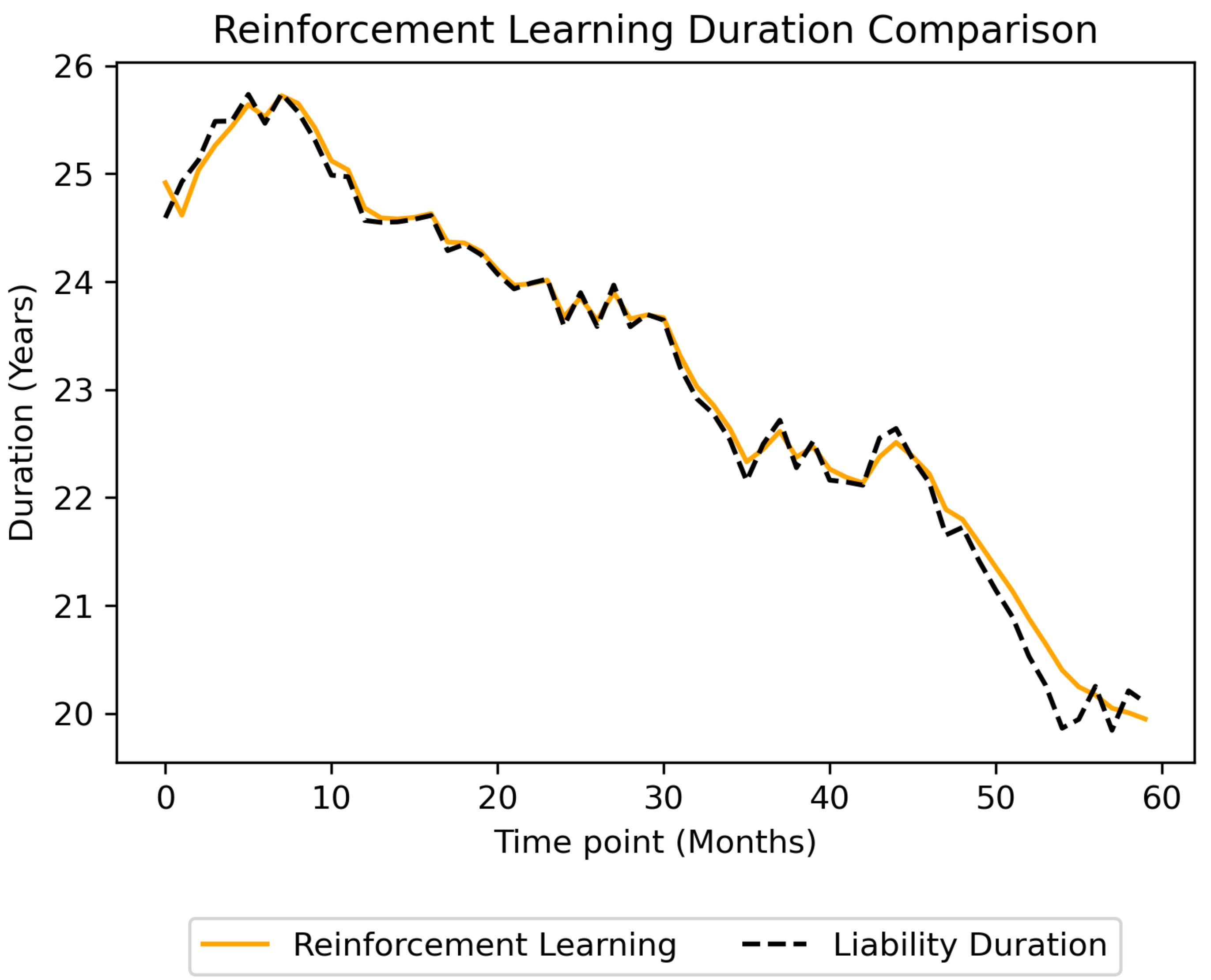
RESULTS

1) DRL PERFORMANCE VS REDINGTON IMMUNISATION

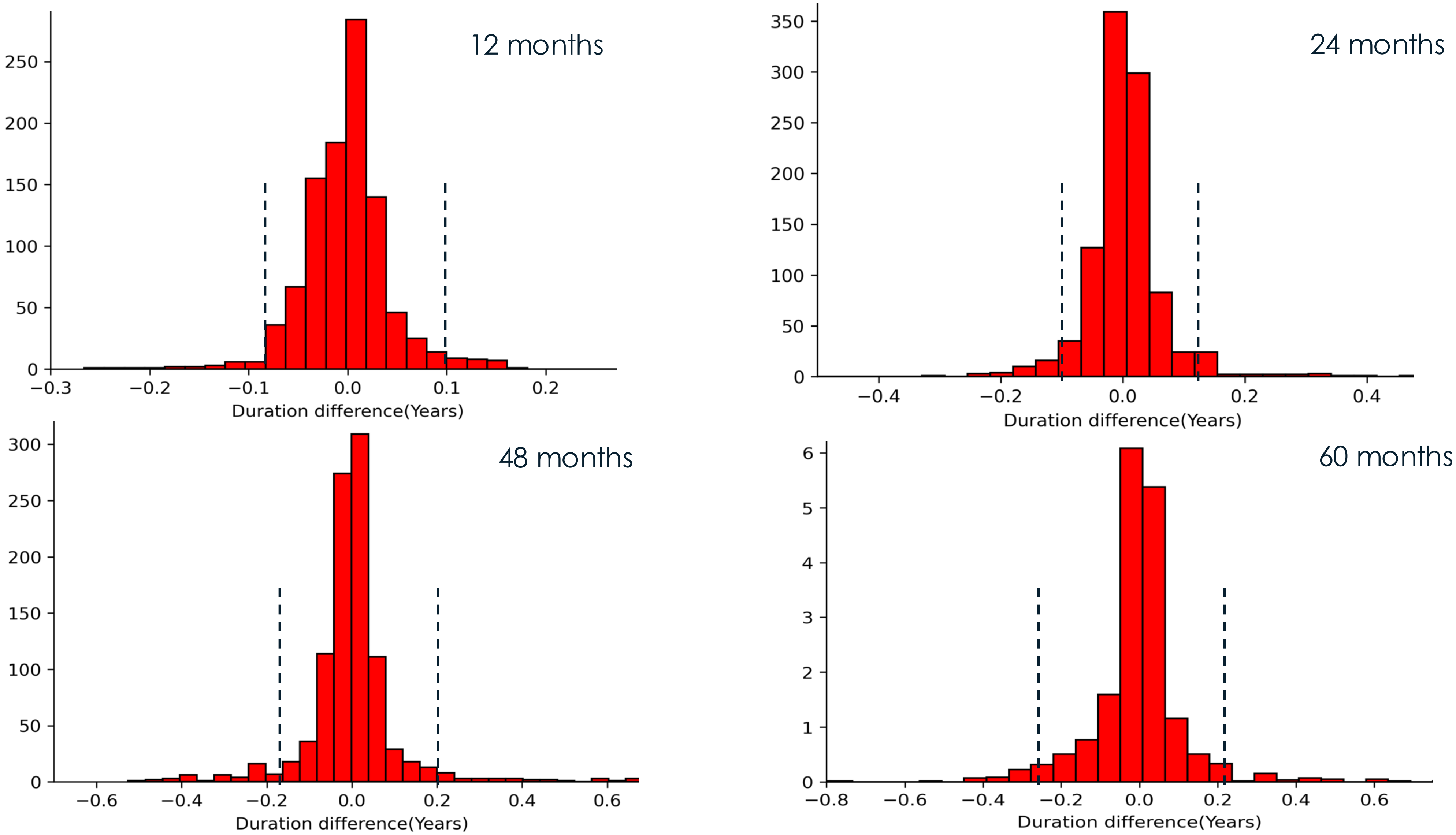
DRL ALM VS CONVENTIONAL ALM EXAMPLE



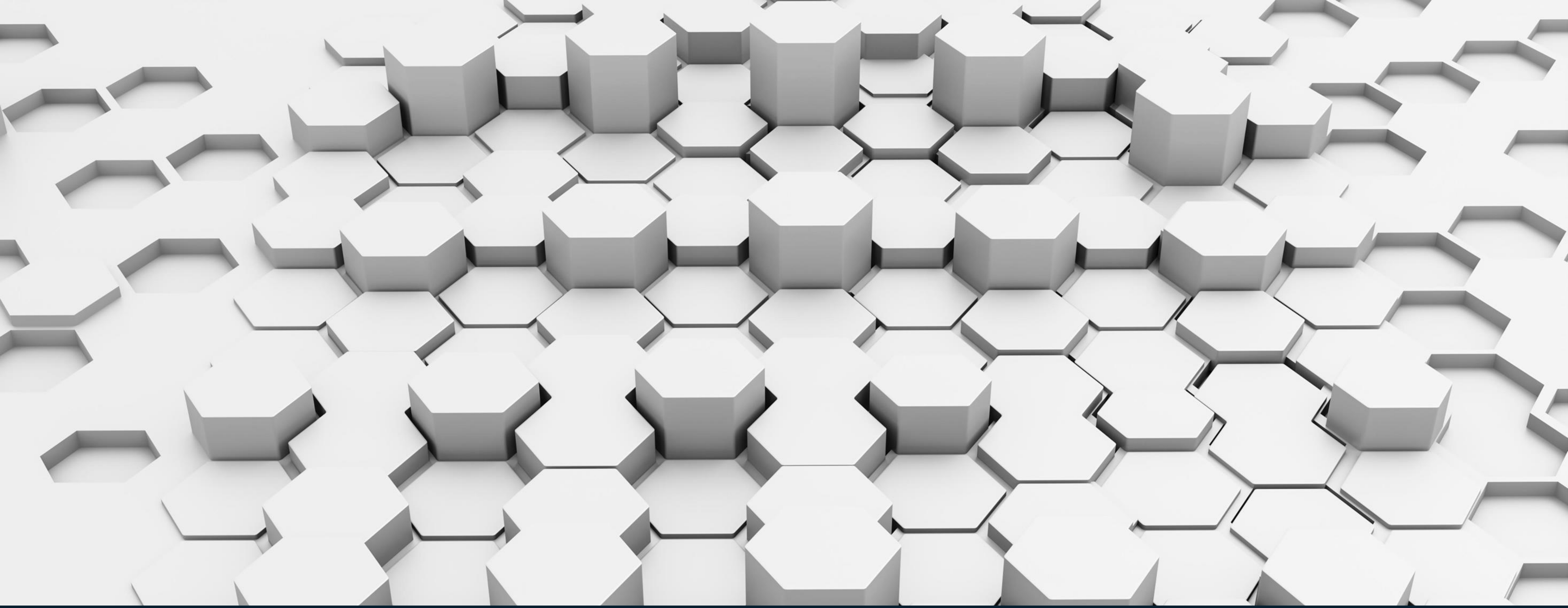




DRL ALM VS CONVENTIONAL AGGREGATED



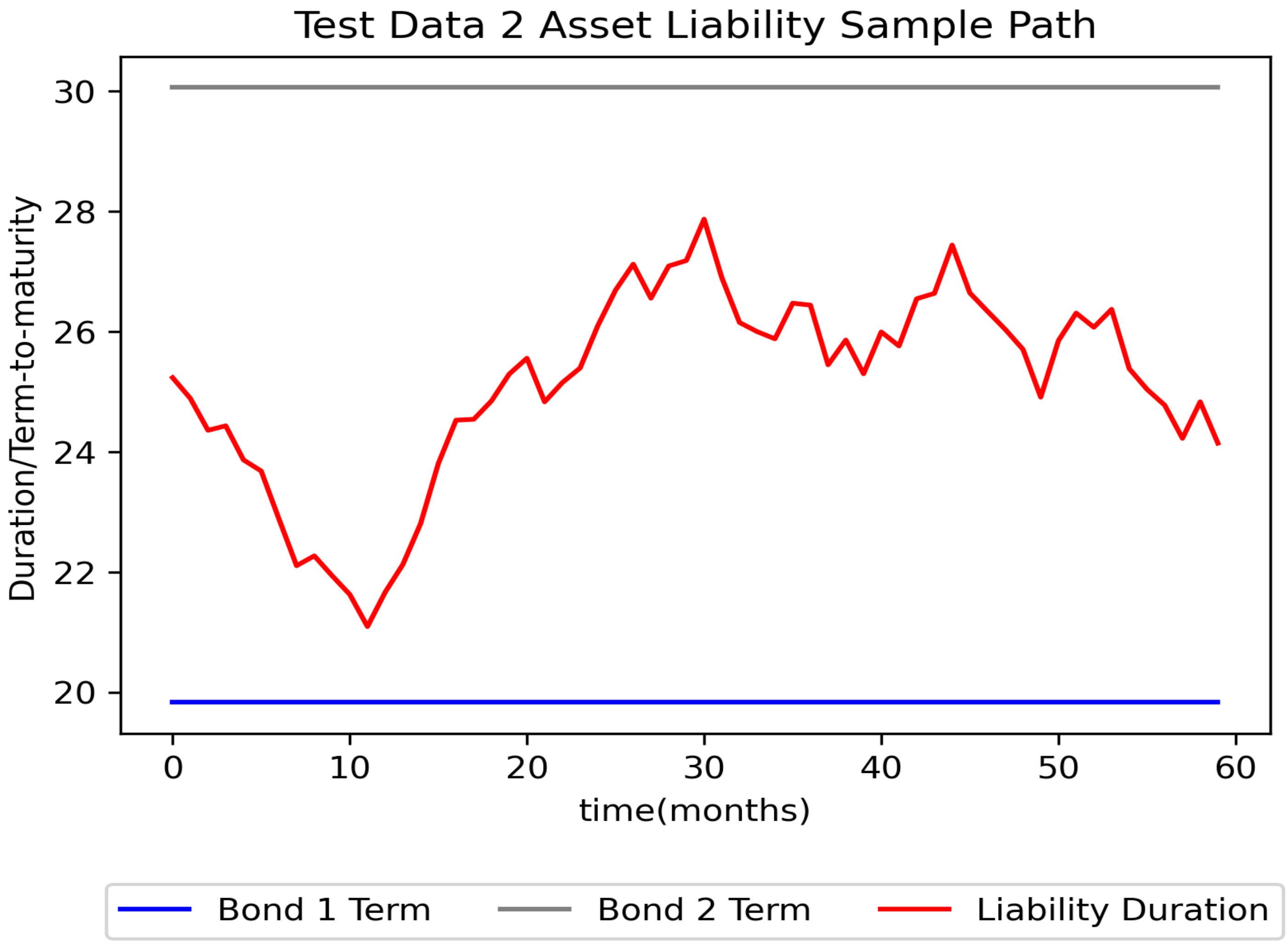
95% of DRL ALM outcomes and Redington immunisation are within 1% of each other



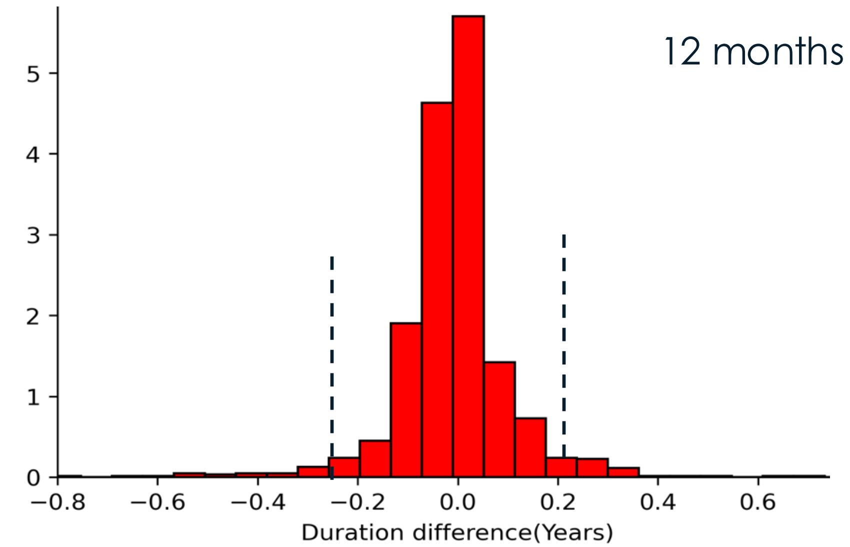
RESULTS

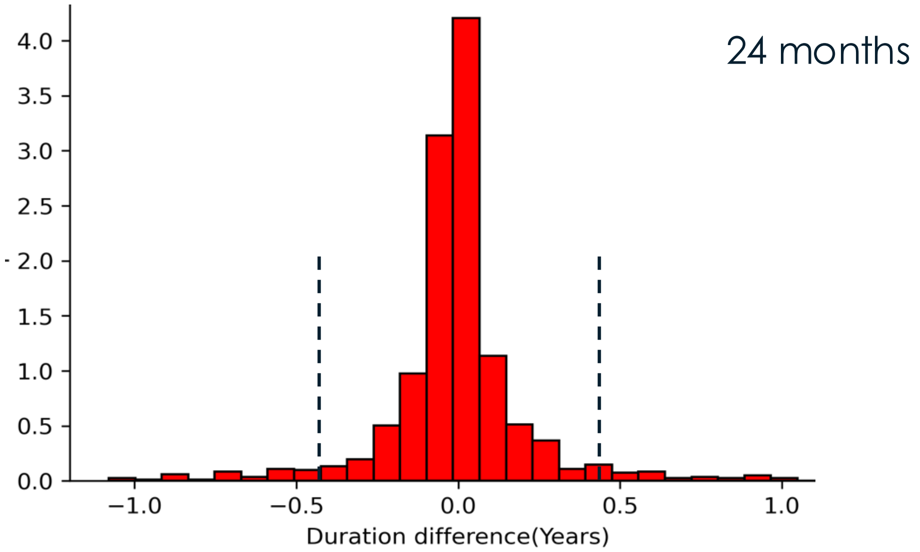
2) DRL ALM STRESS TESTING & ADAPTABILITY

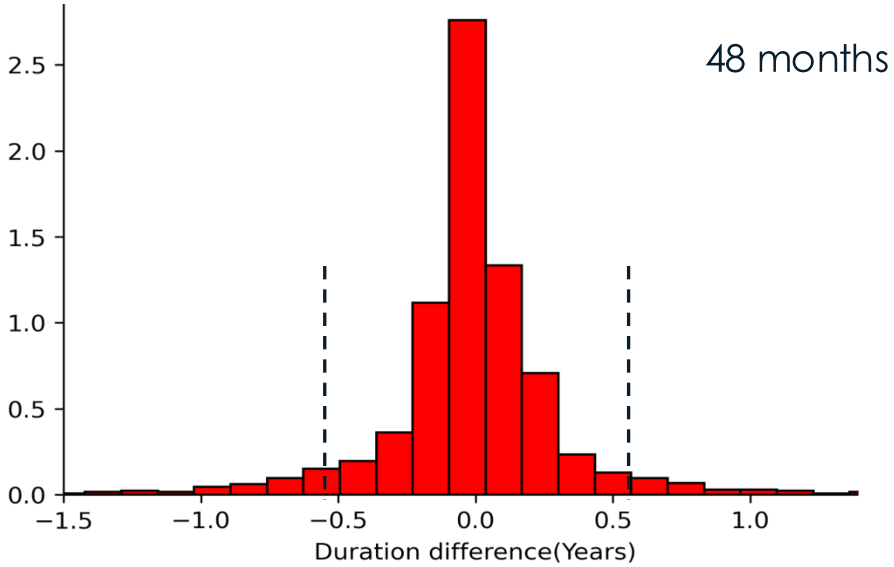
STRESS TESTING SCENARIO EXAMPLE

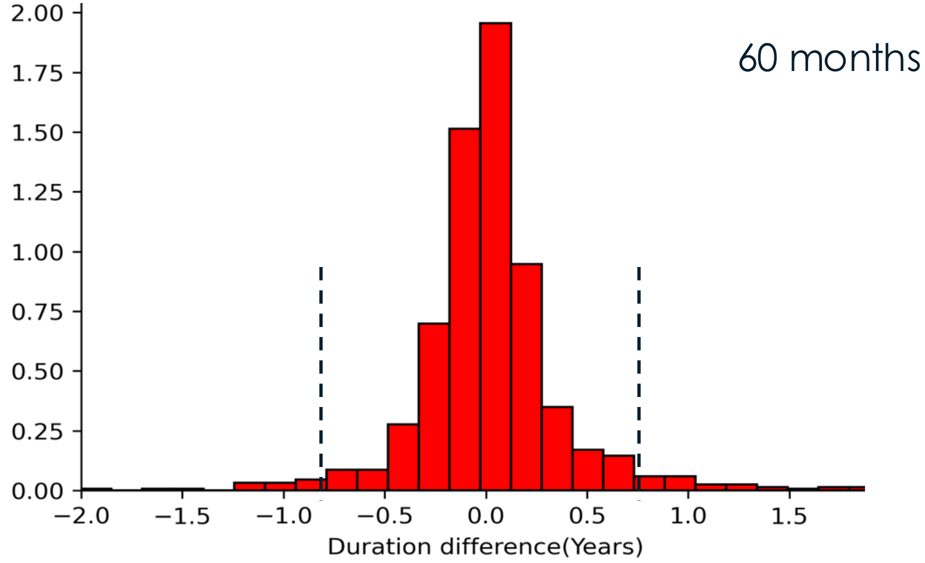


DRL ALM STRESS TESTING RESULTS AGGREGATE

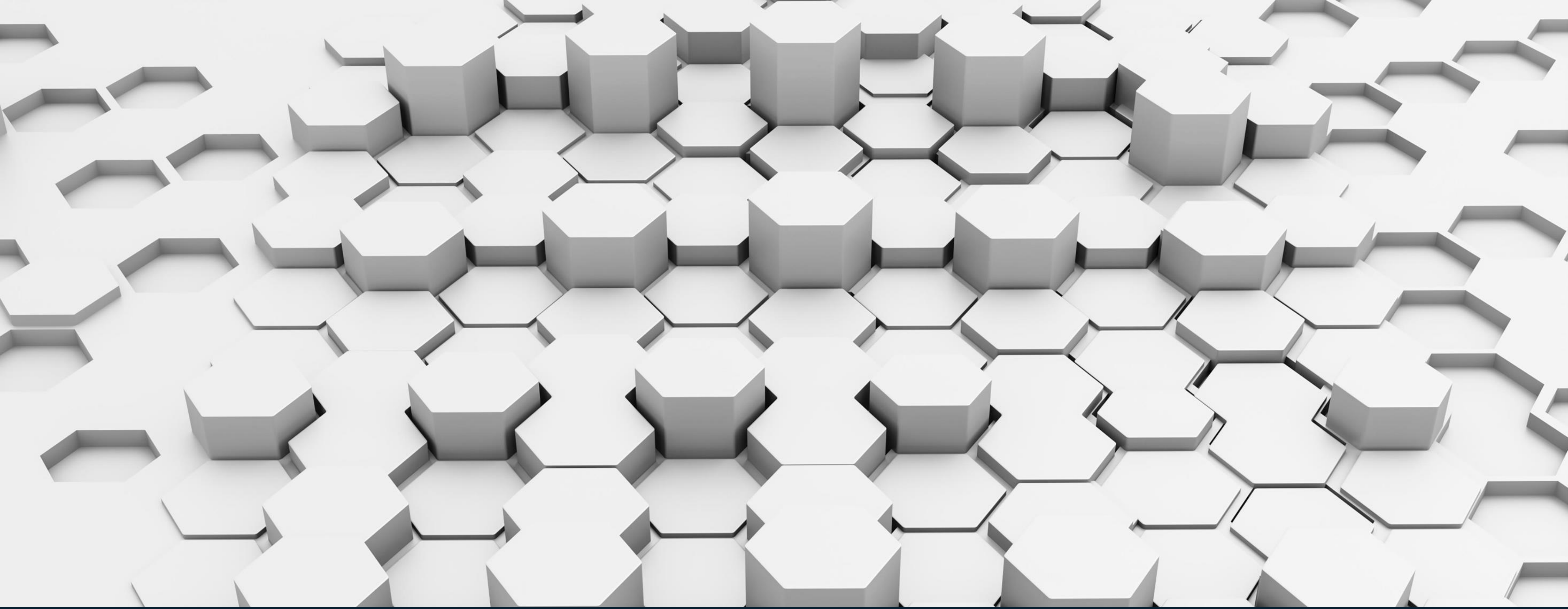








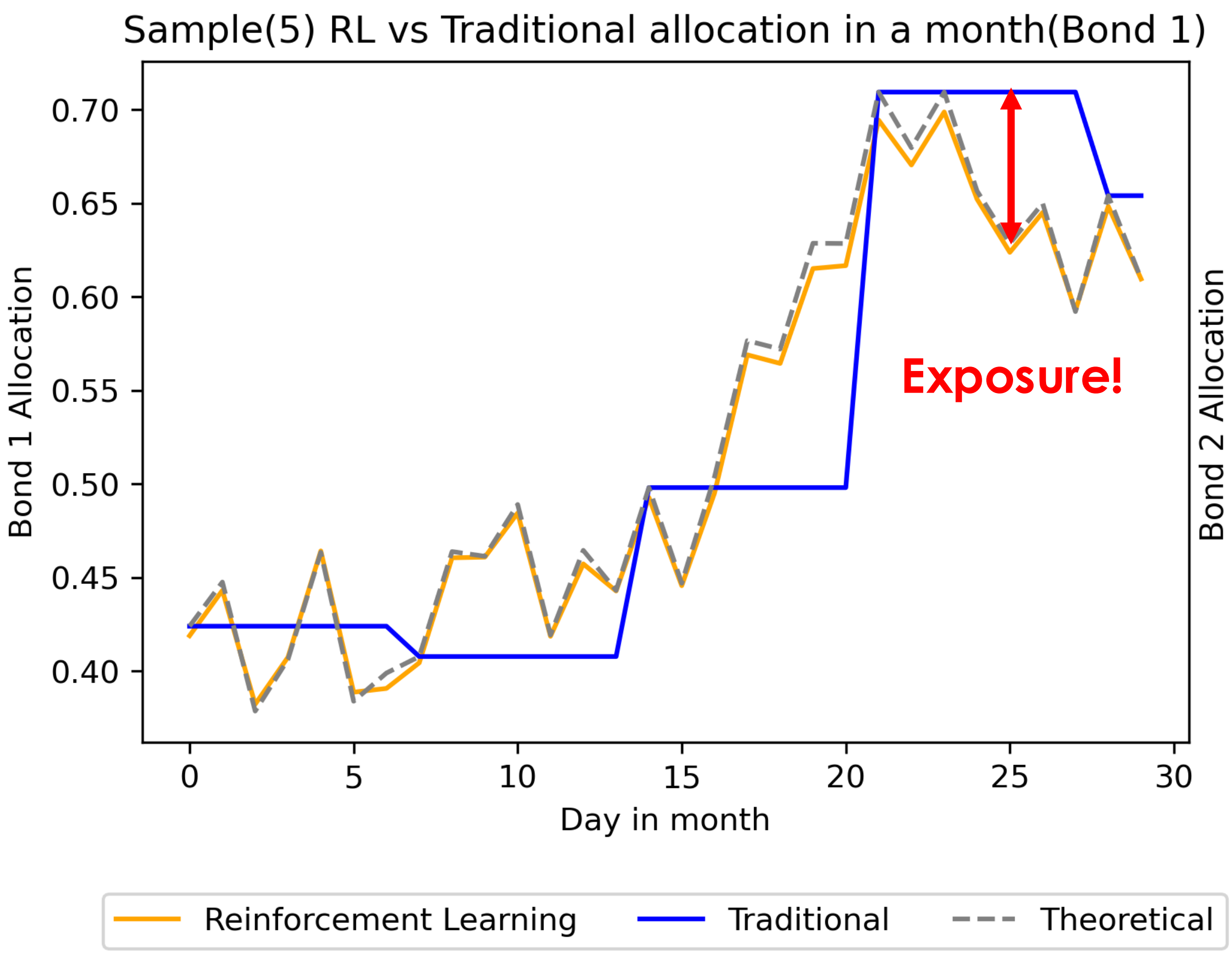
95% of DRL ALM outcomes are within 2% of the appropriate duration outcomes

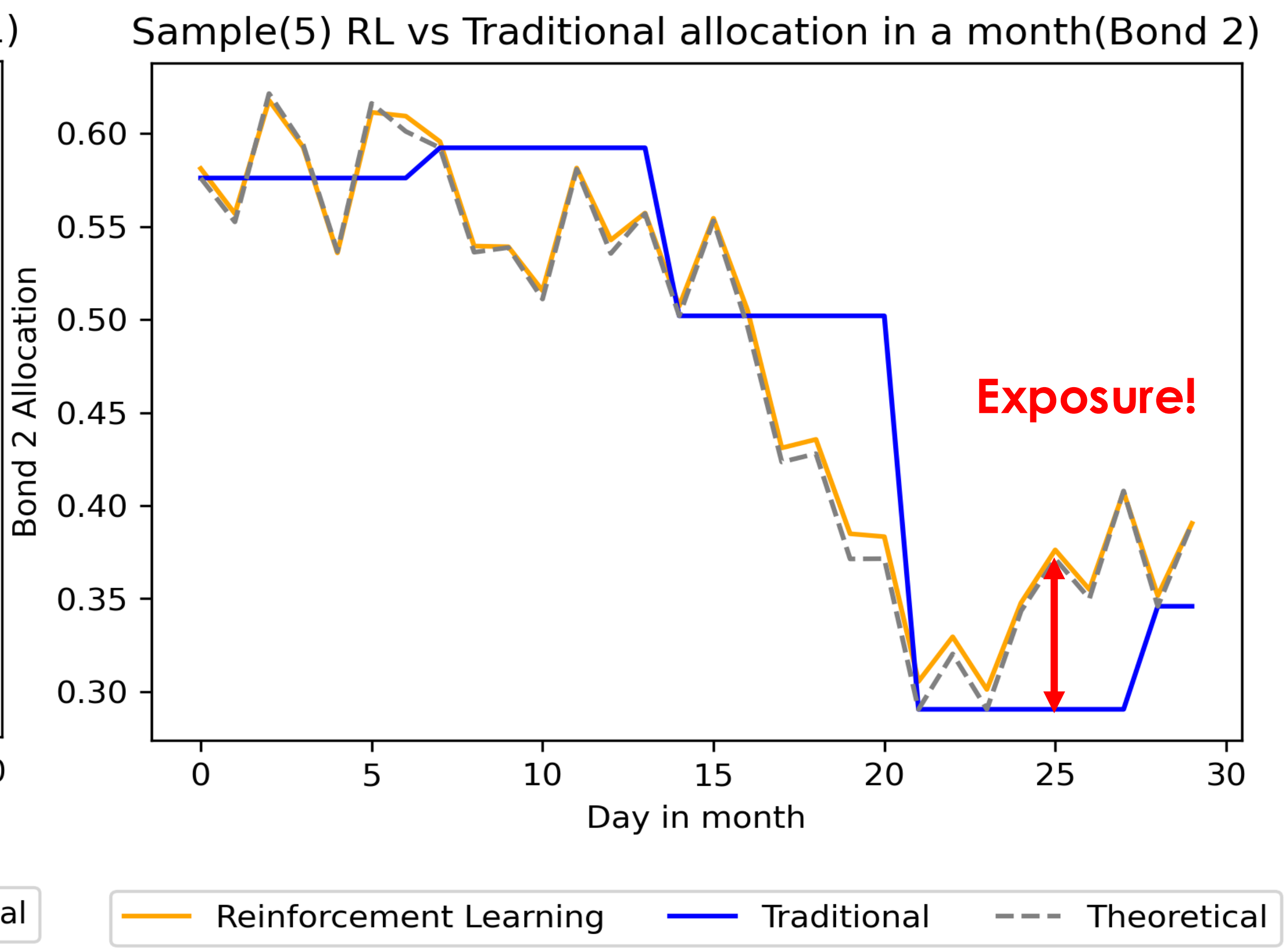


RESULTS

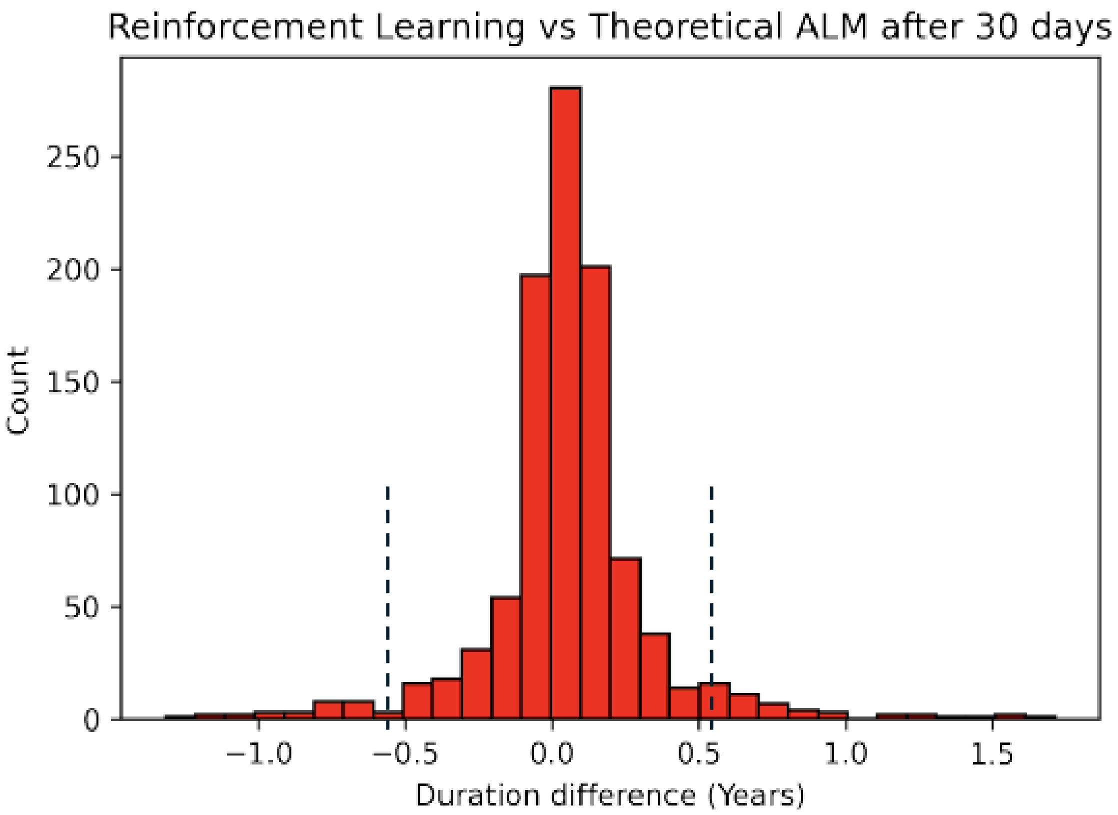
3) COMPARISON TO A BENCHMARK DSTRATEGY

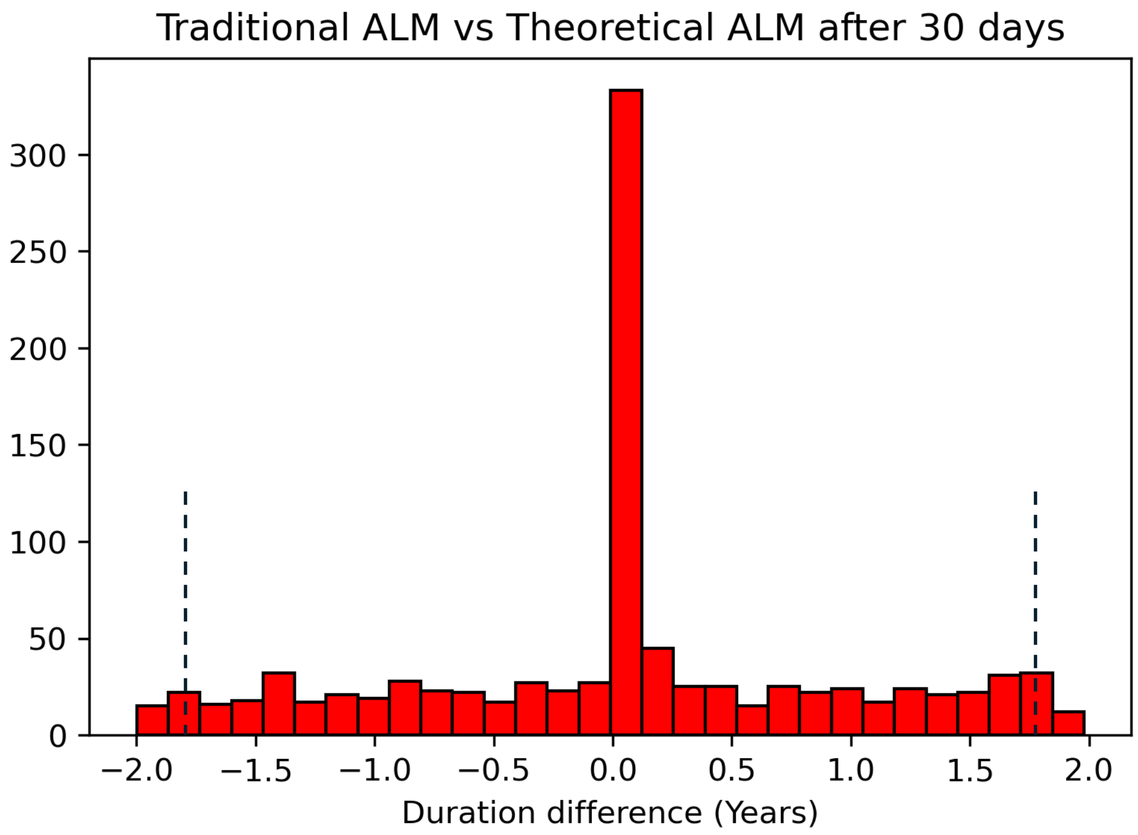
DRL ALM VS BENCHMARK STRATEGY EXAMPLE





DRL ALM VS CONVENTIONAL STRATEGY AGGREGATED





DRL ALM approach had ALM outcomes 3 times less sensitive to interest changes under similar conditions

SUMMARY

1. DRL ALM achieves at least the same level of performance as Redington immunisation under stable conditions

2. DRL ALM is more robust in extreme market conditions

3. DRL ALM significantly out-performs practical traditional strategies

4. Other RL relative strengths
   - Automated & continuously learns
   - Less reliance on theory
   - Interoperable & scalable
   - Multi-objective optimisation

RL USE CASES TO EXPLORE

1 INVESTMENT PORTFOLIO ALLOCATION

2 USER EXPERIENCE & BEHAVIOUR

3 PRICING & UNDERWRITING

4 DISTRIBUTION & CLIENT RETENTION

"..we cannot leave AI only to developers.."

-

Larry Summers

THANK YOU

takurawekwete@gmail.com

https://www.linkedin.com/in/takura-wekwete-asael/