Price and Hedge Discovery of Bermudan Swaptions under Hull-White Dynamics using Q-Learning

Samarth Kadaba

SMBC Capital Markets, New York, NY

August 24, 2023

Problem

Objective: Determine optimal price of (payer) Bermudan swaption using underlying swap trajectories (S_t) and replicating portfolios Π_t with terminal payoffs $\Pi_T = H_T(S_T)$ [1]

1. Preliminaries on Black-Scholes Merton (BSM) model
2. Dynamic Programming (DP) Solution for QLBS
3. Data-driven model-free QLBS
4. Considerations for interest-rate dynamics
5. Methods for pricing/hedging European and Bermudan swaptions

[1] I. Halperin (2019). *QLBS: Q-Learner in the Black-Scholes(-Merton) Worlds*. arXiv: 1712.04609 [q-fin.CP]

Q-Learning uner Hull-White Dynamics

Important Black-Scholes-Merton Assumptions

- Random market movements (Markov assumption)
- Evolves through Brownian motion (known dynamics)

Benefits of Q-Learning

- Provides optimal price **and** hedge actions
- Can be relaxed: 1) non-constant volatility, 2) incorporate dividends, etc. via reward structure
- $Q_{initial}$ (via backward solve) converges in limit to BSM solution

(Simplified) RL Approach via Maximization:

$$\begin{aligned} Q_t^* = \max_{a_t \in A} \quad & \mathbb{E}_t R_t + \gamma \max_{a_{t+1}} Q_{t+1}^* \\ \text{s.t.} \quad & \Pi_T = H_T(S_T) \end{aligned} \tag{1}$$

Q_t^* conditioned on previous state, action value. R_t a function of -Var(Π_t) and current portfolio return.

DP Solution for QLBS (focus of methods herein)

QLBS proposes independent optimization w.r.t. each variable (state, action) of objective - i.e we maximize w.r.t. a_t with (2) then fix optimal actions to solve for Q_t^* with (3), an **on-policy** learner.

$$\min \quad \mathbb{G}_t(a_t) = \sum_{k=1}^{N_{MC}} -(a_t^*(X_t^k)\Delta S_t^k + \gamma\lambda(\hat{\Pi}_{t+1}^k - a_t^*(X_t^k)\Delta\hat{S}_t^k)^2) \quad (2)$$

$$\min \quad \mathbb{F}_t(Q_t) = \sum_{k=1}^{N_{MC}} (R_t(a_t^*) + \gamma \max_{a_{t+1}} Q_{t+1}^*(a_{t+1}) - Q_t^{*,k}(a_t))^2 \quad (3)$$

In practice, consider optimal values a_t^* and Q_t^* parameterized by coefficients of continuous basis functions.

(3) is simply regressing Q table on immediate rewards plus discounted future rewards (Bellman equation).

Solution via Fitted Q-Learning

In this setting, do not consider reward or state dynamics to be known, hence **model-free** and **off-policy**.

Consider Q value updates from single step observations $(X_t^n, a_t^n, R_t^n, X_{t+1}^n)$ From Bellman Optimality, solve LS problem:

$$\min \quad \mathbb{L}_t(Q_t) = \sum_{k=1}^{N_{MC}} \left(R_t(a_t^*) + \gamma \max_{a_{t+1}} Q_{t+1}^*(a_{t+1}) - Q_t^*(a_t) \right)^2 \tag{4}$$

Adaptation to Interest-Rate Derivative Products

Added complexity from interest rate dynamics:

1. Construction of yield curve from market instruments
2. Evolution of term structure via short-rate model
3. Bond pricing for underlying swap value

Added complexity from early exercise:

1. Model as bakset of European swaptions
2. Computation of continuation-values (cost of exercising)

Yield Curve Construction from Market Instruments

Build forward curve from forward-Libor and forward-swap rates (rate periods listed below).

Libor rates	**Swap rates**
1 mo.	1 yr.
3 mo.	2 yr.
6 mo.	3 yr.
12 mo.	5 yr.
	10 yr.
	15 yr.
	20 yr.
	30 yr.

Qantlib API used to obtain short rates from bootstrapped curve which are then used to caculate zero-coupon bond prices.

Short-rate Evolution

Hull-White (HW) model (5) used to model short-rate evolution based on calibration to log-normal (black-implied) swaption volatilities - some negligible error from not using normal vols.

HW models the short-rate dynamics under the **risk-neutral** measure.

$$dr(t) = [\theta(t) - \alpha(t)r(t)]dt + \sigma(t)dW(t) \tag{5}$$

We consider a simplified (1-factor) HW model with constant volatility, mean-reversion.

Maturity (yr.) x Forward time (mo.) Vols (%)

	1	3	6	12	24	60	120
5	34	33.59	34.27	35.93	38.07	34.69	29.67
10	31.6	32.08	32.74	33.77	35.11	32.11	29.84

Pricing Swaps and European Swaptions

Swap valuation formula (6) gives present value of a fixed and floating-leg bonds according to specified payment (tenor) structure.

$$V_{swap}(t) = \sum_{n=0}^{N-1}(P(t, T_n) - P(t, T_{n+1}) - \tau_n k P(t, T_{n+1})) \qquad (6)$$

1. $P(t, T)$ denotes zero coupon bond price at time t with maturity T
2. k denotes fixed rate (value which makes $V_{swap}(0) = 0$)
3. $n = 0...N-1$ denotes the N payment periods on the bond legs

Then, $V_{swaption}(t)$ with expiry T_0 is given by (7)

$$V_{swaption}(t) = \beta(t) E_t(\beta(T_0)^{-1} V_{swap}(T_0)^+) \qquad (7)$$

which is simply the positive part of the swap value at time T_0 deflated by the risk-neutral numeraire.

Priors on Bermudan and Longstaff-Schwartz

We have shown how to price European swaptions via Monte-Carlo using paths of the underlying swap value and (7). Equivalently, we can use DP QLBS (i.e. equations 2 and 3) to achieve the same result as [1] shows.

We will now turn to Bermudan swaptions and consider handling discrete early exercise clauses.

Longstaff-Schwartz Method

- Find optimal stopping time along each path via regression
- Parameterize future expected (Bermudan) values based on prior realizations of all paths
- Continuation value of exercising at time $T - 1$ is discounted bermudan value at T regressed on basis function (here a polynomial of the swap rates)

Basket of Europeans and Longstaff-Schwartz

$$y_p = \frac{B(T_m)}{B(T_{m+1})} V(T_{m+1}, S^p(T_{m+1})) \tag{8}$$

$$x_{pk} = \phi_k(T_m, S^p(T_m)) \tag{9}$$

$$y_p = \sum_{k=1}^{K} \beta_k x_{pk} + \epsilon_p \tag{10}$$

Thus, our *dependent* variables are the discounted Bermudan values, computed backwards in time from the latest-expiring European in our so-called "basket."

We assume some random, independent error ϵ_p to be oberved in our swaption payoffs. Our continuation values then given by...

$$\sum_{k=1}^{K} \beta_k \phi_k(T_m, S^p(T_m)) \tag{11}$$

Pricing in Longstaff Schwartz

The price of a bermudan at time T is the maximum of its intrinsic value (discounted payoffs) and computed continutation value at T.

This is expressed in (12) below. [2]

$$P(T) = max(P(T), C(T)) \tag{12}$$

Where $P(T)$ denotes the time T intrinsic value of the bermudan (i.e. the equivalent European price).

$C(T)$ denotes the time T continuation value (i.e. a backward propogated price with added premium for the option to exercise).

[2] L. B. Andersen (1999). "A simple approach to the pricing of Bermudan swaptions in the multi-factor libor market model". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.155208

Results

Summary: QLBS accurately prices exotic interest rate derivative products while providing efficient hedging on the underlying instruments.

1. Recovering zero coupon bond (ZCB) prices
2. Replicating forward swap values
3. Pricing European swaptions via Monte Carlo
4. Pricing and hedging via QLBS

Metrics for Performance Evaluation

We consider primarily the relative error associated with Monte Carlo (path-dependent) and analytical forms of pricing.

- ZCB prices (discussed in following slides) compare analytical formula with numerical approximation
- Replicating swaps are constructed via Quantlib APIs and Black-Scholes formulas are used to price these instruments which is then compared with our simulated, path-derived prices.
- Swaptions are prices on the Jamishidan engine and compared with our QLBS results
- Tree based engines are used to price replicating, synthetic Bermudan swaptions which mimic our own instrument construction

Errors are quoted in percents and sometimes basis points.

Recovering ZCB Prices

To validate short-rate evolution we evaluate (13) which is the generic bond pricing formula.

$$P(0, T) = exp(-\int_0^T r(\tau)d\tau) \tag{13}$$

Rel. Error - QuantLib (bps)

	1	3	6	12	24	60	120
5	0.00	0.08	0.09	0.09	-0.48	-0.35	2.09
10	0.00	0.09	0.10	0.10	-0.46	-0.94	-1.75

Recovering Swap Prices

To validate swap pricing, we replicate vanilla instruments through QuantLib libraries and price under equivalent terms/conditions.

We consider 5 year and 10 year swaps below with varying forward times.

Rel. Error - QuantLib (%)

	1	3	6	12	24	60	120
5	-0.19	0.15	-0.32	0.07	-0.3	0.087	-0.33
10	0.13	00.16	0.31	0.25	0.35	0.80	-0.5

Recovering Swaption Prices via Monte Carlo

We compare Monte Carlo derived European swaption prices with those derived analytically from Quantlib (which uses Jamishidan swaption pricing).

We consider options into 5 and 10 year swaps with varying expirys (chosen based on which instruments were most liquid).

Rel. Error - QuantLib (%)

	1	3	6	12	24	60	120
5	-6.78	-10.18	-5.66	-10.19	-5.19	-8.75	-2.27
10	-8.17	-3.04	-7.91	-4.08	-6.73	-4.13	-7.61

Pricing European and Bermudan Swaptions using QLBS

Rel. Error - Euro (%)

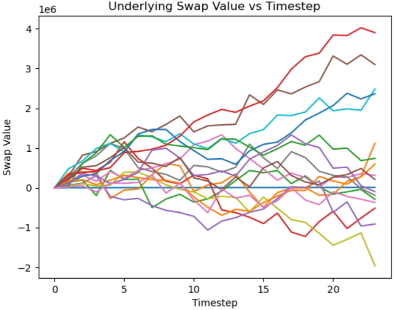
	1	3	6	12	24	60	120
5	-13.2	-12.9	-10.0	-0.6	-7.3	-2.0	3.6
10	-15.1	-15.3	-14.3	-10.3	-11.0	-5.9	N/A

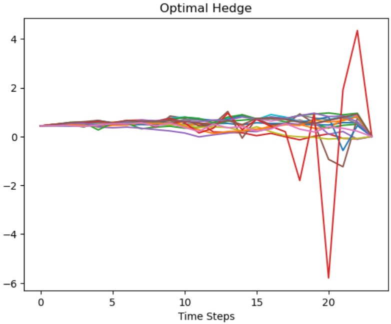
Rel. Error - Berm (%)

	1	3	6	12	24	60	120
5	9.5	-6.3	-15.7	-14.1	-11.5	-6.7	-1.9
10	-1.65	-2.01	-15.69	-22.5	-20.8	-16.3	N/A

Trajectory Visualization I - 1M x 5y

Underlying Trajectories and Optimal Hedging

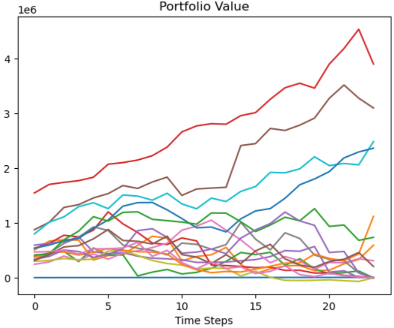


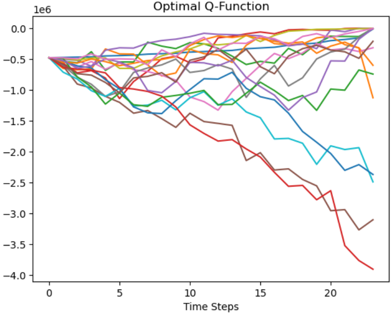


Swap value deflated by money-market account (risk-neutral numeraire) is a martingale.

Optimization Visualization I - 1M x 5y

Portfolio Values and Q-Values

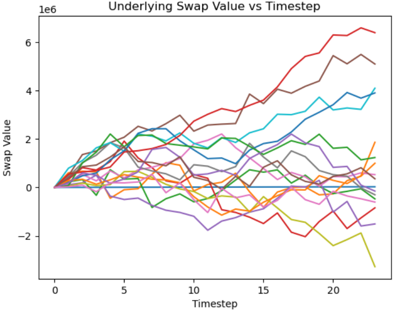


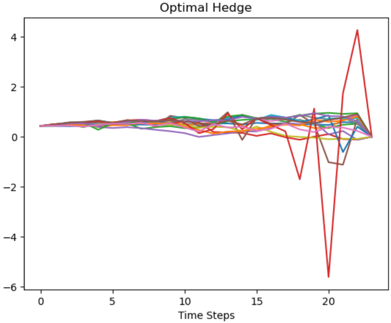


Observe expected increase in portfolio value and near-monotonic decrease in -Q value (increase +Q).

Trajectory Visualization II - 1M x 10y

Underlying Trajectories and Optimal Hedging

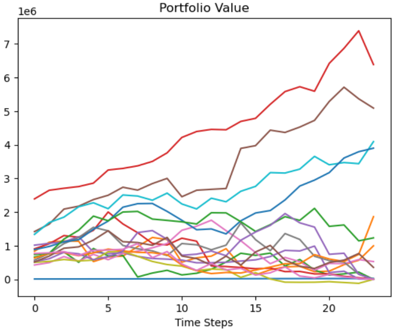


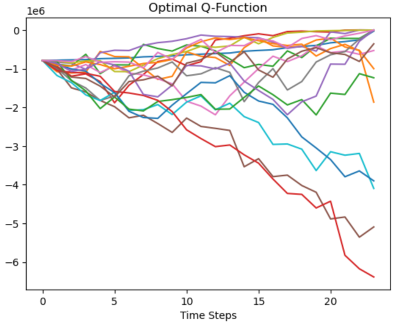


In real-world settings, transaction costs would enforce smooth hedging which can be achieved with certain regularization on optimal actions.

Optimization Visualization II - 1M x 10y

Portfolio Values and Q-Values





Hedge Performance and Efficiency I - 1M x 5y

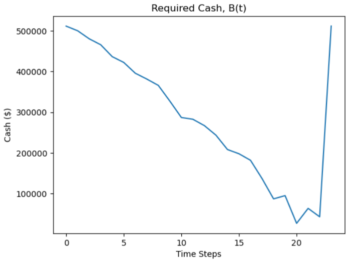


Figure: On notional of $100M, we require an initial bank account of $500K for optimal hedging.

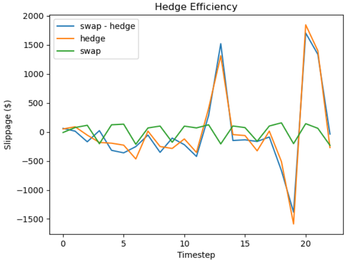


Figure: Our total slippage, i.e. unprotected ΔV_{swap} is $\$9,822.61$ or $427.07 per timestep[3].

3 J. H. Hoencamp, S. Jain, and D. Kandhai (2022). "A semi-static replication approach to efficient hedging and pricing of callable IR derivatives". In

Hedge Performance and Efficiency II - 1M x 10y

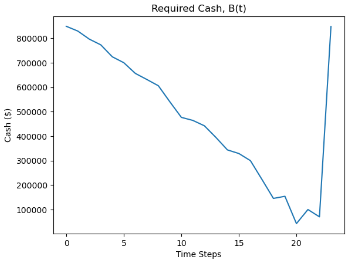


Figure: On notional of \$100M, we require an initial bank account of \$800K for optimal hedging.

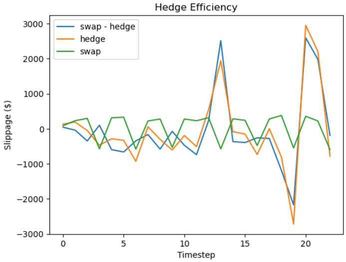


Figure: Our total slippage, i.e. unprotected ΔV_{swap} is $\$16,334.83$ or \$710.21 per timestep.

Hedge Performance and Efficiency III

Another look at Hedge **Effectiveness**

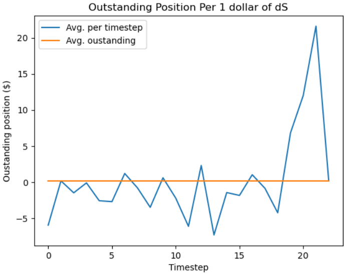


Figure: 1M x 5y outstanding per 1$ change in underlying (ΔS).

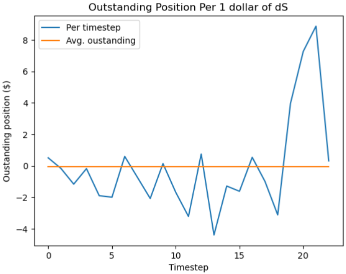


Figure: 1M x 10y outstanding per 1$ change in underlying (ΔS).

We assume risk on a per-timestep basis however our average net-position per dollar change in the underlying is approx. \$0.2.

Sparse First-Differences with L1 Regularization

Model tuning can help priortize specific trading objectives We add weighted L1 regularization to a_t^*, and vectorize w.r.t. time.

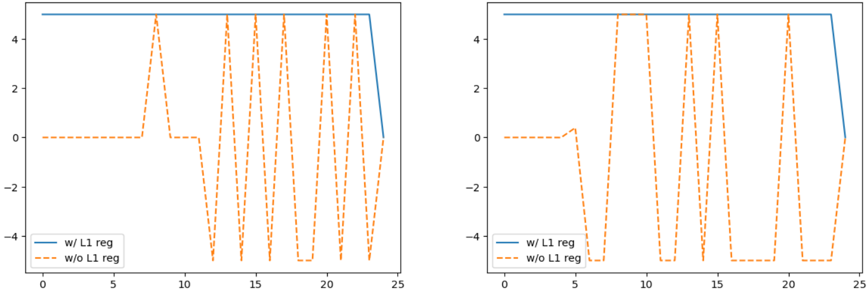


Figure: L1 regularization yields minimal optimality gap with sparse hedge first differences; shown for coefficients related to two features.

Implementation Details

Some parameters and assumptions were made to normalize experimental results.

- Fixed and floating-leg payment schedules identical
- Quarterly payments
- $100M Notional
- Hedge re-balancing across 24 timesteps
- 8192 Sobol (low-discrepancy) sequences were used to simulate short-rate paths (computation time around 30 secs)
- 3% risk-free rate
- All swaptions were evaluated ATM
- Purely risk-based hedging

Bibliography

Andersen, L. B. (1999). "A simple approach to the pricing of Bermudan swaptions in the multi-factor libor market model". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.155208.

Halperin, I. (2019). *QLBS: Q-Learner in the Black-Scholes(-Merton) Worlds*. arXiv: 1712.04609 [q-fin.CP].

Hoencamp, J. H., S. Jain, and D. Kandhai (2022). "A semi-static replication approach to efficient hedging and pricing of callable IR derivatives". In.