

Kannan, Aswin; Kreimeier, Timo; Walther, Andrea

Article — Published Version

On solving nonsmooth retail portfolio maximization problems using active signature methods

Computational Management Science

Suggested Citation: Kannan, Aswin; Kreimeier, Timo; Walther, Andrea (2025) : On solving nonsmooth retail portfolio maximization problems using active signature methods, Computational Management Science, ISSN 1619-6988, Springer Berlin Heidelberg, Berlin/Heidelberg, Vol. 23, Iss. 1, <https://doi.org/10.1007/s10287-025-00544-3>

This Version is available at:

<https://hdl.handle.net/10419/342954>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



On solving nonsmooth retail portfolio maximization problems using active signature methods

Aswin Kannan^{1,2} · Timo Kreimeier^{1,3} · Andrea Walther¹

Received: 16 July 2024 / Accepted: 2 September 2025 / Published online: 13 November 2025
© The Author(s) 2025

Abstract

The retail industry is governed by crucial decisions on inventory management, discount offers like promotions and stock clearing as so-called markdowns, presenting two sets of optimization problems. The former is an estimation problem, where the underlying objective is to predict the coefficients of demand (sales) elasticity with respect to product prices. The latter is the dynamic revenue maximization problem, which takes in the coefficients of demand as inputs. While both tasks present nonsmooth optimization problems, the latter is a challenging nonlinear problem in massive dimensions. This is further subject to constraints on inventory, inter-product relationships, and price bounds. Traditional approaches to solve such problems relied on using reformulations and approximations, thereby leading to potentially suboptimal solutions. In this work, we retain the nonsmooth structure generated by the max type (or equivalently absolute value type) function and solve the resulting problem in its abs-quadratic form, i.e., in a quadratic matrix-vector-product based representation including linear arguments in the abs-evaluation. Subsequently, we present an adaptation of the Constrained Active Signature Method (CASM) that explicitly exploits this abs-quadratic structure of the problem yielding the Quadratic Constrained Active Signature Method (QCASM). In the process, we also guarantee convexity of the objectives under some mild realistic assumptions on the market demand and structure. Two real world retail examples (UK and US market data from 2017-2019) and one simulated use-case are studied from an empirical standpoint. Numerical results demonstrate good performance of QCASM and further show that such solvers can be used significantly by the retail science community in the future.

Keywords Retail · Nonsmooth optimization · Revenue maximization · Pricing · Active signature method

1 Introduction and survey

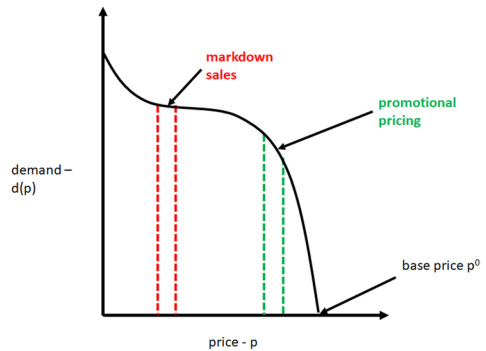
Retail markets primarily gravitate towards revenue maximization. While the revenue can simply be expressed as a direct product of price and sales, these two quantities are actually more complicated in reality (Cohen 2017; Cohen et al. 2016; Kannan et al. 2020) as they depend on multiple factors like consumer demand, time-line, inventory levels, and relationships between SKUs. Here, SKU refers to Stock Keeping Unit, or in simpler terms product ID. It can also be noted that pricing and consumer demand vary significantly from one SKU to another, e.g., pricing of a T-shirt may differ significantly from that of rice or sugar. Consumer demand is usually a consequence of prices, social aspects, geographical indicators, and many other criteria (Roy 1938; Chan et al. 1990). Retailers also face challenges in maintaining sufficient inventory levels at their respective outlets that are subject to several logistic factors, practical weather aspects, and storage limitations. Within the same SKU, the demand towards different sizes can vary. For example, the demand for a medium size T-shirt may be higher than that of a large or small T-shirt. Proposing a different SKU to each such size is deemed to be arduous and less robust, since this can significantly increase the problem dimension. One way to handle such variations is by deploying the stock response factor (Li et al. 2022), which models demand as a function of both customer preferences and available inventory levels.

Another crucial aspect with retail markets is the aspect of discounts (Cohen 2017; Cohen et al. 2016; Kannan et al. 2020; Kannan and Panchamgam 2020). Discounts are usually offered in two levels, respectively referred to as promotions and markdowns. Promotions are lower levels of discounts intended to increase product sales temporarily. These reductions are usually enforced when a new competitive product is introduced or if sudden changes in inventory are required. Products are also promoted when there is a sudden random surge/drop in customer demand levels. Markdowns on the other hand are offered during inventory clearance. These prices are considerably lower than the base and are brought forward as a mechanism to either discard older stocks or when replacing the SKU, e.g., when terminating a product from the catalogue. There exist constraints on the number of times promotions or markdowns are offered, since they are tied to several internal policies and regulations of the corresponding businesses. Lastly, there also exist inter-item constraints within multiple SKUs, e.g., SKU “A” cannot be promoted when SKU “B” is promoted.

It is practically obvious and well established that consumer demand is monotonically decreasing in terms of price, predominantly following an S-like pattern as shown in Fig. 1, (Roy 1938), or Cohen et al. (2016). Here, vaguely the region between the markdowns and the base price is concave. The overall curve can be neither convex or concave and the patterns are not the same as above for all SKUs.

Usually, the windows of promotions and markdowns respectively are between 70-80 percent and 30-50 percent of the base price. However, this is not rigid and the window can be wider/shorter depending on the product type. Quite often, products are priced between the base price and promotional levels. From the standpoint of analysis, both convex (Song et al. 2008; Huang et al. 2013) and concave (Chen et al. 2006; Dossche et al. 2010; Huang et al. 2013) instances of demand-price curves have

Fig. 1 Demand-price curve for regular products/SKUs



been studied in detail. A detailed study on the different types of demand models and the related examples are presented in Sect. 2.

Product prices can either belong to a discrete ladder or be continuous. Say, $p_i^t \in \{1.1, 2.5, 3\}$ represents a discrete ladder. This is prevalent in fashion-type settings. In such settings, the number of price points considered are limited and the problem is reformulated as a mixed integer linear programming problem. The resulting optimization tasks are either solved using standard mixed-integer linear programming solvers like Gurobi for smaller dimensions (Cohen 2017; Subramanian and Sherali Oct. 2010) or solved using heuristics-based approaches for larger dimensional instances (Hua et al. 2021; Ma and Fildes 2017). We focus on solving the continuous version of this problem, allowing for prices to take any intermediate value between the lower and upper bounds. For more details on such models, the reader is asked to refer to (Yaghin Ghasemy Ghomi et al. 2014; Chan et al. 1990).

Underlying Nonsmoothness Demand functions are generally unknown and product-specific, typically estimated from data. While some studies use machine learning to model demand involving both price and design attributes, this work focuses only on price. Linear models offer simplicity but can be limiting, while nonlinear models raise tractability concerns. To strike a balance, the paper adopts piecewise linear demand functions, which capture retail scenarios effectively while remaining computationally manageable.

In problems involving historical data, piecewise linear functions have been found to offer a practical way of representing the demand-price relationship (Cohen 2015, Chapter 4), Kannan et al. (2020); Kannan and Panchamgam (2020). For cases where the underlying demand is convex or concave, corresponding piecewise linear approximations can be constructed accordingly.

$$d_i^t(p_i^t) \approx \max(\tilde{a}_i - \tilde{b}_i p_i^t, a_i - b_i p_i^t), \text{ or } d_i^t(p_i^t) \approx \min(\tilde{a}_i - \tilde{b}_i p_i^t, a_i - b_i p_i^t).$$

Note that $d_i^t(p_i^t)$ denotes the demand of product i at time t , which is dependent on the price p_i^t . Also, a 's and b 's are suitably defined non-negative scalars. It is worth noting that the above expressions can naturally accommodate more than two pieces without loss of generality. Beyond data-driven nuances, the revenue maximization problem is inherently nonsmooth-primarily due to reference price effects and stock response behaviors. While reference price effects resemble price caps, stock response factors

depend on inventory levels rather than price. These aspects are elaborated further in Sect. 2.

Active Signature Methods A natural question that would arise is the possible use of slack variables (sales) s_i^t to reformulate the problem into a smoother version by deploying constraints and incorporating terms of the form $p_i^t s_i^t$ in the objective. For example, $d_i^t(p_i^t) = \min(\tilde{a}_i - \tilde{b}_i p_i^t, a_i - b_i p_i^t)$ can be presented as constraints of the form as follows.

$$\{s_i^t \leq \tilde{a}_i - \tilde{b}_i p_i^t, s_i^t \leq a_i - b_i p_i^t, \forall i, t\}.$$

While such reformulations work reasonably well in practice, we note that these do not present guarantees on global optimality and can potentially lead to suboptimal solutions, unless computationally expensive global solvers are used. Some more foundational details related to such formulations are presented in Sect. 2.3. In cases where the demand function $d_i^t(p_i^t)$ is concave, convexity of the objectives is preserved when using the primitive formulation. These optimization problems fall under the category of “Extended Quadratic Linear Programs” (EQLPs). We propose an appropriate adaptation of the active signature method (Griewank and Walther 2018; Kreimeier et al. 2025) as alternative to solve our EQLP. While Kreimeier et al. (2025) focused on solving a problem with piecewise linear objectives, the novelty in this work arises from extending the same to quadratic objectives of a certain structure. We name our solver Quadratic Constrained Active Signature Method (QCASM). More details of this method are explained in Sect. 3. It is important to note that this extension to the piecewise quadratic case allows for the first time to verify minimality with polynomial effort yielding a complexity-wise feasible approach within an algorithm. The corresponding extension of the optimality conditions from the piecewise linear case to the piecewise quadratic setting are detailed in Sect. 3.2.

Contributions: Our focus is on two different questions.

- **Prediction:** Given historical data on price and demand for each SKU, we solve a regression problem to obtain our nonsmooth demand functions. This problem is a piecewise linear optimization problem and takes the form: $\min_{a,b \geq 0} \sum_{i \in \mathcal{I}} \sum_{j=1}^n |f_i(a, b, p) - d_j^{obs}|$, where $f_i(a, b, p)$ is a PLF (piecewise linear function) of the min or max type. We deploy the extended active signature method mentioned above to solve the same. We consider both convex and nonconvex PLFs in this paper, i.e. $\max(\tilde{a} - \tilde{b}p, a - bp)$ and $\min(\tilde{a} - \tilde{b}p, a - bp)$.
- **Revenue maximization:** Setting $d_i^t(p_i^t) = f_i(a, b, p_i^t)$, we deploy QCASM to solve the revenue maximization problem. This takes the form: $\min_{p \in \mathcal{P}} \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{F}} -f_i(a, b, p_i^t) p_i^t$. Note that the constraint set $\mathcal{P}$ and index sets $\mathcal{T}, \mathcal{F}$ are defined appropriately (more details in the next section).

Although prior work has examined relevant datasets, these are not publicly available (Burman et al. 2021; Hekimoglu Sept. 2019). Some open-source alternatives (Chen xxxx; Seaman and Bowman 2021; Skenderi et al. 2021) exist, but they are primarily geared toward clustering tasks or emphasize product attributes other than price. For our study, we adopt three widely used open-source datasets (Cohen et al. 2022; Daq-

ing et al. 2012; Eben-Chaime 2023) that explicitly provide multi-year price-sales data across SKUs for individual retailers. Section 2 introduces the mathematical model, Sect. 3 outlines the proposed method, Sect. 4 presents empirical results, and Sect. 5 concludes the paper.

2 Problem formulation

Our formulation focuses on the continuous version of the pricing and inventory management problem—two core pillars of retail optimization. We incorporate a range of practical constraints, including inventory limits, bounds on promotions and markdowns, interdependencies in item pricing, and the possibility of replenishments. Discrete pricing aspects, such as ladder-type constraints, are outside the scope of this work and are earmarked for future investigation, with the current study serving as a foundational step in that direction. Following the direction of prior studies (Kannan et al. 2020; Cohen 2017), the revenue maximization problem for a firm can be posed as follows.

$$\begin{aligned} \min_{p, u, I} h(p) &= \sum_{i \in \mathcal{I}} \sum_{t \in \mathcal{T}} \left(\underbrace{-p_i^t d_i^t(p_i^t)}_{\text{revenue}} + \underbrace{\phi(d_i^t(p_i^t))}_{\text{costs}} \right) \\ \text{subject to } I_i^{t+1} &= I_i^t + u_i^t - d_i^t(p_i^t) && \text{Inventory Dynamics constraint} \\ \rho_i^l &\leq p_i^t \leq \rho_i^u, \quad \forall t \in \mathcal{T}_p && \text{Promotion constraint} \\ \theta_i^l &\leq p_i^t \leq \theta_i^u, \quad \forall t \in \mathcal{T}_m && \text{Markdown constraint} \\ p_i^t - p_j^t &\geq \kappa_{ij}, \quad \forall \{i, j\} \in \mathcal{N} && \text{Inter-Item constraints} \\ u_i^t &= 0, \quad \forall t \in \mathcal{T}_i && \text{Non-replenishment time-slots} \\ I_i^t, u_i^t, p_i^t &\geq 0, \quad \forall i \in \mathcal{N}, t \in \mathcal{T} && \text{Non-negativity constraints.} \end{aligned} \quad (1)$$

Here, I_i^t , p_i^t , and u_i^t respectively denote the inventory level, price, and replenishment quantity of product i at time t . The set $\mathcal{I}$ represents the collection of products or SKUs, and $\mathcal{T} = 0, 1, \dots, T$ denotes the set of time steps under consideration (typically in days), where T is the total number of such steps. At various times of the year, stores may not receive any replenishments due to regulatory restrictions, transportation challenges, or unforeseen disruptions such as power outages. Seasonal discounts—comprising promotions and markdowns—are applied only at specific time points denoted by $\mathcal{T}_p$ and $\mathcal{T}_m$, respectively. The parameters ρ^l and ρ^u define the lower and upper pricing bounds on promotional prices, while θ^l and θ^u do the same for markdown pricing. Additionally, $\mathcal{N}$ denotes the set of product pairs that are linked through inter-item (relative pricing) constraints. This notation will be followed consistently throughout the document.

An example of inter-item constraints is the following: If milk is priced at 2 Euros per litre, then yogurt should be priced at the least around 1.5 Euros per litre. The demand $d_i^t(p_i^t)$ predominantly follows different trends and can turn out to be discontinuous, nonlinear, and nonconvex. Some real world examples of historical data from a retail store are shown in Fig. 2. Given that there are no closed form expres-

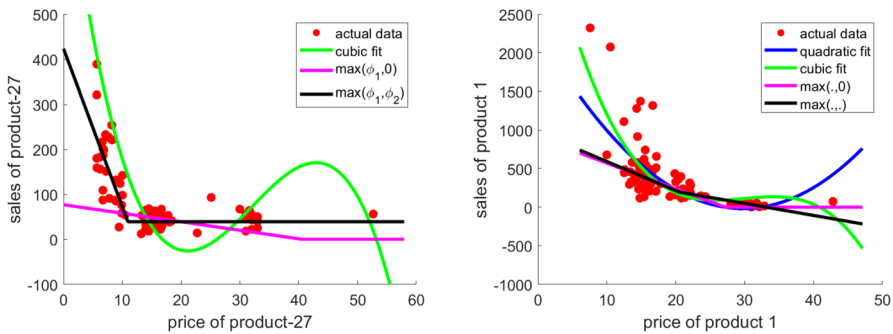


Fig. 2 Examples of demand-price curves for regular products from Cohen et al. (2022)

sions, we attempt to approximate the demand by means of piecewise linear functions. There has been extensive research along these lines for several types of demand functions (Huang et al. 2013). Recent granted patent works (Kannan and Panchamgam 2020) use one such approximation to solve a discrete version of our problem, where $a_i^t(p_i^t) \approx \max_j(h_j(p_i^t))$, $h_j(p_i^t) = a_i^j - b_i^j p_i^t$.

While the above formulation is comprehensive, we focus on a simplified variant for our algorithmic implementation (discussed later). This reduction is primarily to manage complexity and maintain clarity. Importantly, the proposed QCASM algorithm remains applicable to the general formulation, and our intent here is to demonstrate its feasibility. A deeper investigation into computational challenges and broader generalizations is reserved for future work. This section is structured as follows. We begin by outlining various classes of demand-price relationships commonly encountered in practice. We then present the inner regression problem used to estimate demand from historical price-sales data. Finally, we introduce a simplified formulation of the problem in an abs-quadratic form, which aligns with the requirements of the QCASM framework.

2.1 Demand types

In this subsection, we focus on elaborating on different types of demand models that have been deployed in literature from the context of portfolio optimization. We also signify the importance of modeling demand with piecewise linear functions. We begin by exploring fundamental economic models that define demand as a function of reference prices (Chen 2016). These models naturally encompass simpler frameworks where demand is constrained to be non-negative. In certain retail portfolios, pricing structures exhibit relative pricing effects, where demand drops significantly once a price crosses a certain threshold. This phenomenon is commonly studied in both game-theoretic contexts (Fibich et al. 2003) and single-agent optimization settings (Greenleaf 1995). Moreover, considerable attention has been given to such pricing structures in marketing-driven retail environments with multiple sales channels (Chiang et al. 2003). Retail industries often involve both physical outlets and online stores. While indirect channels typically show minimal reference price effects, direct channels tend to exhibit pronounced demand behaviors that require such nonsmooth

pricing effects. These nonsmooth behaviors are particularly relevant in direct-to-consumer channels, where the consumer's sensitivity to price changes can be starkly different across different price ranges. Such complex pricing dynamics can significantly impact sales strategies and optimization models in retail. Some examples of such functions are presented below.

$$d^t(p^t) = a - bp^t - \gamma(p^t - p_{\text{ref}}), \quad \gamma = \begin{cases} \gamma_{\text{ref}}, & p^t > p_{\text{ref}} \\ 0, & p^t \leq p_{\text{ref}} \end{cases}.$$

Positive sales: In simpler cases, where the demand function is linear and smooth, it can still be noticed that the sales have to be non-negative (Kunz and Crone 2014). This is dictated by

$$s^t = \max(d^t, 0) = \max(a - bp^t, 0).$$

It can be easily observed that this is a special case of the model with reference price effects. A natural question that would arise is whether this behavior can be simply invoked in the problem by plugging in s^t instead of d^t in the objective and imposing constraints $s^t \geq 0$ and $s^t \geq d^t$. However, such a formulation induces bilinear terms of the form $p^t s^t$ in the objective, and we lose the benefits of the primitive formulation as discussed earlier.

From a business perspective, this is crucial in that when demand is approximated by linear models, there is a possibility of a negative demand mathematically. While an optimizer technically will stay away from such a region, this difficulty is obviated when using a $\max(\cdot, 0)$ type function to approximate demand data. This is one of our crucial models for numerical results (as shown later). At the same time, when using other PLF approximations, we obtain our price bounds from data. These upper bounds on price are reasonable in practice in that they inherently keep the demand positive.

Zero price-elasticity: In wholesale outlets, demand responses can be significantly low or even zero for certain levels of price. For example, assume that the regular price of a milk is 1.2 Euros per litre and the demand at this level is 1000 units. If the price drops to 0.5 Euros, the demand increases to 5000 units. If the price drops further down to 0.2 Euros, the demand still stays at 5000 units. Beyond a point, the society becomes “in-elastic” to price changes. This effect can be described by $d(p) = \min(a - bp, \tilde{d})$. This setting is also analogous to price-caps in Cournot based regimes (Jean and Patrick 2019), where prices are defined as a consequence of demand, e.g., $p(q) = \min(a - bq, \tilde{p})$. In both of these cases, the revenue functions lead to convex optimization problems.

Polynomial Models: In addition to reference price models, polynomial functions offer a relatively straightforward way to represent demand. A few illustrative examples are presented below.

$$\underbrace{d_i^t(p_i^t) = a_i - b_i(p_i^t)^\gamma}_{\text{concave for } \gamma \geq 1}, \quad \underbrace{d_i^t(p_i^t) = (a_i - b_i p_i^t)^\gamma}_{\text{concave for } \gamma \leq 1 \text{ and convex for } \gamma > 1}, \quad \underbrace{d_i^t(p_i^t) = a_i(p_i^t)^{-b_i}}_{\text{convex}}, \quad (2)$$

Note that a_i , b_i , and γ are positive scalars. A broader discussion on various demand-pricing functional forms can be found in (Huang et al. 2013). Traditional demand models treat demand as a static function of current price, disregarding historical data. In contrast, dynamic models account for the influence of past prices on present demand, as explored in works such as Cohen (2015, 2017). An illustrative example follows.

$$d_i^t(p_i^t) = (p_i^t)^{\alpha_1} (p_i^{t-1})^{\alpha_2} \dots (p_i^{t-K})^{\alpha_{K+1}}.$$

While the dynamic model captures historical pricing effects more realistically, it poses significant computational challenges for large-scale revenue maximization. To maintain tractability, we therefore focus on a static formulation, where demand d_i^t depends solely on the current price p_i^t . Beyond polynomial and reference price-based formulations, the literature offers a wide array of demand models-including logarithmic, exponential, logit, and Multiplicative Competitive Interaction (MCI) models. For clarity and ease of analysis, we categorize these broadly into convex and non-convex classes. As we will see, many of these models can be closely approximated by specific types of piecewise linear functions (PLFs), facilitating both modeling flexibility and computational tractability.

2.1.1 Convex demand

An exhaustive survey of convex models has been described in Huang et al. (2013). This type of behavior stems in both competitive (oligopolistic) and single-firm based settings. In the case of competition, it has been well accepted that demand either falls under the category of discrete choice or Multiplicative Competitive Interaction (MCI) models. Discrete choice models are either logistic (Kunz and Crone 2014) or probit type (Chintagunta 2001), which are mostly concave. As an alternative to logit models, the MCI model (Nakanishi and Cooper 1974; Kunz and Crone 2014) has also been studied in great detail. A static example (dropping time suffixes) of MCI type models from the standpoint of competition is given in Eq. (3). Here, p_{-i} refers to prices of agents (retailers) other than i . Logarithmic type models (log-log) of the form given by Eq. (3) have also been studied in literature (Cohen 2017). As mentioned earlier, Cohen (2017) also presents an MCI model in the context of dynamic and non-competitive regimes. In addition to the earlier discussion, figure 2 also depicts convex behavior of demand.

$$\text{MCI: } d_i(p_i, p_{-i}) = a_i p_i^{b_i} \prod_{j \neq i} p_j^{b_j}, \quad \text{Log-Log: } \log(d_i(p_i^t)) = a_i \log(p_i^t) + b_i. \quad (3)$$

We further highlight a few important model classes from literature as follows.

- Linear type: These take a simple form $d(p) = a - bp$ and have been extensively studied (Mills 1959). Nonsmooth counterparts related to price caps and reference prices have also been studied (to be discussed later).
- Iso-elastic type: These have a sharper elasticity or response towards price (Karlin 1962) and have also been widely studied in literature. These are of the form

$$d(p) = ap^{-b}.$$

- Power Linear type: These are inverse demand functions that are represented as follows (Song et al. 2008): $d(p) = (a - bp)^\gamma$, $\gamma > 1$.
- Hybrid versions (Lau et al. 2003): They are combinations of any of the above three forms of functions (additive or composition).

Smooth and convex demand functions can be approximated by a max over multiple linear functions, i.e.,

$$d(p) = \max_i (a_i + b_i p), \quad a_i, b_i \geq 0.$$

One of the main disadvantages with this setting is that the revenue function, which is a product of negative price times demand happens to be nonconvex. In this relation, we present the following lemma on convexity. We note that the related extensions follow by direct inspection.

Lemma 2.1 *Let $f(p) = -p \cdot \max(a - bp, 0)$ and let $a, b \geq 0$. Then, the function $f(p)$ is not convex over the complete domain $-\infty \leq p \leq \infty$.*

Proof Let us consider p_1 and p_2 such that $a - bp_1 > 0$ and $a - bp_2 < 0$. That is, $p_2 > \frac{a}{b} > p_1$. Let $p_3 = \alpha p_1 + (1 - \alpha)p_2$ for $\alpha \in (0, 1)$. It is easy to observe that $p_2 > p_1$ and $f(p_1) < f(p_2) = 0$. This further implies

$$\alpha f(p_1) + (1 - \alpha)f(p_2) = \alpha f(p_1) < 0. \quad (4)$$

Noting the strict inequalities, it is easy to observe that there exists $\bar{\alpha} \in (0, 1)$ such that

$$p_3 = \bar{\alpha} p_1 + (1 - \bar{\alpha})p_2, \quad \text{where } p_3 = \frac{a}{b}.$$

Using Eq. (4), this implies that $f(p_3) = f(\alpha p_1 + (1 - \alpha)p_2) = 0 > \alpha f(p_1) + (1 - \alpha)f(p_2)$. This completes the proof. $\square$

Remark For max over three functions and higher, recursive arguments can be made by using expressions similar to $\max(f_1, f_2, f_3) = \max(\max(f_1, f_2), f_3)$.

Stock Response Factor: This setting is very common in domains pertaining to fashion. Here, the differences between sales and available inventory can be enormous (Corsten and Gruen 2005) and only a subset of products are kept on display. As an example, consider a firm that may have 100 white T-shirts in their stock. The size is say, spread over “Small”, “Medium”, “Large”, and “XL”. The distribution may not be exactly 25 each. The display may start, e.g., with 10 T-shirts per size. During the course of the day this could be reduced to one T-shirt of size “Small”. This may not exactly correspond to the situation in the inventory. Usually there are minor price dif-

ferences between sizes among the same brand, but these are assumed to follow fixed patterns. Then, the sales capacity is limited by the following expression

$$s_i^t \leq \text{srf}_i^t(I_i^t) \cdot \underbrace{d_i^t(p_i^t)}_{\text{constant}},$$

where the function $\text{srf}(\cdot)$ denotes the stock response factor (SRF), which captures how inventory levels influence demand. The srf function is typically nonlinear, and a widely accepted approach in the literature is to approximate it using a set of piecewise linear functions (Li et al. 2022), as follows: $\text{srf}_i^t(I_i^t) = f_{\text{nonlin}}(I_i^t) \approx \max_j (b_j I_i^t + a_j)$, $a_j, b_j \geq 0$. Here, the objective is to maximize sales and prices are assumed to be fixed as per seasonal rules within a firm. Note that the above inequality automatically presents a nonconvex constraint in s and I . We do not focus on “srf” in this work, but mention it in this subsection, because it follows a similar problem structure as our case of convex demand.

2.1.2 Nonconvex demand

Literature’s focus has also been significant on concave demand type models. These may be classified vaguely into three forms as follows. This classification is not existent in literature and we merely state it for the ease of understanding.

- Logit models: These choice models as stated previously in Chintagunta (2001); Kunz and Crone (2014) have predominantly been the mainstay of literature. Similar to price p , we note that product attributes a_p like design specification become very relevant to the question. An example of a logit model is given as follows.

$$d(p, a_p) = \eta_p \left(1 + e^{-U(p, a_p)} \right)^{-1}.$$

The utility function $U(\cdot)$ is usually linear or quadratic. Consider a winter-wear product (shawl or gloves). Let p and a_1 denote the price and number of woolen layers respectively. Then, the utility function may take the form $U(p, a_p) = 0.1p - 0.01a_1$. The equation above is a case restricted to a single-firm with the customer presented with two options, namely purchase (high utility) and no-purchase (the scalar term “1” indicates the no-purchase option or no-utility). The purchase levels are fractions or probability levels, which are in turn multiplied by the market population η_p . We note that this extends to a multinomial-logit model (Chan et al. 1990) when there are multiple firms, each with an exponential component defined in the denominator. However, we do not focus on product attributes in this work.

- Exponential models (Chen and Simchi-Levi 2004): These models can be defined by $d(p) = a - be^{\gamma p}$, where $a, b, \gamma \geq 0$ are scalars. Exponential models are structurally slightly different from logit models.
- Logarithmic models (Chen et al. 2006; Ray et al. 2005): Their structure is also similar to exponential models and they take the form: $d(p) = (\log a - bp)^\gamma$.

- Polynomial models (Song et al. 2008): These models are more commonly deployed in several industries and are given by $d(p) = (a - bp)^\gamma$ for $\gamma < 1$. Above, we had discussed cases where these models are convex for $\gamma > 1$. Concave demand models of the form $d(p) = a - bp^\gamma$ have also been studied (Chen et al. 2006).

Some examples of data depicting nonconvex demand functions are given in Fig. 3.

We note that when demand is concave, then the resulting minimization problem happens to be convex. Prior to moving to the regression problem, we present the formal statement on convexity as follows.

Lemma 2.2 *Let $f(p) = -p \cdot \min(\tilde{a} - \tilde{b}p, a - bp)$ and let $\tilde{a}, \tilde{b}, a, b \geq 0$. Then, the function $f(p)$ is convex over the complete domain $-\infty \leq p \leq \infty$.*

Proof Using a similar argument along the lines of lemma 2.1, this result follows. The proof can also be taken from lemma 2.4, which is discussed later in this section. $\square$

2.2 Regression subproblem

It can be observed that building “parametric” models based on data automatically leads to regression based settings. Given that we already have piecewise linear functions in our model to be estimated, we move along the path of l1-regression, which further presents piecewise linear functions in the absolute value form, i.e., using $|\cdot|$. The benefit of such an l1-type setting is that the learning problem yields an optimization formulation with piecewise linear objectives, which can be handled by our QCASM framework. Moving to l2-type regression (or polynomial and more complex models) further makes the problem nonlinear, obviously necessitating the use of more stylized solvers and reformulations. We do not intend to compare the complexity versus accuracy behavior of multiple regression models in this work and leave this large area for future research. Since all three classes of PLFs yield structurally similar formulations, we focus on the regression problem corresponding to PLFs

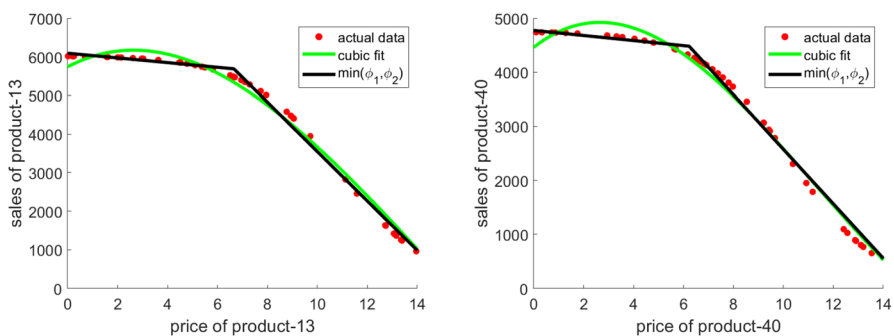


Fig. 3 Concave example of a demand-price curve for products for the logit demand from Eben-Chaime (2023)

of the form $d_i^t(p_i^t) \approx \min(\tilde{a}_i - \tilde{b}_i p_i^t, a_i - b_i p_i^t)$. Given historical price-sales data $\{p_i^{k,sl}, d_i^{k,sl}\}_{k=1}^K$, the corresponding regression problem can be written as follows.

$$\begin{aligned} \min_{a, \tilde{a}, b, \tilde{b}} \quad & \sum_{i \in \mathcal{I}} \sum_{k=1}^K \left| \min(\tilde{a}_i - \tilde{b}_i p_i^{k,sl}, a_i - b_i p_i^{k,sl}) - d_i^{k,sl} \right| \\ \text{subject to} \quad & \tilde{a}_i, a_i, \tilde{b}_i, b_i \geq 0 \quad \forall i \in \mathcal{I}. \end{aligned} \quad (5)$$

Here, $\tilde{a}_i, a_i$ and $\tilde{b}_i, b_i$ denote the intercept and slope of the demand function, respectively. The pair $\{p_i^{k,sl}, d_i^{k,sl}\}$ represents the historical price and corresponding demand observed in the k^{th} record. We denote the price and demand values in data by index “sl” to distinguish the same from decision variables p_i^t and d_i^t . We note that the regression problem also need not necessarily be convex by means of the following corollary.

Corollary 2.3 *Let $h(a, b) = |\max(a - bp, 0) - d|$ for scalars $p, d \geq 0$. Then, $\min_{a, b \geq 0} h(a, b)$ does not necessarily present a convex problem.*

Proof Using the same argument as in lemma 2.1, this result follows. The steps are straightforward. $\square$

We move to the next subsection to analyze the formulation of the larger revenue maximization problem from the angle of a composition of absolute value type functions. As we will see later through the next section, this fits into the general framework of EQLPs, for which our solver QCASM can be deployed.

2.3 Larger problem in abs-quadratic form

We deliberately omit inter-item constraints, replenishment decisions, and full inventory trajectory constraints in this formulation, while retaining only the initial and final inventory bounds. This simplification is motivated by three key considerations. First, inter-item pricing constraints—such as substitution or bundling effects—are often domain-specific and can vary significantly across retailers, making it difficult to define a unified and generalizable formulation. Second, replenishment modeling introduces supply-side dynamics and operational lead-time effects, which, while important, are peripheral to the core pricing optimization structure being studied here. Third, modeling the complete inventory trajectory—i.e., inventory levels at every time point—can significantly increase computational complexity without altering the essential economic intuition captured by initial and final bounds. Importantly, these modeling choices do not compromise the generality of our formulation or the applicability of our solution method. The proposed QCASM framework is structurally compatible with extensions that include inter-item constraints, replenishment policies, and richer inventory dynamics. These components can be layered on top of our formulation with appropriate modifications to the constraint sets, without disrupting the tractability or

convergence guarantees of the algorithm. Our current focus on a more streamlined problem is intended to highlight the methodological core and ensure computational feasibility in empirical settings. These richer extensions are a natural direction for future work and system-level integration.

We state our formulation for the case (1) as minimum over two functions. Note that the flipping of signs (maximum over two functions) and subtraction of linear terms (maximum over a linear function and 0) lead to the other two types of PLFs.

$$\begin{aligned} \min_{p, I} f(p) &= - \sum_{i \in \mathcal{I}} \sum_{t=0}^{T-1} p_i^t \min(\tilde{a}_i - \tilde{b}_i p_i^t, a_i - b_i p_i^t) = \sum_{i \in \mathcal{I}} \sum_{t=0}^{T-1} p_i^t \max(\tilde{b}_i p_i^t - \tilde{a}_i, b_i p_i^t - a_i) \\ \text{subject to : } I_i^{t+1} &= I_i^t + u_i^t + \max(\tilde{b}_i p_i^t - \tilde{a}_i, b_i p_i^t - a_i) \\ p_{i,l} &\leq p_i^t \leq p_{i,u}, & \forall t \in \mathcal{T} & \quad \text{Bound constraints on price} \\ I_i^t &\geq 0, & \forall i \in \mathcal{I}, t \in \mathcal{T} & \quad \text{Non-negativity constraints} \\ I_i^T &\leq \bar{I}, & \forall i \in \mathcal{I} & \quad \text{Final-inventory constraint.} \end{aligned}$$

For the sake of clarity and ease with notation, we integrate both promotion and mark-down type constraints into a single set of bound constraints. Additionally, we unify the lower and upper bounds for promotions (ρ) and markdowns (θ) into p_l and p_u respectively. Introducing additional variables, the problem can be rewritten in the following abs-quadratic form:

$$\begin{aligned} \min_{I, p, s, \hat{s}} \hat{f}(p, s, \hat{s}) &= \sum_{i \in \mathcal{I}} \sum_{t=0}^{T-1} p_i^t (b_i p_i^t - a_i) + \frac{1}{2} \sum_{i \in \mathcal{I}} \sum_{t=0}^{T-1} p_i^t (s_i^t + \hat{s}_i^t) \\ \text{subject to } s_i^t &= (\tilde{b}_i - b_i) p_i^t - \tilde{a}_i + a_i, & \forall t \in \mathcal{T} & \quad \text{(EQ-class-1)} \\ 2I_i^{t+1} &= 2I_i^t + 2u_i^t + 2b_i p_i^t - 2a_i + s_i^t + \hat{s}_i^t, & \forall t \in \{0, \dots, T-1\} & \quad \text{(EQ-class-1)} \\ p_{i,l} &\leq p_i^t \leq p_{i,u}, & \forall i \in \mathcal{I}, t \in \mathcal{T} & \quad \text{(INQ-class-1)} \\ I_i^t &\geq 0, & \forall i \in \mathcal{I}, t \in \mathcal{T} & \quad \text{(INQ-class-1)} \\ I_i^T &\leq \bar{I}, & \forall i \in \mathcal{I} & \quad \text{(INQ-class-1)} \\ \hat{s}_i^t &= |s_i^t|, & \forall i \in \mathcal{I}, t \in \mathcal{T} & \quad \text{(EQ-class-2),} \end{aligned} \quad (6)$$

using $\max(s_i^t, 0) = (s_i^t + |s_i^t|)/2$. Here, note that (INQ-class-1) refers to inequality constraints of a particular form handled by QCASM (described later) and (EQ-class-1) and (EQ-class-2) refer to two similar classes of equality constraints. For the ease of parsing and clarity, we maintain the same terminology in both (6) and (13) discussed in Sect. 3 for constraint classes (EQ-class-1), (EQ-class-2), and (INQ-class-1).

Benefits of the Abs-smooth formulation: Although introducing additional variables increases the problem size, it allows a direct fit into our solver by leveraging a linear relationship between variables s and $\hat{s}$ through absolute value constraints. This structure enables the tracking of signatures—the signs of s —which, while distinct from active-set methods, serve a similar role in identifying nonsmooth points in the constraints. An alternative reformulation using slack variables (termed the inq

reformulation) replaces $\max(s, 0) = t$ with inequality constraints: $\bar{s} \geq s$, $\bar{s} \geq 0$, and $\bar{s} = t$. However, this approach results in a relaxation-not an exact equivalent-especially problematic when the original problem is convex, as convexity may be lost and solutions may diverge. In contrast, our abs-quadratic formulation preserves convexity when the original problem is convex, offering an exact reformulation and maintaining theoretical convergence guarantees for the QCASM method. While empirical comparisons with inq formulations may be promising (even in nonconvex settings), they are beyond this paper's scope. Our focus remains on the theoretical soundness and implementation of QCASM, which is applicable to both convex and empirically to nonconvex problems. Additionally, when demand is modeled as the minimum of two linear functions, convexity of the reformulated problem is guaranteed as stated in the forthcoming lemma.

Lemma 2.4 *Consider problem (6) and the function $\hat{f}$. Assume that there exists a feasible point $(p, s, \hat{s})$ that satisfies all the constraints (EQ and INQ). Additionally, let $p \geq 0$. Then, in the feasible region governed by the constraints (EQ and INQ), the function $\hat{f}$ is convex.*

Proof Exploiting the structure of $\hat{f}$, it suffices to prove the convexity of the following function

$$\tilde{f}(p) = 2p(bp - a) + p((\tilde{b} - b)p - \tilde{a} + a + |(\tilde{b} - b)p - \tilde{a} + a|),$$

for given and fixed values of $\tilde{a}, a, \tilde{b}, b \in \mathbb{R}$. Note that we drop the subscripts of i pertaining to products for the simplicity of notation for this discussion. Let $\tilde{b} = b + \delta, \tilde{a} = a + \gamma$. Then, $\tilde{f}$ can be further simplified to

$$\tilde{f}(p) := (2b + \delta)p^2 + |\delta p^2 - \gamma p|.$$

Setting $\lambda = \frac{\gamma}{\delta}$, then it suffices to prove the convexity of the function $\hat{g}(p)$ for $\lambda \geq 0$.

$$\hat{g}(p) := p^2 + |p^2 - \lambda p|, \quad \lambda \geq 0.$$

Here, we note that for $\lambda < 0$, the $|\cdot|$ sign (absolute value sign) becomes redundant (since $p \geq 0$) and the convexity of $\hat{f}(p)$ follows. Observe that the condition $p \geq 0$ aligns seamlessly with our framework, as prices are inherently lower-bounded by p_l as indicated in (6). Let p_A and p_B refer to any two arbitrary prices that satisfy the bounds, i.e., the inequality constraints INQ in (6). Then, it suffices to show that

$$\hat{g}(\alpha p_A + (1 - \alpha)p_B) \leq \alpha \hat{g}(p_A) + (1 - \alpha)\hat{g}(p_B) \quad (7)$$

holds. We prove this by an exhaustive case-by-case analysis. Let $p_C = \alpha p_A + (1 - \alpha)p_B$ denote the convex combination of p_A and p_B .

- $p_A, p_B \geq \lambda$: Then, inequality (7) follows from the convexity of the quadratic function $(2p^2 - \lambda p)$.

- $p_A, p_B \leq \lambda$: Here, $\tilde{g}(p_A) = \lambda p_A$ and $\tilde{g}(p_B) = \lambda p_B$ holds. Then, the result follows from the convexity of the linear function λp .
- $p_A \leq \lambda \leq p_C \leq p_B$: For the expressions in inequality (7) one obtains

$$\begin{aligned}
 \tilde{g}(\alpha p_A + (1 - \alpha)p_B) - \alpha \tilde{g}(p_A) - (1 - \alpha) \tilde{g}(p_B) &= 2((\alpha p_A + (1 - \alpha)p_B)^2 - \alpha \lambda p_A - (1 - \alpha) \lambda p_B) \\
 &= 2(\alpha^2 p_A^2 + p_B^2 (\alpha^2 - \alpha) + 2\alpha(1 - \alpha)p_A p_B - \alpha \lambda p_A) \\
 &= 2(\alpha^2 p_A^2 - p_B^2 \alpha(1 - \alpha) + 2\alpha(1 - \alpha)p_A p_B - \alpha \lambda p_A) \\
 &\leq 2(\alpha^2 p_A^2 - p_B^2 \alpha(1 - \alpha) + 2\alpha(1 - \alpha)p_A p_B - \alpha p_A^2) \\
 &= 2(-\alpha(1 - \alpha)p_A^2 - \alpha(1 - \alpha)p_B^2 + 2\alpha(1 - \alpha)p_A p_B) \\
 &= -2\alpha(1 - \alpha)(p_A - p_B)^2 \leq 0,
 \end{aligned}$$

proving the assertion for this case.

- $p_A \leq p_C \leq \lambda \leq p_B$: Then one obtains for the expressions in inequality (7) that

$$\begin{aligned}
 \tilde{g}(\alpha p_A + (1 - \alpha)p_B) - \alpha \tilde{g}(p_A) - (1 - \alpha) \tilde{g}(p_B) &= \lambda(\alpha p_A + (1 - \alpha)p_B) - \alpha \lambda p_A - (1 - \alpha)(2p_B^2 - \lambda p_B) \\
 &= -(1 - \alpha)2p_B^2 + 2\lambda(1 - \alpha)p_B = 2(1 - \alpha)p_B(\lambda - p_B) \leq 0,
 \end{aligned}$$

proving the assertion for this case.

Hence, the convexity of our objective $f(\cdot)$ is shown. $\square$

We conclude this section by mentioning that the above lemma on convexity does not hold for the other two instances of PLF given by the $\max(\tilde{a} - \tilde{b}p, a - bp)$ and $\max(a - bp, 0)$ types.

3 Optimization approach

In this section, we will present a solution algorithm for the optimization problem posed previously. Since this algorithm is not specific and exclusive to the problem in this paper and builds on a previously published algorithm, the notation in this section is detached from the other sections.

For solving unconstrained piecewise linear optimization problems, Andreas Griewank and Andrea Walther presented the Active Signature Method (ASM) in Griewank and Walther (2018). The basic idea of this method is to explicitly exploit the structure of the given nonsmoothness caused by the absolute value. For this purpose, the argument space is decomposed into polyhedra and then a sequence of quadratic optimization problems is solved on them. A given starting point determines the initial polyhedron and then optimality conditions help to identify a suitable neighboring polyhedron until a local minimizer is reached.

In Kreimeier (2023) this concept was extended to the Constrained Active Signature Method (CASM) for piecewise linear optimization problems with piecewise linear equality and inequality constraints. Here, an active set strategy is used to handle the constraints. For both ASM and CASM, necessary and sufficient optimality con-

straints, which can be verified in polynomial time, can be derived, see (Griewank and Walther 2018; Kreimeier 2023). Furthermore, in the same references, convergence in finitely many steps is proven. Optimality conditions for the more general nonlinear, nonsmooth case are considered in Hegerhorst-Schultchen and Steinbach (2020).

In this section, first we will introduce some notations and then present the Constrained Active Signature Method for Abs-quadratic Problems, short QCASM, an algorithm derived from CASM to handle additional quadratic components in the objective function.

3.1 Notation and definitions

The basis for developing the new optimization method is the so-called abs-linear form, which was introduced in Griewank (2013) and is defined as follows.

Definition 3.1 (abs-linear form, switching vector) A continuous piecewise linear function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is in *abs-linear form* if $y \equiv f(x)$ is given by

$$y = d + a^\top x + b^\top z, \quad (8a)$$

$$z = c + Zx + Mz + L|z|, \quad (8b)$$

with $x \in \mathbb{R}^n$ the argument vector, $z \in \mathbb{R}^s$ the vector of switching variables, called switching vector, and constants $a \in \mathbb{R}^n$, $b, c \in \mathbb{R}^s$, $d \in \mathbb{R}$, $Z \in \mathbb{R}^{s \times n}$, $L, M \in \mathbb{R}^{s \times s}$, where the last two matrices are strictly lower triangular. Eq. (8b) is called switching system.

Here, $|z|$ denotes the component-wise absolute value of a vector z . In the context of optimization, one can set $d = 0$ without loss of generality. Furthermore, we assume that all components z_i , $1 \leq i < s$, are used as arguments of the absolute value for simplicity of notation. It follows with the identities

$$\min(x, y) = \frac{1}{2} (x + y - |x - y|) \quad \text{and} \quad \max(x, y) = \frac{1}{2} (x + y + |x - y|)$$

and the representation of piecewise linear function shown, e.g., in (Scholtes 2012, Proposition 2.2.2), that every piecewise linear function has a representation in an abs-linear form, see also (Griewank and Walther 2020, Lemma 2). However, this representation might be not unique.

We denote by $\alpha(x) \equiv \{i \in \{1, \dots, s-1\} \mid z_i(x) = 0\}$ the set of indices of so-called active switching variables and by $P_\alpha \equiv (e_i^\top)_{i \in \alpha} \in \mathbb{R}^{|\alpha| \times s}$ the projection onto the active components of $z(x)$, where e_i is the i th unit vector of appropriate size. For each $x \in \mathbb{R}^n$ we define the signature vector by $\sigma(x) \equiv (\text{sgn}(z_1(x)), \dots, \text{sgn}(z_s(x))) \in \{-1, 0, 1\}^s$. The corresponding signature matrix is given by $\Sigma(x) = \text{diag}(\sigma(x))$. As in Griewank (2013), the signature vectors are used to define a polyhedral structure.

Definition 3.2 ((extended) signature domain) For a fixed $\sigma \in \{-1, 0, 1\}^s$, define

$$\mathcal{P}_\sigma \equiv \{x \in \mathbb{R}^n \mid \operatorname{sgn}(z(x)) = \sigma\} \subset \bar{\mathcal{P}}_\sigma \equiv \{x \in \mathbb{R}^n \mid \Sigma z(x) = |z(x)|\}. \quad (9)$$

The set $\mathcal{P}_\sigma$ is called signature domain and the set $\bar{\mathcal{P}}_\sigma$ extended signature domain. As shown in (Griewank 2016, Proposition 2), the union of all possible extended signature domains equals $\mathbb{R}^n$. By definition, the extended signature domains do not have common interior points. Hence, they form a decomposition of $\mathbb{R}^n$. The points where the underlying piecewise linear function is not differentiable are located exclusively on the boundaries of the polyhedra. Motivated by the structure in $\mathbb{R}^2$ these boundaries are also called kinks, see for example (Griewank and Walther 2018; Kreimeier 2023) for more details.

To incorporate also quadratic components in the objective function, we now extend the definition of abs-linear functions as follows.

Definition 3.3 (abs-quadratic form) A continuous piecewise smooth function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is in *abs-quadratic form* if $y \equiv f(x)$ is given by

$$y = x^\top Q_1 x + x^\top Q_2 z + z^\top Q_3 z + a^\top x + b^\top z + d, \quad (10a)$$

$$z = c + Zx + Mz + L|z|, \quad (10b)$$

with $x \in \mathbb{R}^n$ the argument vector, $z \in \mathbb{R}^s$ the switching vector, and constants $Q_1 = Q_1^\top \in \mathbb{R}^{n \times n}$, $Q_2 \in \mathbb{R}^{n \times s}$, $Q_3 = Q_3^\top \in \mathbb{R}^{s \times s}$, $a \in \mathbb{R}^n$, $b, c \in \mathbb{R}^s$, $d \in \mathbb{R}$, $Z \in \mathbb{R}^{s \times n}$, $L, M \in \mathbb{R}^{s \times s}$, where the last two matrices are strictly lower triangular. The chosen name is motivated by quadratic optimization problems, where the objective function is quadratic and the constraints are linear. In the further course of the paper, Eq. (10a) serves as the objective function of the considered optimization problem and Eq. (10b) represents one of the constraint sets. Note that in contrast to the piecewise linear case, not every piecewise quadratic function has an abs-quadratic form, as shown by the following example.

Example 3.4 The function $f_1(x_1, x_2) = |x_1 x_2| = |x_1| |x_2|$ can be represented by the abs-quadratic form

$$y = (z_1 \quad z_2 \quad z_3 \quad z_4) \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \quad (11)$$

$$z = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} |z_1| \\ |z_2| \\ |z_3| \\ |z_4| \end{pmatrix} \quad (12)$$

setting the remaining matrices and vectors to zero. However, the function $f_2(x_1, x_2) = |x_1 x_2 + 1|$ cannot be written in abs-quadratic form, since the switching

equation does not allow a quadratic term. Furthermore, the term $x_1x_2 + 1$ cannot be decomposed into linear factors over $\mathbb{R}$.

3.2 Description of the Algorithm

First, we state the optimization problem for which we will present and analyze a solution algorithm in this paper, i.e., the *constrained abs-quadratic optimization problem* (13) given by

$$\begin{aligned} \min_{x \in \mathbb{R}^n, z \in \mathbb{R}^s} \quad & x^\top Q_1 x + x^\top Q_2 z + z^\top Q_3 z + a^\top x + b^\top z + d \\ \text{s.t.} \quad & 0 = g + Ax + Bz + C|z|, \\ & 0 \geq h + Dx + Ez + F|z|, \\ & z = c + Zx + Mz + L|z|, \end{aligned} \quad (13)$$

with an abs-quadratic target function and $g \in \mathbb{R}^m, h \in \mathbb{R}^p, A \in \mathbb{R}^{m \times n}, B, C \in \mathbb{R}^{m \times s}, D \in \mathbb{R}^{p \times n}$ as well as $E, F \in \mathbb{R}^{p \times s}$. For a detailed investigation of necessary and sufficient optimality conditions we refer to Hegerhorst-Schultchen and Steinbach (2020); Griewank and Walther (2016); Kreimeier (2023).

For later use, we introduce

$$H : \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^s \rightarrow \mathbb{R}^p, \quad (x, z, |z|) \mapsto h + Dx + Ez + F|z|$$

as the function for the inequality constraints. Furthermore, we define $\omega = \text{sgn}(H(x, z, |z|))$ to identify the active inequalities and $\Omega = \text{diag}(\omega)$ as the projection onto these active inequality constraints.

The proposed optimization method is a feasible point approach. Hence, a feasible starting point x_0 is required that determines the initial polyhedron $\mathcal{P}_\sigma$ with $\sigma = \sigma(x_0)$ to be examined. For such a fixed σ , we will now describe the steps of the optimization algorithm:

Computing a descent direction for given σ and ω Following the approaches presented in Griewank and Walther (2018); Kreimeier (2023), on each polyhedron defined by σ taking account the required signs of z , we obtain the optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n, z \in \mathbb{R}^s} \quad & x^\top Q_1 x + x^\top Q_2 |\Sigma|z + z^\top |\Sigma|Q_3 |\Sigma|z + a^\top x + b^\top |\Sigma|z + d \\ \text{s.t.} \quad & 0 = g + Ax + B|\Sigma|z + C\Sigma z, \\ & 0 \geq h + Dx + E|\Sigma|z + F\Sigma z, \\ & 0 = |\Sigma|z - \tilde{c} - \tilde{Z}x, \\ & 0 \leq \Sigma z, \end{aligned} \quad (14)$$

with $\tilde{Z}$ and $\tilde{c}$ defined as

$$\tilde{Z} = (I_s - M - L\Sigma)^{-1}Z \quad \text{and} \quad \tilde{c} = (I_s - M - L\Sigma)^{-1}c. \quad (15)$$

Due to the fixed signature vector, this optimization problem is smooth, has a quadratic target function and linear constraints. When a solution exists, it can be solved with a standard QP method. However, we want to exploit the structure provided by the signature vector as additional feature. For this purpose, we apply standard KKT theory to Equation (14). With Lagrange multipliers $\delta \in \mathbb{R}^m, \nu \in \mathbb{R}^p, \lambda \in \mathbb{R}^s$ and $\mu \in \mathbb{R}^s$, we obtain the following necessary optimality conditions

$$0 = 2x^\top Q_1 + z^\top |\Sigma| Q_2^\top + a^\top + \delta^\top A + \nu^\top D - \lambda^\top \tilde{Z}, \quad (15a)$$

$$0 = x^\top Q_2 |\Sigma| + 2z^\top |\Sigma| Q_3 |\Sigma| + b^\top |\Sigma| + \delta^\top (B|\Sigma| + C\Sigma) + \nu^\top (E|\Sigma| + F\Sigma) + \lambda^\top |\Sigma| - \mu^\top \Sigma, \quad (15b)$$

$$0 = g + Ax + B|\Sigma|z + C\Sigma z, \quad (15c)$$

$$0 \geq h + Dx + E|\Sigma|z + F\Sigma z, \quad (15d)$$

$$0 = |\Sigma|z - \tilde{c} - \tilde{Z}x, \quad (15e)$$

$$0 \leq \Sigma z, \quad 0 \leq \mu, \quad 0 = \mu^\top \Sigma z, \quad (15f)$$

$$0 \leq \nu, \quad 0 = \nu^\top (h + Dx + E|\Sigma|z + F\Sigma z). \quad (15g)$$

Multiplying Eq. (15b) by Σ from the right and using the sign condition for μ yields

$$0 \leq \mu^\top |\Sigma| = x^\top Q_2 \Sigma + 2z^\top |\Sigma| Q_3 \Sigma + b^\top \Sigma + \delta^\top (B\Sigma + C|\Sigma|) + \nu^\top (E\Sigma + F|\Sigma|) + \lambda^\top \Sigma. \quad (16)$$

Due to the complementarity condition $\mu^\top \Sigma z = 0$, this inequality must hold as an equation. Hence, it follows that

$$-b^\top \Sigma = x^\top Q_2 \Sigma + 2z^\top |\Sigma| Q_3 \Sigma + \delta^\top (B\Sigma + C|\Sigma|) + \nu^\top (E\Sigma + F|\Sigma|) + \lambda^\top \Sigma.$$

Thus with $\omega = \text{sgn}(H(x, z, |z|))$ and $\Omega = \text{diag}(\omega)$ denoting as before the projection onto the active inequality constraints, a solution x of the optimization task (14) together with the corresponding z and the Lagrange multipliers must satisfy the linear system

$$\begin{bmatrix} 2Q_1 & Q_2 |\Sigma| & -\tilde{Z}^\top & A^\top & D^\top \\ \Sigma Q_2^\top & 2\Sigma Q_3 |\Sigma| & \Sigma & \Sigma B^\top + |\Sigma| C^\top & \Sigma E^\top + |\Sigma| F^\top \\ \tilde{Z} & -|\Sigma| & 0 & 0 & 0 \\ A & B|\Sigma| + C\Sigma & 0 & 0 & 0 \\ \bar{\Omega} D & \bar{\Omega} (E|\Sigma| + F\Sigma) & 0 & 0 & \Omega \end{bmatrix} \begin{bmatrix} x \\ z \\ \lambda \\ \delta \\ \nu \end{bmatrix} = - \begin{bmatrix} a \\ \Sigma b \\ \tilde{c} \\ g \\ \Omega h \end{bmatrix}, \quad (17)$$

where $\bar{\Omega} = I_p - |\Omega|$ forces the active inequalities to vanish. The matrix Ω in the right lower corner ensures that ν is zero for the inactive inequality constraints. A solution of this linear system must exist as soon as the optimization task (14) has at least one solution. At the end of this section, we will discuss options to ensure the existence of such a solution. For the time being, we assume that we can compute a solution $(\hat{x}, \hat{z}, \lambda, \delta, \nu)$ of this linear system and define for the current iterate x and z

$$\Delta x := \hat{x} - x \quad \text{and} \quad \Delta z := \hat{z} - z$$

as directions towards the next iterate.

Computing a step size β The optimization problem (13) has only (piecewise) linear constraints. This situation was already analyzed for the CASM solver. Hence the step size strategy from CASM (Kreimeier (2023)) can be reused here. Nevertheless, we sketch the step calculation briefly for a complete picture of the algorithm.

For a solution $(\hat{x}, \hat{z})$ of Eq. (17), we must check whether $\sigma(\hat{x}) = \sigma$ is still valid and that the inequality constraints of Eq. (15a) still hold to ensure feasibility. For this purpose, we calculate two step sizes. The first one is the step length from the current iterate x in the direction Δx to a possible kink, i.e., a sign change in one component of z . It is denoted by β^z and defined as

$$\beta^z = \inf_{1 \leq l \leq s} \left\{ \beta_l^z \equiv \frac{-z_l}{\hat{z}_l - z_l} \mid (\hat{z}_l - z_l)\sigma_l < 0 \right\} \in (0, \infty]. \quad (18)$$

If $\beta^z < \infty$ the first index for which the minimum is attained is denoted by j^z . For $\beta^z < 1$, there exists a blocking kink such that we can not take the full step to the solution $\hat{x}$. For a given signature vector σ with $\sigma_i \neq 0$ and $x \in \mathcal{P}_\sigma$ it follows that $z_i(x) \neq 0$. Thus, $\beta^z > 0$ must hold.

The second step size is the step length from the current iterate x in the direction Δx to a possible inequality constraint $H_l(x, z, \Sigma z)$, $1 \leq l \leq p$, that becomes active. Similar to the computation of β^z , this step size β^H is given by

$$\beta^H = \inf_{1 \leq l \leq p} \left\{ \beta_l^H \equiv \frac{H_l}{H_l - \hat{H}_l} \mid (\hat{H}_l - H_l)\omega_l < 0 \right\} \in (0, \infty], \quad (19)$$

where $H \equiv H(x, z, \Sigma z)$, $\hat{H} \equiv H(\hat{x}, \hat{z}, \Sigma \hat{z})$ and l denotes the l th component of H and $\hat{H}$, respectively. We denote by j^H the smallest index for which the minimum is attained. If $\beta^H < 1$, there exists a blocking inequality constraint, i.e., the solution $\hat{x}$ is not feasible. Therefore, the new iterate x^+ should be chosen such that the j^H th components of $H(x^+, z^+, \Sigma z^+)$ and $\omega(x^+)$ drop to zero in comparison to $H(x, z, \Sigma z)$ and ω , respectively. Setting $\omega_{j^H}^+ = 0$ changes the optimality system (15a) and a new solution of system (17) has to be computed. If $\beta^H \leq \beta^z$ then we have $z_{j^H} \hat{z}_{j^H} \geq 0$ such that the iterate $\hat{x}$ is still contained in $\mathcal{P}_\sigma$, i.e., $\sigma(\hat{x}) = \sigma$ is still valid. Since all active constraints are encoded in ω , one must have $\beta^H > 0$.

Finally, we determine the actual step size

$$\beta = \min\{\beta^z, \beta^H, 1\} \in (0, 1], \quad (20)$$

where the upper bound 1 on β ensures with the update

$$x^+ = (1 - \beta)x + \beta\hat{x} = x + \beta\Delta x$$

that the next iterate is feasible for the optimization task (14) and the current σ . As can be seen, the case $\beta < 1$ corresponds to the activation of a kink or an inequality constraint, respectively. Therefore, we will refer to this situation as a restriction of σ or ω , respectively. If $\beta = 1$, one has for the new iterate $x^+ = \hat{x}$ that $\sigma(x^+) = \sigma$ and $\omega(x^+) = \omega$. In this case, x^+ is called *signature stationary* since the two signature vectors are kept.

Checking the optimality If x^+ is signature stationary on the current polyhedron $\mathcal{P}_\sigma$, it fulfills the necessary optimality condition of the optimization task (14) on $\mathcal{P}_\sigma$. Then, one has to check whether x^+ is a minimizer of (13). If this is the case the iteration stops. Otherwise, the optimization continues in one of the neighboring polyhedra $\mathcal{P}_{\tilde{\sigma}}$ with $\tilde{\sigma} \succ \sigma$ where

$$\tilde{\sigma} \succeq \sigma \quad \text{holds if} \quad \tilde{\sigma}_j \sigma_j \geq \sigma_j^2 \quad \text{for } j = 1, \dots, s.$$

Such a $\tilde{\sigma}$ can be decomposed into $\sigma + \gamma$ with $|\sigma|^\top |\gamma| = 0$. Replacing Σ in the optimality conditions (15a) by the corresponding $\Sigma + \Gamma$ and using

$$z_{\tilde{\sigma}}(x) = z_{\sigma+\gamma}(x) = (I_s - M - L\Sigma - L\Gamma)^{-1}(c + Zx) = (I_s - \tilde{L}\Gamma)^{-1}(\tilde{c} + \tilde{Z}x)$$

it follows that most of the equations still hold for the current values of $\hat{x}$, $\hat{z}$ and λ . Just Eq. (16) changes in that it has as many new nontrivial components as γ which are given by

$$\begin{aligned} 0 \leq \mu^\top |\Gamma| &= \hat{x}^\top Q_2 \Gamma + 2\hat{z}^\top Q_3 \Gamma + b^\top \Gamma + \delta^\top (B\Gamma + C|\Gamma|) + \nu^\top (E\Gamma + F|\Gamma|) + \lambda^\top (I_s - \tilde{L}\Gamma)\Gamma \\ &= (\hat{x}^\top Q_2 + 2\hat{z}^\top Q_3 + b^\top + \delta^\top B + \nu^\top E + \lambda^\top) \Gamma + (\delta^\top C + \nu^\top F - \lambda^\top \tilde{L})|\Gamma|. \end{aligned}$$

This condition is violated if and only if there exists at least one index k with $1 \leq k < s$ such that

$$\gamma \equiv - \operatorname{sgn}(\hat{x}^\top Q_2 + 2\hat{z}^\top Q_3 + b^\top + \delta^\top B + \nu^\top E + \lambda^\top) e_k$$

satisfies

$$0 > (\delta^\top C + \nu^\top F - \lambda^\top \tilde{L})e_k - |\hat{x}^\top Q_2 + 2\hat{z}^\top Q_3 + b^\top + \delta^\top B + \nu^\top E + \lambda^\top| e_k \quad \text{and} \quad \sigma_k = 0. \quad (21)$$

This optimality condition can be verified as a matrix-vector product in polynomial time. For the unconstrained and abs-linear case a similar complexity result was shown in Griewank and Walther (2018) and for the general case of abs-smooth and constrained problems in Hegerhorst-Schultchen and Steinbach (2020).

The overall algorithm Combining all the steps described above, one obtains Algo. 1 as an extension of (Kreimeier et al. 2025, Algorithm 1) for the piecewise linear case. Once more, the resulting algorithm consists of three main parts: First, the computation of the search direction (cf. line 2 of Algo. 1). Second, computing the step size and in case of blocking kinks and/or inequality constraints restrict σ and/or ω , respectively (cf. line 3-6). Third, checking the optimality and relaxing kinks or constraints in case of nonoptimality (cf. line 7-14).

Require: Feasible start point $x \in \mathbb{R}^n, n \in \mathbb{N}, s, m, p \in \mathbb{N} \cup \{0\}, a \in \mathbb{R}^n, b, c \in \mathbb{R}^s, Z \in \mathbb{R}^{s \times n}, L, M \in \mathbb{R}^{s \times s}$ strictly lower triangular, $Q_1 = Q_1^T \in \mathbb{R}^{n \times n}, Q_2 \in \mathbb{R}^{n \times s}, Q_3 = Q_3^T \in \mathbb{R}^{s \times s}, g \in \mathbb{R}^m, h \in \mathbb{R}^p, A \in \mathbb{R}^{m \times n}, B, C \in \mathbb{R}^{m \times s}, D \in \mathbb{R}^{p \times n}, E, F \in \mathbb{R}^{p \times s}, \beta = 0$

Set: $z := z(x)$ via Eq. (8b), $\sigma := \sigma(x)$ and $\omega := \omega(x)$

```

1: loop
2:   If possible compute  $(\hat{x}, \hat{z}, \lambda, \delta, \nu)$  by solving Eq. (17) else return ERROR
3:   Compute  $\beta^z$  via Eq. (18),  $\beta^H$  via Eq. (19) and  $\beta$  via Eq. (20)
4:   Set  $(x^+, z^+) = (1 - \beta)(x, z) + \beta(\hat{x}, \hat{z})$ 
5:   if  $\beta^H = \beta$  then Restrict  $\omega$  ▷ Add constraint
6:   if  $\beta^z = \beta$  then Restrict  $\sigma$  ▷ Add kink
7:   if  $\beta = 1$  then ▷  $x^+$  is feasible signature stationary
8:     if  $\nu \not\geq 0$  then
9:       Relax  $\omega$ , set  $\beta = 0$  ▷ Drop constraint
10:    else ▷  $x^+$  is feasible signature optimal
11:      if Eq. (21) holds true then
12:        Relax  $\sigma$ , set  $\beta = 0$  ▷ Drop kink
13:      else ▷  $x^+$  is local optimal
14:        return  $(x^+, z^+)$  ▷ Problem solved
15:   Set  $(x, z) = (x^+, z^+)$ 

```

Algorithm 1 Quadratic constrained active signature method (QCASM)

As can be seen from Line 2 of Algo. 1, in each iteration of QCASM one system of linear equations, see Equation (17), has to be solved. The dimension of this system depends on the dimensions n, s, m and p , i.e., the number of variables, the number of switching variables as well as the number of equality and inequality constraints, respectively. Depending on the software used, properties such as the block structure (see also (Kreimeier 2023, Lemma 4.9) for the linear case) or sparsity can be exploited to reduce the overall computational effort. In the applications considered so far, the matrices occurring in (13) are usually very sparse. Assuming a feasible starting point, the system is always consistent, so that a solution of (17) always exists. The complexity of the remaining steps in Algo 1 is negligible.

Finite convergence of QCASM For the piecewise linear case, convergence properties of (C)ASM have already been analyzed in detail in Griewank and Walther (2018); Kreimeier (2023). The basic idea is that by decomposing $\mathbb{R}^n$ there are only finitely many polyhedra and on one polyhedron a solution is found in finitely many steps. However, the essential difference lies in the quadratic parts of the objective function. For the piecewise linear case considered in Griewank and Walther (2018); Kreimeier (2023) the resulting linear problems on each polyhedron were extended by a quadratic penalty term with a positive definite matrix Q to ensure the existence of a minimizer in each polyhedron. To derive the same convergence results for the optimization task (13), a similar requirement would be that the matrix

$$Q = \begin{pmatrix} Q_1 & \frac{1}{2}Q_2 \\ \frac{1}{2}Q_2^\top & Q_3 \end{pmatrix}$$

is positive definite. However, this is not necessarily the case for the examples considered in the earlier sections of this paper. On the other hand, it was shown in the Lemmas 2.1 and 2.4 that the target function is convex on the feasible domain. Therefore, we will now analyze the convergence of QCASM under more general convexity assumptions including also the case that the matrix Q is positive definite. For this purpose, first we define a regularity condition, which has already been introduced (Griewank and Walther 2019) for the unconstrained case and in Hegerhorst-Schultchen and Steinbach (2020) for the constrained case. It corresponds to the well-known LICQ in the smooth case.

Definition 3.5 (LIKQ (constrained case)) Let a constrained optimization problem of the form (13) and a signature vector $\sigma \in \{-1, 0, 1\}^s$ be given. Then the linear independent kink qualification (LIKQ) holds at a feasible point x of the corresponding optimization problem (14) if the active Jacobian

$$\mathcal{J}_\sigma \equiv \begin{bmatrix} A + B|\Sigma|\tilde{Z} + C\Sigma\tilde{Z} \\ P_{\mathcal{I}}(D + E|\Sigma|\tilde{Z} + F\Sigma\tilde{Z}) \\ P_\alpha\tilde{Z} \end{bmatrix} \in \mathbb{R}^{(m+|\mathcal{I}|+|\alpha|) \times n} \quad (22)$$

has full row rank $m + |\mathcal{I}| + |\alpha|$. Here, $\mathcal{I} \equiv \mathcal{I}(x)$ collects the indices of the active inequality constraints at x and the projection onto the active components of $H(x)$ is defined as $P_{\mathcal{I}} \equiv (e_i^\top)_{i \in \mathcal{I}} \in \mathbb{R}^{|\mathcal{I}| \times p}$ with e_i denoting the i th unit vector of appropriate size.

Now, we present one of the main contributions of this paper.

Theorem 3.6 (Finite convergence of QCASM) Let an optimization problem of the form (13) be given and suppose that LIKQ holds at every feasible point. Furthermore, assume that the target function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ of (13) is convex on the compact feasible set $\mathcal{F}$. Then, Algorithm 1 terminates for any feasible starting point $x \in \mathbb{R}^n$ after finitely many iterations at a global minimizer of (13).

Proof Since f is continuous and $\mathcal{F}$ is compact, there exists a minimizer of the optimization task (13). For a fixed given σ , the same statement holds for the optimization task (14), such that one can always compute a solution of Eq. (17) and Algorithm 1 does not terminate with an error. Moreover, it is clear that if the current iterate x_i is not a minimizer of (13), then a descent direction must exist. Now, we show that such a descent direction is determined by Algorithm 1.

Since standard KKT theory was used to derive the system of Eq. (17), its solution is a stationary point for the currently active set of constraints, encoded by the matrix Ω , and the active switches, encoded by the matrix Σ . Since f is convex on the feasible set, such a stationary point is already a minimum for the restricted subproblem defined by the current Ω and Σ . If the solution of (17) yields a point $\hat{x}$ that is not

equal to the current iterate x , then $\Delta x = \hat{x} - x \neq 0$ defines a direction of descent. If $\hat{x} = x$, but x is not a minimizer of the original problem (13), then active inequality constraints and/or switches must be dropped. Since we can ensure in the next iterate that the step length β is larger than zero, a cycling can not happen. Hence, standard active set theory yields the convergence to a minimizer of the optimization task (14) in finitely many steps.

There are only finitely many polyhedra. Furthermore, if the signature vector σ is modified by Algo. 1, the value of the target function is consistently reduced in the next iterate. Hence, the signature vector can be modified only finitely many times leading to a finite convergence of the overall algorithm. $\square$

4 Numerical validation

This section is organized into two subsections. The first subsection discusses about a few possible data sources that have been used in the literature. This is included for the benefit of the community and aims at showing the importance of such datasets. We also note that most of these datasets are proprietary and are far from being published in open-source repositories. The second subsection presents the numerical performance of our proposed approach on three retail datasets. Two out of the three datasets correspond to real data and one refers to simulated data.

4.1 Data sources

We divide this subsection further into three parts, where the first one elaborates on case studies that are proprietary, the second one on the datasets we use for this work, and the third one on other possible sources that can be used for the future.

4.1.1 Proprietary industry data

We present two use-cases that specifically focus on problems pertaining to discounts and pricing.

- **Assortment Optimization Case (2017):** This work presents an inventory management problem, where the demand is a consequence of product prices. Here, promotions and markdowns follow a slightly different structure and apply not to the product pricing, but instead to their shelf assortment within the store. Here, product sales is limited by an upper bound, which is a consequence of some pre-defined rules set by the store.

$$d_i^t = \min(\hat{d}_i^t(p_i^t), \bar{d}_i^t), \quad \hat{d}_i^t(p_i^t) = \max(a_i - b_i p_i^t, 0).$$

Note that $\bar{d}_i^t$ is a constant, which is determined based on the various shelf widths and heights, specific to the store. Recent work (Hekimoglu et al. 2019) investigates this

setting in detail based on real business data from a Turkish grocery store. However, the data has not been made publicly available.

- **Markdown Optimization Case (2020/2021):** Recent work (Burman et al. 2021) considers the promotions planning problem (discrete version) in spirit similar to our proposed formulation. However, the setting is slightly more complex, where the objective includes metrics that are not restricted to only revenue. This data is from a European retail outlet (details also anonymized) and is not publicly available.

4.1.2 Open source repositories

While there are multiple datasets available from the standpoint of retail, e.g., market basket data, product attribute data, inventory data, and geographical variation data, very few actually correspond to historical price-sales entries. Of these, we describe two crucial and widely used datasets and one simulated dataset as follows.

- **Tech-Gadget Retailer (Cohen’s dataset):** This dataset is anonymized and refers to real data from a US based retailer (Cohen et al. 2022). It comprises 44 SKUs and 4400+ entries, with individual columns for price and sales. This dataset has been used as one of the standards for demand forecasts, e.g., by Deng et al. (2022); Cohen et al. (2019). The time stamps for the records range from the years of 2016 to 2018 for 98 weeks.
- **Online Retail:** This refers to data from the UCI repository (Daqing et al. 2012) and contains all the transactions between Jan. 2010 and Sep. 2011 for a real UK based online retail company (details similarly anonymized). The customers are wholesalers and the products/SKUs are “gifts” (occasion specific). The dataset has a few additional indicators like geography. It contains more than 2500 SKUs and roughly 50,000 records (transactions).
- **Logit:** Here, we simulate data based on previous work on logit models (Eben-Chaïme 2023). While the logit model is well established and possibly the most widely used, most of these models also contain product attributes or competition related additions. The model takes the form: $d(p) = \{1 + \exp\{-\psi(p)\}\}^{-1}$. We use the same parameters from Eben-Chaïme (2023), adding minor randomization with respect to the products. Our data is generated for 50 products and 50 weeks.

4.1.3 Other references

Some datasets from Chen (xxxx); Seaman and Bowman (2021) and beyond Skenderi et al. (2021) also contain the necessary price-sales entries. However, they consist of information on product attributes in addition to prices. These datasets have been mainly analyzed in literature from the standpoint of clustering or association rule mining. Demand forecasting in these cases has also been studied, but over these comprehensive input features. These datasets hence do not ideally suit our case study.

4.2 Numerical results

We present four sets of results in this discussion. The numbers denoting the revenue are expressed in the respective currency units (US-Dollars or Pounds). First, we analyze the progressive performance of the QCASM algorithm for the different datasets. Second, we discuss the accuracy of the three different piecewise linear functions on each of these datasets. Third, we evaluate the performance of our optimization algorithm against standard baseline pricing strategies as well as randomly generated price sets. Finally, we benchmark our QCASM-based approach against the conventional state-of-the-art solver Gurobi. To integrate the primitive nonsmooth optimization problem into this solver framework, we reformulate it into a smooth mixed-integer optimization problem. Further details on this transformation are provided in the final subsection of this section. All our computations were done on a quadcore Linux machine (Ubuntu 15.0) with 2.7Ghz speed and 32 GB RAM. We consider a simple version of the revenue maximization problem, which is continuous and further ignore inter-item constraints. Prices are bounded from above and below, where upper bounds indicate the base price. The lower bounds and the interior values of price indicate discounts. The price bounds are set based on historical data. The initial inventory and replenishment levels are set to be proportionate to the total historical sales. All the datasets considered in our study and the related implementation details are available online¹.

Additional Specifics: Beyond the hardware and platform details, we highlight several key aspects of our implementation. In implementing QCASM, we begin with a feasible starting point-specifically, a vector of ones representing base prices, as it yields higher initial revenue than full discounts (zeros). This choice also aligns with the goal of evaluating revenue gains from offering discounts. Given the large size of our datasets, especially UCI, we report revenue outcomes for a representative subset of products, selected from low, medium, and high revenue segments. This approach provides a comprehensive view of QCASM's performance across the portfolio, with a slightly larger sample used for UCI to reflect its scale. For consistency in presentation, we order the product columns in ascending or lexicographic order. We proceed with detailed performance analysis in the following paragraphs. The paper presents numerical results based on a simplified version of the problem (6), which excludes inter-item constraints and stock response factors. This simplification helps keep the focus on the core idea. Including these additional constraints would not change the fundamental theory behind deploying QCASM. However, exploring these empirical aspects further is suggested as a future research direction.

Iterative Performance: We begin our analysis with Cohen's dataset. For this dataset, the sales data obviously appeared to follow a non-concave trajectory and hence the choice of $\max(.,.)$ is practical. Given the 44 SKUs, QCASM was used to solve both the regression and revenue maximization problems. The termination criterion for QCASM was 200 iterations or attainment of optimal error tolerance levels. Table 1 shows the revenue of different products at different stages of the QCASM algorithm. The estimation problem is relatively simple in that it consumes

¹ Please follow this link for all details: <https://scm.cms.hu-berlin.de/aswinkannan1987/retail-code-base>.

Table 1 Iterative progress of QCASM for Cohen's problem

Iteration	Revenue (in multiples of 100,000)					
	Product 4	Product 5	Product 7	Product 8	Product 12	Total
1	0.07529	0.04287	0.11084	1.14629	0.56539	24.4017
10	0.07620	0.04380	0.11968	1.16121	0.56647	25.4976
20	0.07734	0.04496	0.13073	1.17985	0.56783	26.8673
50	0.08075	0.04844	0.16388	1.23578	0.57191	30.9767
70	0.08121	0.04891	0.18598	1.27306	0.57191	32.3248
80	0.08121	0.04891	0.19703	1.29170	0.57191	32.8250
100	0.08121	0.04891	0.21913	1.32899	0.57191	33.8253

a significantly lesser number of iterations. Hence, we ignore reporting the related iterative performance. Here, coefficients a and b refer to the values from the fit $d_i^t(p_i^t) \approx \max(a_i - b_i p_i^t, \tilde{a}_i - \tilde{b}_i p_i^t)$.

It can be observed that convergence is reached in very few iterations for some products. This is very much expected since the variability in price and demand levels is huge across the product portfolio. In the case of nearly one-third of the products, the number of iterations did not exceed 10. For the UCI dataset, a similar iterative behavior is reported in Table 2. Here, coefficients a and b are fitted with respect to the curve $d_i^t(p_i^t) = \max(a_i - b_i p_i^t, 0)$ and the second problem (revenue maximization) is solved using the same expression. For this case, the sales data had a large fraction of "0" values and product returns, and the choice of $\max(\cdot, 0)$ is the most practical. The total revenue is reported in the last column. Similar results are reported in Table 3 for another set of products.

For the logit dataset, a similar iterative behavior is reported in Table 4. Here, coefficients a and b are fitted with respect to the curve $d_i^t(p_i^t) = \min(a_i - b_i p_i^t, \tilde{a}_i - \tilde{b}_i p_i^t)$ and the revenue maximization problem is solved accordingly. This solves the issue of concavity with respect to maximization of revenue and as we will see later also provides the best possible fit. The choice of the min type PLF for the logit case can be observed to be quite straightforward.

Comparison of models: In synopsis, the three piecewise linear models give a good estimate on the nature of demand for each of our datasets. If a min type function is more accurate, then demand can be possibly more concave. Table 5 shows the accuracy error (mean absolute error) for all the different possible scenarios. It follows that the $\max(\cdot, \cdot)$ fits the Cohen problem finer, and $\min(\cdot, \cdot)$ suits the concave demand problem.

Solution Quality: For the revenue maximization problem, we compare our results with three standard (and rudimentary) settings. One is setting the prices throughout at the base level (upper bound). The second setup refers to a randomly generated price vector between the lower and upper bounds, while the third one refers to the mid point between the lower and upper price bounds. It can be easily noticed from Table 6 that there is a significant improvement in running our algorithm QCASM on the actual nonsmooth optimization problem. At this point, we do not intend to compare the performance of our algorithms against other smooth reformulations or other quadratic optimization methods. We just demonstrate as a proof of concept that the non-

Table 2 Iterative progress of QCASM for UCI's problem

Iteration	Revenue									
	P-263	P-324	P-377	P-441	P-599	P-686	P-756	P-780	P-1060	Total
1	1536.60	3307.20	257.92	792.48	1234.48	1012.94	132.60	1939.60	515.84	5857376.16
10	3119.66	4852.65	425.60	1056.24	2616.66	1903.86	326.01	2726.41	875.51	7692706.11
20	5098.49	6784.47	635.20	1385.94	4344.38	3017.51	567.77	3709.92	1325.10	9987238.74
50	11034.96	12579.92	1264.01	2375.05	9527.54	6358.47	1293.04	6660.46	2673.87	16803284.78
70	11826.49	13352.65	1347.85	2506.93	10218.63	6803.94	1389.75	7053.87	2853.70	17766752.17
80	11826.49	13352.65	1347.85	2506.93	10218.63	6803.94	1389.75	7053.87	2853.70	17798875.94
100	11826.49	13352.65	1347.85	2506.93	10218.63	6803.94	1389.75	7053.87	2853.70	17824970.72
150	11826.49	13352.65	1347.85	2506.93	10218.63	6803.94	1389.75	7053.87	2853.70	17839027.50

smooth problem can be solved efficiently by the QCASM solver. The scheme is also theoretically convergent when demand functions follow some structure on concavity.

Benchmarking and Comparisons We reformulate the primitive nonsmooth optimization problem in (1) into a smooth form by introducing binary variables. However, it is important to note that the resulting formulation may be nonconvex, even if the original problem is convex. We begin by expressing the demand function $d_i^t(p_i^t)$ as:

$$d_i^t(p_i^t) = \max(\tilde{a}_i - \tilde{b}_i p_i^t, a_i - b_i p_i^t).$$

Let $v_i^t = d_i^t(p_i^t)$ denote an auxiliary variable representing demand or sales, and let $\nu_i^t \in \{0, 1\}$ be a binary variable used to identify the active piece in the piecewise linear function. Using a sufficiently large scalar M , we can reformulate the above as a system of inequalities:

$$\begin{aligned} v_i^t &\geq \tilde{a}_i - \tilde{b}_i p_i^t, \\ v_i^t &\geq a_i - b_i p_i^t, \\ v_i^t &\leq \tilde{a}_i - \tilde{b}_i p_i^t + M(1 - \nu_i^t), \\ v_i^t &\leq a_i - b_i p_i^t + M\nu_i^t, \\ \nu_i^t &\in \{0, 1\}. \end{aligned}$$

Here, M is a suitably large constant to enforce the disjunctive logic in the reformulation. In the case where the maximum is replaced by a minimum over the same affine terms, the representation can be modified accordingly. For instance,

$$v_i^t = -d_i^t(p_i^t) = \max(\tilde{b}_i p_i^t - \tilde{a}_i, b_i p_i^t - a_i) = -\min(\tilde{a}_i - \tilde{b}_i p_i^t, a_i - b_i p_i^t),$$

can be written as follows.

$$\begin{aligned} v_i^t &\geq \tilde{b}_i p_i^t - \tilde{a}_i, \\ v_i^t &\geq b_i p_i^t - a_i, \\ v_i^t &\leq \tilde{b}_i p_i^t - \tilde{a}_i + M(1 - \nu_i^t), \\ v_i^t &\leq b_i p_i^t - a_i + M\nu_i^t, \\ \nu_i^t &\in \{0, 1\}. \end{aligned}$$

For our comparisons, we use the standard benchmark, namely Gurobi Optimization (2020). For our experiments in this regard, we adopt the min-type piecewise linear function (PLF) approximation. This choice is primarily motivated by the fact that our theoretical guarantees are strongest in the convex setting, and, as discussed in the introductory section, a significant portion of demand-side models in the literature naturally lead to convex formulations. Although our theoretical framework extends to nonconvex problems-ensuring local optimality-the current implementation of our solver remains at a prototype stage. In contrast, Gurobi is a mature and highly optimized solver that incorporates a suite of algorithms and heuristics spe-

Table 3 Iterative progress of QCASM for UCI's problem (second set)

Iteration	Revenue								
	P-1145	P-1169	P-1201	P-1316	P-1404	P-1433	P-1476	P-1541	P-1913
1	1325.48	187.13	309.40	1505.40	1549.08	1677.52	1491.85	388.96	895.44
10	3816.33	555.90	1068.77	4884.62	2033.86	2001.66	3248.09	1650.53	4273.68
20	6929.90	1016.87	2017.98	9108.65	2639.84	2406.83	5443.38	3227.48	8496.49
50	16270.60	2399.77	4865.62	21780.72	4457.78	3622.35	12029.27	7958.35	21164.91
70	17516.03	2584.16	5245.30	23470.33	4700.18	3784.42	12907.38	8589.13	22854.03
80	17516.03	2584.16	5245.30	23470.33	4700.18	3784.42	12907.38	8589.13	22854.03
100	17516.03	2584.16	5245.30	23470.33	4700.18	3784.42	12907.38	8589.13	22854.03
150	17516.03	2584.16	5245.30	23470.33	4700.18	3784.42	12907.38	8589.13	22854.03

Table 4 Iterative progress of QCASM for the Logit problem

Iteration	Revenue - in multiples of 100,000							
	P-8	P-11	P-21	P-24	P-25	P-33	P-36	P-38
1	13.6753	3.7833	8.4056	7.7593	9.2360	7.5832	9.9388	9.5455
10	14.5872	5.6423	9.3681	9.7020	10.1433	9.3743	11.2419	9.9612
20	15.7272	7.9662	10.5713	12.1304	11.2774	11.6132	12.8707	10.4809
50	19.1471	14.9376	14.1807	19.4154	14.6798	18.3299	17.7572	12.0400
70	21.4270	15.8672	16.5871	20.3867	16.9481	19.2254	18.4088	13.0794
80	22.5669	15.8672	17.7902	20.3867	18.0822	19.2254	18.4088	13.5991
100	24.8468	15.8672	20.1965	20.3867	20.3505	19.2254	18.4088	14.6385
150	25.5308	15.8672	20.9184	20.3867	21.0310	19.2254	18.4088	14.9504

Table 5 Comparison of different piecewise linear functions

Dataset	Mean Absolute Error (Scaled)		
	Model - $\max(., 0)$	Model - $\max(., .)$	Model - $\min(., .)$
Cohen	46.4562	42.7222	47.8897
UCI	19.9365	6.1840	8.6981
Concave-Demand	5.6640	5.6637	0.7258

Table 6 Comparison of different piecewise linear functions

Dataset	Revenue			
	Base price	Random price	Midpoint price	QCASM
Cohen	2440173.82	2970553.95	2846046.30	4774583.20
UCI	5857376.16	15807897.04	12736676.95	17841259.92
Concave-Demand	34268272.32	88747048.32	76545493.35	96117974.13

cifically designed to handle nonconvexity. Additionally, it benefits from its backend being implemented in compiled languages, offering substantial speed advantages. Our solver, by comparison, is written in MATLAB and lacks these low-level performance optimizations at this stage of development. For the $\min(., .)$ -type approxima-

tion, incorporating the associated switching constraints into (1) yields the following optimization formulation:

$$\begin{aligned}
 \min_{p, I, v, \nu} \quad & \sum_{i \in \mathcal{I}} \sum_{t=0}^{T-1} p_i^t \cdot v_i^t \\
 \text{s.t.} \quad & I_i^{t+1} = I_i^t + u_i^t + v_i^t & \forall i \in \mathcal{I}, t = 0, \dots, T-1, \\
 & v_i^t \geq \tilde{b}_i p_i^t - \tilde{a}_i & \forall i \in \mathcal{I}, t \in \mathcal{T}, \\
 & v_i^t \geq b_i p_i^t - a_i & \forall i \in \mathcal{I}, t \in \mathcal{T}, \\
 & v_i^t \leq \tilde{b}_i p_i^t - \tilde{a}_i + M(1 - \nu_i^t) & \forall i \in \mathcal{I}, t \in \mathcal{T}, \\
 & v_i^t \leq b_i p_i^t - a_i + M\nu_i^t & \forall i \in \mathcal{I}, t \in \mathcal{T}, \\
 & \nu_i^t \in \{0, 1\} & \forall i \in \mathcal{I}, t \in \mathcal{T}, \\
 & p_{i,l} \leq p_i^t \leq p_{i,u} & \forall i \in \mathcal{I}, t \in \mathcal{T}, \\
 & I_i^t \geq 0 & \forall i \in \mathcal{I}, t \in \mathcal{T}, \\
 & I_i^T \leq \bar{I} & \forall i \in \mathcal{I}.
 \end{aligned}$$

Note that we have only inventory and bound constraints throughout our numerical results section. As earlier, we proceed with the Logit dataset for the $\min(\cdot, \cdot)$ type approximation. We evaluate algorithmic performance based on two primary criteria: computational time and solution quality. All comparisons are conducted using Gurobi as a benchmark solver. While earlier evaluations of our proposed QCASM method emphasized its iterative behavior, the current experiments apply a time-based stopping criterion to report both solution quality and convergence characteristics. For our computational experiments, we increase the problem size by varying T , the number of time steps, from $T = 52$ to $T = 780$ in increments of 104 (i.e., 2 years per step). The computational budget naturally differs with problem size. For smaller instances, we limit the runtime to a few hundred seconds, while for the largest problems we allow up to 3500 (s), where “(s)” denotes seconds. The variable count in QCASM grows from approximately 10,000 to over 130,000, whereas in Gurobi, the number of variables remains roughly similar across instances-but with about half of them being binary, marking a key distinction. The problem structurally decomposes in the products, and we present the problem in this form to both QCASM and Gurobi for consistency. Results are summarized in Table 7. The third and fourth columns report the objective values (revenues) obtained under comparable computational budgets (shown in the second column). We use `cputime` as a unified metric, recognizing that QCASM is implemented in MATLAB, while Gurobi uses compiled backends in C/C++. Given the fundamentally different algorithmic strategies employed, we consider `cputime` time a more meaningful metric than iteration counts: while QCASM identifies kinks and structural transitions in the continuous space, Gurobi applies advanced Branch-and-Bound (BnB) and Branch-and-Cut (BnC) procedures over a binary-variable-augmented space. Although both solvers exhibit similar asymptotic performance in terms of solution quality, QCASM consistently reaches high-quality solutions faster. It typically begins slower than Gurobi, but then steadily improves,

Table 7 Comparison of QCASM with Gurobi on the Logit dataset for PLFs of the form $\min(.,.)$

Time Horizon	Computational Time (s)	Revenue - in multiples of 100,000 dollars/ units	
		Gurobi	QCASM
T=52	10	961.1295	961.1797
	20	961.1295	961.1797
	30	961.1295	961.1797
	40	961.1295	961.1797
T=156	50	288.3049	288.3539
	100	288.3049	288.3539
	150	288.3049	288.3539
	200	288.3049	288.3539
T=260	60	439.2096	390.2729
	100	439.2096	480.3758
	150	480.2955	480.3758
	300	480.2955	480.3758
	400	480.2955	*
	600	480.2955	*
T=364	800	480.2955	*
	100	614.7760	467.6873
	200	614.7760	504.8708
	300	614.7760	609.4616
	400	614.7760	671.4470
	600	614.7760	*
	800	614.7760	*
T=468	1000	614.7760	*
	100	722.4738	455.0444
	200	722.9352	689.1916
	300	723.4192	715.0395
	400	723.8193	808.9788
	600	724.8932	808.9788
	800	724.8932	*
	1100	724.8932	*
T=572	3500	724.8932	*
	100	883.0291	495.1663
	250	965.5139	678.1153
	400	965.5139	817.3932
	700	965.5139	1030.5200
	900	965.5139	*
	1100	965.5139	*
	1500	965.5139	*
	3500	965.5139	*

Table 7 (continued)

Time Horizon	Computational Time (s)	Revenue - in multiples of 100,000 dollars/ units	
		Gurobi	QCASM
T=676	100	1043.5798	560.8207
	300	1043.5798	752.1873
	500	1043.5798	816.6001
	800	1043.5798	1169.3449
	1200	1043.5798	*
	1500	1043.5798	*
	1800	1043.5798	*
	3500	1043.5798	*
T=780	100	1204.1306	611.0686
	300	1204.1306	787.1123
	500	1204.1306	862.2696
	800	1204.1306	1097.9531
	1200	1204.1306	1270.2441
	1500	1204.1306	*
	1800	1204.1306	*
	3500	1252.4200	*

eventually surpassing Gurobi's performance as dimensionality increases. This difference is rooted in formulation: QCASM evolves in a continuous search space where complexity increases more gracefully, whereas Gurobi operates in a discrete space with rapidly growing combinatorial complexity. For higher values of T , the improvements by Gurobi become marginal. Although it is guaranteed to converge to the global optimum asymptotically, we impose a practical timeout of 3000 (s). For small instances like $T = 52$, Gurobi naturally has an edge, but the advantages of QCASM emerge clearly as the scale increases. In particular, Gurobi's progress stagnates beyond $T = 364$, $T = 468$, and $T = 572$, often showing minimal gain after a decent solution is found. An asterisk "*" in the table denotes instances where QCASM has converged to a global solution, after which no further runs were conducted. A deeper profiling of the solvers' computational characteristics is left for future work. However, these findings already highlight that directly solving the original nonsmooth formulation via QCASM yields significant advantages in convex regimes-compared to reformulating the problem with binary variables and relying on commercial mixed-integer solvers.

Remark Our comparison with Gurobi is currently restricted to the convex setting. For nonconvex problems, our solver QCASM is still in an early stage and requires further development. Firstly, as a MATLAB-based prototype, QCASM lacks the performance benefits that could be gained from a C/C++ implementation-porting to such a compiled language could yield significant speedups. Secondly, Gurobi is a mature and production-grade solver that incorporates a range of advanced techniques, including sequenced algorithms and global optimization heuristics. In contrast, QCASM presently guarantees only local optimality in the nonconvex regime. Consequently, it may not always match Gurobi's solution quality in such settings. That said, Gurobi itself is not immune to challenges-it can also converge to local optima and exhibits scalability

issues in higher dimensions, particularly due to the growing number of binary variables, as seen in earlier tables. At this stage, our intention is simply to demonstrate that QCASM functions reliably on both convex and nonconvex instances, and that it scales well in convex cases, where convergence theory is also satisfied. A deeper treatment of nonconvex scenarios, including improved heuristics and C/C++ portability, is planned for future versions of this work.

5 Conclusion

This work considered nonsmooth quadratic optimization problems in retail revenue management. A continuous variant of the multi-period problem with inventory constraints and price bounds was analyzed and some preliminary statements on convexity were stated. The problem also required the estimation of piecewise linear functions that fit the historical data on price-sales leading to a similar optimization problem. A convergent active signature method that deploys the structure of the nonsmoothness was proposed for both problems in question. For a proof of concept, three practical datasets were studied. While the first two cases originate from real retail data, the third one was simulated from the widely used logit model. Preliminary results demonstrated a good behavior in iterative performance and revenue improvement in comparison to other trivially chosen points. Future work aims to expand on deploying active signature methods for generalizations of our problem involving discrete prices and product design attributes.

Acknowledgements The research was funded partly by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy — The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID:@390685689).

Author contributions All authors worked on designing the algorithm, theoretical analysis, and outlining the numerical results.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allen RGD (1938) Mathematical analysis for economists. English, Macmillan London
- Burman V et al (2021) Markdown pricing for a large scale retailer. In: 2021 IEEE International Conference on Big Data (Big Data). pp. 1817–1826. <https://doi.org/10.1109/BigData52589.2021.9671641>
- Chan CS., Desarbo WS, Harker PT (1990) Product Positioning under Price Competition. *Manage Sci* 36(2):175–199. issn: 00251909, 15265501. <http://www.jstor.org/stable/2661454>
- Chen X, Simchi-Levi D (2004) Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: the finite horizon case. *Oper Res* 52(6):887–896
- Chen X et al (2016) Dynamic stochastic inventory management with reference price effects. *Oper Res* 64(6):1529–1536. <https://doi.org/10.1287/opre.2016.1524>
- Chintagunta PK (2001) Endogeneity and heterogeneity in a probit demand model: estimation using aggregate data. *Mark Sci* 20(4):442–456. <https://doi.org/10.1287/mksc.20.4.442.9751>
- Corsten D, Gruen T (2005) On Shelf Availability: An Examination of the Extent, the Causes, and the Efforts to Address Retail Out-of-Stocks. pp. 131–149. ISBN: 3-540-22611-7. https://doi.org/10.1007/3-540-27059-0_9
- Chiang WK, Chhajer D, Hess J (2003) Direct marketing, indirect profits: a strategic analysis of dual-channel supply-chain design. *Manage Sci* 49:1–20. <https://doi.org/10.1287/mnsc.49.1.1.12749>
- Chen YF, Ray S, Song Y (2006) Optimal pricing and inventory control policy in periodic-review systems with fixed ordering cost and lost sales. *Naval Research Logistics (NRL)* 53(2):117–136. <https://doi.org/10.1002/nav.20127>
- Cohen MC (2015) Pricing for retail, social networks and green technologies. PhD thesis. Massach Inst Tech
- Cohen MC et al (2017) The impact of linear optimization on promotion planning. *Oper Res* 65(2):446–468. <https://doi.org/10.1287/opre.2016.1573>
- Cohen MC, Perakis G, Pindyck RS (2016) Pricing with Limited Knowledge of Demand. In: Proceedings of the 2016 ACM Conference on Economics and Computation. EC '16. Association for Computing Machinery, p. 657. <https://doi.org/10.1145/2940716.2940734>
- Cohen MC, Philip ZR, Jiao K (2019) Data aggregation and demand prediction. *ERN: Stat Decision Theory; Oper Res (Topic)*
- Cohen MC et al (2022) Demand prediction in retail a practical guide to leverage data and predictive analytics. Springer Series in Supply Chain Management, 14
- Chen Daqing Online retail dataset. <https://www.kaggle.com/datasets/lakshmi25npathi/online-retail-dataset/metadata>
- Chen D, Sain LS, Guo K (2012) Data mining for the online retail industry: a case study of RFM model-based customer segmentation using data mining. *J Database Mark Cust Strategy Manag* 19(3):197–208. <https://doi.org/10.1057/dbm.2012.17>
- Dossche M, Heylen F, Van den Poel D (2010) The kinked demand curve and price rigidity: evidence from scanner data. *Scand J Econ* 112(4):723–752. <https://doi.org/10.1111/j.1467-9442.2010.01621.x>
- Deng Y, Li Y, Jeannette Song J-S (2022) A unified parsimonious model for structural demand estimation accounting for stockout and substitution. SSRN
- Eben-Chaime M (2023) Demand Models, Revenue Curves and Profit. *Athens J Bus Econ* 9(2):221–230. <https://doi.org/10.1287/msom.2015.0561>
- Fibich G, Gavious A, Lowengart O (2003) Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effects. *Oper Res* 51(5):721–734. <https://doi.org/10.1287/opre.51.5.721.16758>
- Greenleaf EA (1995) The Impact of Reference Price Effects on the Profitability of Price Promotions. *Mark Sci* 14(1):82–104. issn: 07322399, 1526548X. <http://www.jstor.org/stable/184209>
- Griewank A (2013) On stable piecewise linearization and generalized algorithmic differentiation. *Optim Methods Softw* 28(6):1139–1178. <https://doi.org/10.1080/10556788.2013.796683>
- Griewank A et al (2016) On Lipschitz optimization based on gray-box piecewise linearization. *Math Program* 158:383–415
- Griewank A, Walther A (2016) First- and second-order optimality conditions for piecewise smooth objective functions. *Optim Methods Softw* 31(5):904–930. <https://doi.org/10.1080/10556788.2016.1189549>
- Griewank A, Walther A (2018) Finite convergence of an active signature method to local minima of piecewise linear functions. *Optim Methods Softw*. <https://doi.org/10.1080/10556788.2018.1546856>

- Griewank A, Walther A (2019) Relaxing kink qualifications and proving convergence rates in piecewise smooth optimization. *SIAM J Optim* 29(1):262–289. <https://doi.org/10.1137/17M1157623>
- Griewank A, Walther A (2020) Beyond the oracle: opportunities of piecewise differentiation. *Numerical Nonsmooth Optimization: State of the Art Algorithms*. Ed. by Adil M. Bagirov et al. Cham: Springer International Publishing, pp. 331–361. https://doi.org/10.1007/978-3-030-34910-3_10
- Hegerhorst-Schultchen LC, Steinbach MC (2020) On first and second order optimality conditions for abnormal NLP. *Optim* 69(12):2629–2656. <https://doi.org/10.1080/02331934.2019.1626386>
- Hekimoglu M et al (2019) Assortment optimization with log-linear demand: Application at a Turkish grocery store. *J Retail Consum Serv* 50:199–214. <https://doi.org/10.1016/j.jretconser.2019.04.007>
- Hua J et al (2021) Markdowns in E-Commerce Fresh Retail: A Counterfactual Prediction and Multi-Period Optimization Approach. *CoRR* abs/2105.08313. [arxiv:2105.08313](https://arxiv.org/abs/2105.08313)
- Huang J, Leng M, Parlar M (2013) Demand functions in decision modeling: a comprehensive survey and research directions. *Decis Sci* 44(3):557–609. <https://doi.org/10.1111/dec.12021>
- Karlin S, Carr CR (1962) Prices and Optimal Inventory Policy. *Studies in Applied Probability and Management Science*. Ed. by Kennet J. Arrow, Samuel Karlin, and Herbert Scarf. Stanford (Calif.): Stanford University Press, Chap. 10:159–172
- Kunz T, Crone S (2014) Demand models for the static retail price optimization problem – a revenue management perspective. *OpenAccess Series in Informatics* 37. <https://doi.org/10.4230/OASICS.SICOR.2014.101>
- Kreimeier T (2023) Solving Constrained Piecewise Linear Optimization Problems by Exploiting the Abs-Linear Approach. PhD thesis. Humboldt-Universität zu Berlin
- Kreimeier T, Walther A, Griewank A (2025) Constrained piecewise linear optimization by an active signature method. *Mathematical Optimization for Machine Learning: Proceedings of the MATH+ Thematic Einstein Semester 2023*, De Gruyter, , pp. 93–106.
- Kannan A, Panchamgam K (2020) Computerized promotion and markdown price scheduling. US Patent 10,776,803
- Kannan A, Panchamgam K, Wu S-M (2020) Computerized promotion and markdown price scheduling. US Patent 10,528,903
- Lau AHL, Lau HS (2003) Effects of a demand-curve's shape on the optimal solutions of a multi-echelon inventory/pricing model. *Eur J Oper Res* 147(3):530–548. [https://doi.org/10.1016/S0377-2217\(02\)00291-6](https://doi.org/10.1016/S0377-2217(02)00291-6)
- Li H et al (2022) Large-scale price optimization for an online fashion retailer. *Innovative Technology at the Interface of Finance and Operations: Volume II*. Springer International Publishing, pp. 191–224. https://doi.org/10.1007/978-3-030-81945-3_8
- Ma S, Fildes R (2017) A retail store SKU promotions optimization model for category multi-period profit maximization. *Eur J Oper Res* 260(2):680–692. <https://doi.org/10.1016/j.ejor.2016.12.032>
- Mills ES (1959) Uncertainty and Price Theory. *The Q J Eco* 73(1):116–130. issn: 00335533, 15314650. <http://www.jstor.org/stable/1883828>
- Nakanishi M, Cooper LG (1974) Parameter estimation for a multiplicative competitive interaction model: least squares approach. *J Mar Res* 11(3):303–311. issn: 00222437. <http://www.jstor.org/stable/3151146>
- Optimization G, LLC (2020) Gurobi optimizer reference manual. https://www.gurobi.com/wp-content/plugins/hd_documentations/documentation/9.0/refman.pdf
- Ray S, Li S, Song Y (2005) Tailored supply chain decision making under price-sensitive stochastic demand and delivery uncertainty. *Manage Sci* 51(12):1873–1891. issn: 00251909, 15265501. <http://www.jstor.org/stable/20110474>
- Song Y, Ray S, Li S (2008) Structural properties of buyback contracts for price-setting newsvendors. *Manuf Serv Oper Manag* 10(1):1–18. <https://doi.org/10.1287/msom.1060.0140>
- Scholtes S (2012) Introduction to piecewise differentiable equations. *Springer-Briefs in Optimization*. Springer New York
- Seaman B, Bowman J (2021) Applicability of the M5 to Forecasting at Walmart. *Inter J Forec*. issn: 0169-2070. <https://doi.org/10.1016/j.ijforecast.2021.06.002>
- Skenderi G et al (2021) Well googled is half done: multimodal forecasting of new fashion product sales with image-based google trends. <https://doi.org/10.48550/ARXIV.2109.09824>
- Subramanian S, Sherali H (2010) A fractional programming approach for retail category price optimization. *J Global Optim* 48:263–277. <https://doi.org/10.1007/s10898-009-9491-2>
- Tirole J, Rey P (2019) Price Caps as Welfare-Enhancing Cooperation. *J Polit Econ* 127(6):3018–3069. <https://doi.org/10.1086/702014>

Yaghin GR, SMT Fatemi Ghomi , Torabi SA (2014) Enhanced joint pricing and lotsizing problem in a two-echelon supply chain with logit demand function. *Int J Prod Res* 52(17):4967–4983. <https://doi.org/10.1080/00207543.2014.885665>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Aswin Kannan^{1,2} · Timo Kreimeier^{1,3} · Andrea Walther¹

✉ Aswin Kannan
aswin.kannan@hu-berlin.de; aswin.kannan@iiitb.ac.in

¹ Institut für Mathematik, Humboldt-Universität zu Berlin, Berlin, Germany

² Data Sciences and Artificial Intelligence (and Computer Science), IIIT-Bangalore, Bangalore, India

³ Hannoversche Volksbank eG, Hannover, Germany