Introduction.} Each year, billions of dollars are invested on product innovation and new product launches. But dealing with new products is not easy. Almost 15\% of the total new products launched in the market each year are unsuccessful and cannot sustain for a long time [1]. Dynamic pricing plays a crucial role in the success of a new product [2]. In this paper we focus on the question of optimal pricing strategies when the underlying demand is unknown and further the unknown demand function is non-linear. We tackle the problem of non linear demand under two different information settings: In the first setting, we assume that the retailer has limited information on the consumer features and in the other setting we assume that the retailer has personalized feature information of its customers (these features can be age, order history, region, family size etc). In the limited information setting, we study a new demand structure that models the change in consumer price elasticities over time in a non linear way. An example of this effect in practice is airline tickets. In this case, the customer is willing to pay a higher price as one gets closer to the date of departure for the same ticket class. Similarly, in fashion retail, as we move into the end of the selling horizon, customers tend to pay less for the same item which translates into increasing price elasticities. We model this effect by introducing the notion of elasticity trends. More specifically, we assume that the price elasticity of the customer population changes as a non linear function of time. We analyze the dynamic pricing problem in both cases when either this price elasticity trend is known or can be estimated and propose a dynamic pricing algorithm that has provable regret guarantees.

We also analyze the case of increased customer information in which the retailer has access to customer features. We assume that the demand in this case is a non linear function of the consumer features deviating from the linear demand assumption prevalent in the existing literature. Modeling this problem as the contextual Multi Armed Bandit, we extend the UCB algorithm [3] for non-linear rewards. Assuming that the parametric form of the nonlinear relation between features and demand is known to the retailer and exploit this information structure to develop near optimal policies.

Contributions.} Our main contribution in this work is to propose near optimal dynamic pricing algorithms that do not impose linearity assumptions on demand and optimally learn and earn over the selling horizon. We analyze two different information cases: one with contextual feature
information and another without contextual feature information. For the contextual information free case, we analyze changing consumer behavior by studying non-linear elasticity trends. In the contextual feature information case, we again use the non linearity of the reward combined with ideas from Taylor approximation to extend the UCB algorithm [3] for non linear reward functions. In both cases, we prove analytical guarantees on the regret incurred by our algorithm. We further test the algorithms on synthetic data sets and show the applicability and improvement of our proposed algorithms over existing algorithms.

Literature Review. Our work lies in the intersection of three different fields: (i) pricing with contextual or feature information (ii) non static demand environments and (iii) exploration-exploitation trade-off with non linear rewards. Recent studies [5], [6] have looked at the problem of feature based pricing. Nevertheless, current literature imposes linearity assumption of the unknown demand while we relax this assumption to include any parametric demand form. Similarly, [7], [8] have looked at optimal dynamic pricing under non stationary demand models. In this work, we study a different structure of non linear demand where we use elasticity trends to model this non linearity. Finally, [4], [3] and others have used the MAB paradigm to solve optimal learning problems but literature on contextual learning problems with non linear rewards is sparse.

Results.

- **Analysis of price elasticity trends:** We study the notion of non stationary demand models using commonly observed price elasticity trends in various industries. These elasticity trends guide our parametric demand model and our proposed pricing policy guarantees optimal parameter learning by ensuring enough exploration in the pricing space as in [9].

- **Extension of feature based demand models to the non-linear case:** We extend the feature based dynamic pricing and learning literature to the non linear case where demand is a non linear function of the contextual features. Using the well known UCB algorithm and ideas from Taylor series approximation, we devise a policy that is simple and easy to implement. We construct upper confidence bounds of the reward using linear approximations of the unknown non linear reward function and optimize decisions based on these optimistic estimates of rewards.
Near optimal pricing algorithms with regret guarantees: We develop near optimal pricing algorithms with provable regret guarantees in comparison to a clairvoyant that has full information of the unknown demand model. For the limited feature information case, our proposed algorithm achieves regret of the order of $O(\sqrt{T})$ and for the full feature information case our proposed algorithm achieves regret of $O(\sqrt{Td \log(T)})$ where $T$ is the total time horizon and $d$ is the dimension of the feature vector.

References


