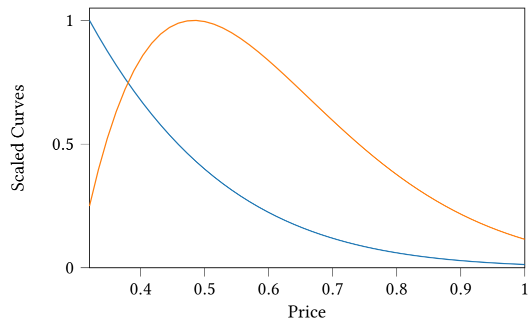

Figure 9: Demand curve used in the noisy experiment and corresponding reward function obtained maximizing profit.

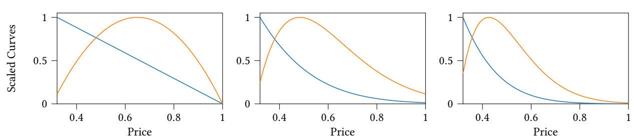

Figure 10: Demand curves used in the non-stationary experiment and corresponding reward functions.

Non-stationary Environment Simulation. In the second experiment, three different volume functions have been used during the different phases of the non-stationary process. The volume functions were:

$$v_1(x) = \frac{3}{10}(1-x),$$
$$v_2(x) = 2e^{-(x+1.2)^{\frac{5}{2}}},$$
$$v_3(x) = 7e^{-(x+1.2)^3}.$$

Their corresponding volumes curves are provided in Figure 10. The first abrupt change substituted the volume function $v_1(x)$ with $v_2(x)$, the second substituted $v_2(x)$ with $v_3(x)$, and the third one $v_3(x)$ with $v_1(x)$. In this set of experiments the noise's standard deviation is $\sigma = 0.001$, and the outliers' percentage is $o = 0\%$.

Algorithm Settings. In the first scenario, the demand curve have been estimated using Bernstein's Polynomial with $Z = 75$. The priors for the Lognormal and Gaussian distribution of the BRL model have been set with $\sigma_h = 0.75$ and $\sigma_h = 0.5$, respectively.

The values for the hyper-parameters have been chosen basing on an independent data set. The sampling procedure described in Section 4 have been applied to the set of margins $\mathcal{M}$ of evenly spaced values over the domain $[0.05, 1.5]$, where $|\mathcal{M}| = 50$.

In the second scenario, we use the same configurations for the Bernstein's Polynomial and the sampling procedure. Conversely, the prior parameters for the Lognormal and Gaussian priors were set to $\sigma_h = 0.75$ and $\sigma_h = 2$, respectively.

Notice that the clairvoyant solution to the problem of maximizing the profits is non-trivial even knowing the real volume functions, due to the fact that the introduction of noise and outliers do not allow to compute it in a closed form solution. We estimated the optimal solution using Monte Carlo approach, *i.e.*, we simulated 10,000 samples from each one of the margins used in the experiments and averaged the values of the profit gained with such a margin. Then we took the maximum over the computed profits as the optimal solution for the problem. Thanks to this approach, the empirical regret is computed as the difference between this value and the one obtained using the analyzed algorithms.

A.4 Algorithm Running Time

The algorithm running time can be analyzed by dividing the process into two phases: first, the distance estimation and the tree structure generation, then, the proper optimal price estimation.

Similarity and Tree Structure. This phase is required to be performed only when there are changes in the catalog of the products. The running time for the distance estimation algorithm is $O(|\mathcal{J}|^2)$ for what concerns the operations required to construct the distance matrix. Building the agglomerative clustering tree structure requires a running time $O(|\mathcal{J}|^2 \log |\mathcal{J}|)$ when using single linkage, and $O(|\mathcal{J}|^3)$ in the general case. It is worth noting that adding a new product to the catalog corresponds to an incremental update of the distance matrix, *i.e.*, adding a new row and column to the matrix consisting of the distance of the new products w.r.t. the previous ones.

Optimal Pricing. The proper estimate of the optimal price must be performed at every time t, as well as the association of a product j with the related meta-product $\mathcal{K}$. This is because the cluster stopping condition is defined over transactions data, which changes over time. Given $|\mathcal{J}|$ products, we must estimate at most (worst-case scenario) $|\mathcal{J}|$ BLR models.