Fitting absolutely monotonic real functions in bounded intervals

Juan Casanova

February 22, 2026

This is the abstract.

Contents

1 Introduction

1.1 Absolutely monotonic functions

A real function is *absolutely monotonic* if all of its derivatives are non-negative. Absolutely monotonic functions grow faster and faster as we approach higher values, curving upwards in every sense. They are growing and convex, though not all growing convex functions are absolutely monotonic.

Convex smooth functions, and in particular fitting convex smooth functions from samples, have been studied extensively [Groeneboom et al., 2001, Lu, 2015, Lebair and Shen, 2015, Nagahara and Martin, 2013, Zhang, 2004, He and Shi, 1998, Aguilera and Morin, 2008, Hannah and Dunson, 2013, Holmström and Petersson, Lachand-Robert and Oudet, 2005, Kopotun et al., 2011, Wang and Ghosh, 2012, Ghosal et al., 2023] due to the multitude of applications in economics, statistics, machine learning, etc. Like other types of *shape-constrained estimation*, we know or wish the resulting function to fulfill certain growth or smoothness properties, but also wish to fit the available data as best as possible.

Absolutely monotonic shape-constrained estimation from fixed samples has not been studied as extensively. The majority of methods that ensure absolute monotonicity of the resulting functions do so *by construction*, and usually as a side effect of tighter restrictions. For example, fitting exponential sums [Holmström and Petersson, 2002, Keller and Plonka, 2021, Chunaev and Safiullin, 2020, Evans et al., 1980, Smith et al., 1976, Kammler, 1976]. These also sometimes presume we can select the points to sample on.

1.2 Motivating problem: reward curves

The motivating problem for this paper is the production of *reward curves from scores in a competitive single player game*. Specifically, Beat Saber[1] is a virtual reality rhythm game released in 2018 in which players play single player maps with a maximum possible score, attempting to get as close as possible. Scores in Beat Saber are often represented as percentages of the maximum possible score. Beat Saber has rich scoring mechanics[2], and a healthy competitive scene with upwards of 60,000 active ranked players[3], which can be observed through its two most popular ranked leaderboards:

---

[1]https://beatsaber.com/

[2]https://bsmg.wiki/ranking-guide.html

[3]Verified using BeatLeader's search functionality at https://beatleader.com/ranking/1?mapsType=all&recentScoreTime=1721433600

ScoreSaber[4] and BeatLeader[5]. Competitive Beat Saber focuses on large pools of *custom* maps made by members of the community (*mappers*) that undergo a ranking process[6] to ensure their competitive viability. For example, Beat-Leader offers over 3500 different ranked maps[7].

Experienced Beat Saber players observe a high diminishing returns difficulty curve. The closer to 100%, the harder it is to improve the score further. There are exactly two 100% scores across all ranked maps in Beat Leader (over 3500 maps, over 60,000 players), on particularly easy maps by particularly skilled players. The difference in difficulty between a 98% and a 99% score in a map is orders of magnitude larger than the difference between a 90% and a 91%. Most scores that do not fail the map are over 70% to begin with.

Different maps do not just have different difficulties, but also different difficulty *curves*, with the difficulty accumulating at different relative ranges of the score spectrum, while always preserving the diminishing returns nature. The problem of finding adequate reward curves for each map is difficult and highly sought-after by the Beat Saber ranked community [Casanova, 2025]. Shape-constrained fitting from player scores is a valuable tool in this context. Naturally, a higher score should always give a higher reward (often quantified as Performance Points (PP)), introducing a growing constraint on the reward curve. Moreover, the diminishing returns nature justifies introducing further growth constraints that increase the rate at which the rewards increase. While convexity constraints could be used, there are growing convex curves that do not adequately reflect the ever increasingly difficult nature of approaching 100% score in Beat Saber maps (see, for example, figure 1). Therefore, we instead consider the more constraining problem of fitting smooth absolutely monotonic curves from sample scores.

1.3   Problem statement

As discussed, we consider the problem of, given a single variable **real valued** set of datapoints in a **bounded interval** (we can assume $[0, 1]$ without loss of generality), finding **absolutely monotonic** functions that best approximate the data. We will, however, compare with some methods that only guarantee convex functions due to their prominence and to understand their relative advantages and disadvantages.
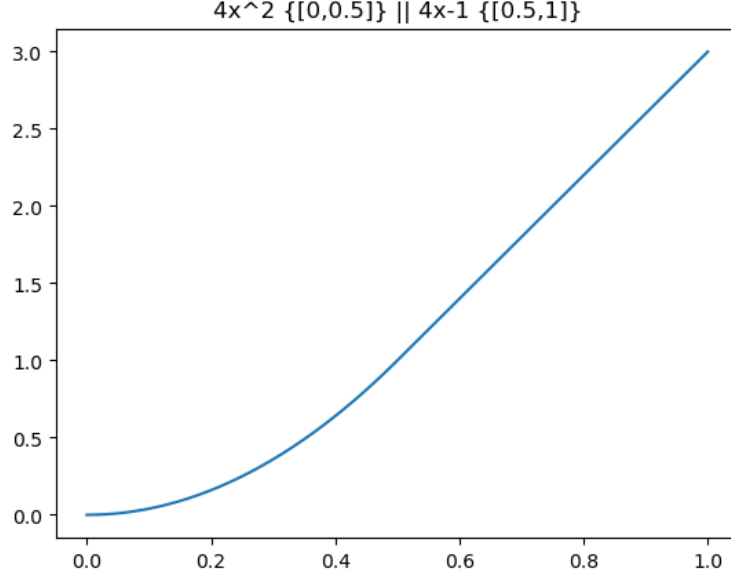
---

[4]https://scoresaber.com/

[5]https://beatleader.xyz/

[6]https://beatleader.wiki/en/ranking/Ranking-your-map

[7]https://beatleader.xyz/leaderboards

Figure 1: Smooth piecewise polynomial that is growing and convex but not absolutely monotonic

We presume noise in the data. In the Beat Saber map context, the x axis indicates the score obtained in the map by the player, and the y axis indicates the reward associated with it, which in the data is represented by the quantified player skill of the player that produced the score. The player skill of a player in a ranked setting represents *their best possible performance*, meaning that better players can sometimes produce worse scores (if not taking the map seriously, being unlucky, or having less than ideal circumstances), but it is far more unlikely that a player will produce a score above their skill level. Therefore, we presume noise in the data will have a strong positive bias (the skill of the player is higher than this score in this map would suggest). In other words, our estimation should ideally favour functions that **underestimate** the data while being as accurate as possible. In order to achieve this, we use asymmetric loss functions in our methods. We consider two different asymmetric loss functions:

1. Pinball loss

$$l_1(r) = \begin{array}{ll} \tau r & \text{if } r \geq 0 \\ (\tau - 1)r & \text{if } r < 0 \end{array} \tag{1}$$

2. Soft floor

$$l_2(r) = |r| + e^{-\tau r} \tag{2}$$

4

One final aspect to consider is how we will use the resulting estimations. Our goal is to use real score data to learn how to best describe the reward curves of different maps. The results of this could then be used to train a different model that looks at the map gameplay properties themselves to estimate those parameters without needing scores. This is useful for newly ranked maps that do not have enough scores yet but need a reward curve. Therefore, we are aiming for a **(semi-)parametric** method that can output a bounded and small number of parameters that represents the fitted curve.

We can summarize these properties in a single problem statement. The problem statement leaves some room for variability to consider the different approaches' advantages and disadvantages.

Problem 1.1. *Given a set of real valued datapoints $\{x_i, y_i\}$, where $x_i \in [0, 1]$, find a function $f \in \mathcal{C}^{\infty}([0, 1])$ that:*

- *is absolutely monotonic in $[0, 1]$.*

- *approximates the datapoints as best as possible with a preference for underestimation.*

- *can be represented (semi-)parametrically with a bounded, low number of parameters.*

The remainder of this document discusses several potential approaches to solving this problem (or slightly relaxed versions of it), discussing their advantages and disadvantages. We also introduce a novel approach, Greedy Lowballing Absolutely Monotonic Polynomial Fitting (GLAMPF), specifically tailored for this problem. We implemented and compared some of these approaches with simulation studies.

Section 2 explains the approaches, with their pros and cons, including the implementation of those that were implemented. Section 3 presents the simulation ande evaluation set up for the comparison of the implemented approaches. Section 4 discusses the results of the evaluation and general discussion on how different methods compare. Section 5 presents the conclusions from this paper.

2 Potential approaches

We discuss a number of reasonable approaches to the presented problem, with their rationales, advantages, and disadvantages. A summary table can be found at the end of this section.

2.1 Positive sums of positive exponentials

Exponential functions with positive exponents are always absolutely monotonic. A deeply related family of functions are *completely monotonic functions*, which are functions with alternating sign derivatives. It is easy to see that $f(x)$ is absolutely monotonic in (a, b) if and only if $f(-x)$ is completely monotonic in $(-a, -b)$. Moreover, Bernstein's theorem characterizes all completely monotonic functions in $[0, +\infty)$, and consequently also all absolutely monotonic functions in $(-\infty, 0]$, as sums of exponentials:

Theorem 2.1 ([Bernstein, 1929]). *Let $f \in \mathcal{C}^\infty([0, +\infty))$. Then, f is completely monotonic if and only if it is the Laplace transform of a non-negative finite Borel measure μ. Explicitly:*

$$f(x) = \int_0^\infty e^{-tx} d\mu(t) \tag{3}$$

or, rephrasing it for absolutely monotonic functions:

Theorem 2.2 ([Bernstein, 1929]). *Let $f \in \mathcal{C}^\infty((-\infty, 0])$. Then, f is absolutely monotonic if and only if there is a non-negative finite Borel measure μ such that:*

$$f(x) = \int_0^\infty e^{tx} d\mu(t) \tag{4}$$

In a few words, any absolutely monotonic function in $(-\infty, 0]$ can be uniquely characterised as an infinite positive sum of exponentials with positive exponents. This also forces f to be analytic because the representation is analytic. The mathematical theory of completely monotonic, absolutely monotonic, and related families of functions is much broader and deeper than what we discuss or use in this paper[Koumandos, 2014].

This is obviously interesting from the point of view of attempting semi-parametric fitting of absolutely monotonic functions: if a function can be expressed as a sum of exponentials, can we select the most significant exponents in this sum to estimate the function? There is one important initial problem with this approach: it only applies to functions that are absolutely monotonic in all of $(-\infty, 0]$. Many absolutely monotonic functions in $[0, 1]$ cannot be extended analytically to $(-\infty, 0]$, and therefore do not admit an exponential sum representation. A simple example of this is monomials $f(x) = x^n$. All of these are absolutely monotonic (and, in fact, analytic) in $[0, 1]$, but cannot be extended to an absolutely monotonic function in $(-\infty, 0]$, and cannot be represented as sums of exponentials.

When attempting to use sums of exponentials to fit datapoints to absolutely monotonic functions, this severely limits the representation power.

Nonetheless, should we accept this limitation, there is extensive literature on the subject of fitting exponential sums.

Prony's method relies on sampling at specific intervals to exploit the structure of an exponential sum[Holmström and Petersson, 2002, Keller and Plonka, 2021, Chunaev and Safiullin, 2020]. This is an effective method for calculating transforms of known functions that can be sampled at prescribed intervals, but is not suitable for fitting a function to random data that does not adjust to this structure and contains noise.

Nonlinear least squares optimization with various suitable setups can be used to approximate a set of parameters (both coefficients and exponents) for an exponential sum that minimizes loss on the datapoints[Holmström and Petersson, 2002, Evans et al., 1980]. This presents challenges such as convergence to local optima (particularly for unusual loss functions), high reliance on the parameterization choice, complex and unpredictable effects of noise in the data, the requirement for the loss function to be smooth, and poor computational scalability to higher number of parameters. It is nonetheless a feasible option for datapoints that can be approximated using exponential sums.

A number of other approaches and variations on the aforementioned approaches for fitting exponential sums exist. All of these share some of the challenges discussed, especially the need for specific sampling, sensitivity to noise, convergence to local optima, and computational issues. For example, [Smith et al., 1976], is outdated, computationally problematic, shows mixed results, and most importantly, requires admitting negative coefficients, losing absolute monotonicity. [Holmström and Petersson, 2002] discusses a wide range of methods for fitting exponential sums. It is possible that the novel greedy lowballing polynomial method introduced in §2.3 can be used with exponential sums as well, but it would incur the aforementioned representation limitation and thus this was not explored.

2.2   Constrained piecewise linear interpolation

While exponential sums ensure absolute monotonicity *by construction*, an opposite approach uses *explicit constraints* to enforce monotonicity. An extreme version of this is to build piecewise linear interpolations that minimize loss with respect to the data points, under, for example, convexity constraints. Convexity can be expressed as optimization constraints by indicating that the second order discrete differences must all be positive:

$$\frac{y_{i+2} - y_{i+1}}{x_{i+2} - x_{i+1}} - \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \geq 0 \tag{5}$$

Despite piecewise linear interpolation having high flexibility allowing low loss fits and being achievable with a wide variety of optimization algorithms (meaning high performance), it has a number of issues. First, higher-order monotonicity constraints involve more and more complex optimization constraints, with strict absolute monotonicity not being possible to express (but rather up to an arbitrary degree of discrete differences). The more higher-order constraints are used, the less the computational attractiveness of the optimization algorithm can be exploited. Moreover, piecewise linear interpolation is only continuous and not differentiable. Finally, this method is inherently nonparametric, with the result being dependent on the number of interpolating points used, meaning that using a small number of parameters will lead to worse fits.

We implemented this method for comparison purposes even if it does not satisfy important conditions of our problem.

2.3 Polynomials

Taylor's theorem is foundational in real analysis and naturally connected to the problem of fitting absolutely monotonic functions. If we restrict our potential functions to analytic functions $\mathcal{C}^\omega$, then the Taylor series at any point of the function explicitly presents all the derivatives of the function at that point. Moreover, if all derivatives at a point are positive and the function is analytic, it is guaranteed that they will only grow, and therefore the function is absolutely monotonic in the whole half-line after that point. We can summarize this in the following lemma, by building analytic functions from their Taylor series at 0:

Lemma 2.1. *An analytic function $f \in \mathcal{C}^\omega([0, +\infty))$ is absolutely monotonic in the positive half-line if and only if it is of the form $\sum_{i=0}^{\infty} a_i x^i$, where every $a_i \geq 0$.*

An important result justifies the restriction to analytic functions to perform approximations on a compact interval $[0, 1]$. An absolutely monotonic function in $[0, 1]$ in particular is absolutely monotonic in $(0, 1)$, and therefore it must be analytic in that open interval. While analyticity may not extend to the endpoints, this is enough to use the Weierstrass theorem to ensure

that the function can be uniformly approximated by absolutely monotonic polynomials. We can summarize all of this in the following lemma:

Lemma 2.2. *A function $f \in \mathcal{C}^\infty([0,1])$ is absolutely monotonic if and only if it can be uniformly approximated by polynomials of the form $\sum_{i=0}^{N} a_i x^i$, where every $a_i \geq 0$.*

In short, we can use non-negative polynomial sums to fit our data, ensuring absolute monotonicity by construction and strong guarantees about approximation capability.

Some important questions remain, and different polynomial fitting methods will be discussed in this section to address these questions. The first is what method to use to obtain good approximation coefficients a_i from the sample without further knowledge about the underlying function (other than absolute monotonicity), especially when the sample is predetermined and the function cannot be sampled at specific points. This also relates to the question of how the fitting method responds to noisy data. Finally, there is the question of how many coefficients are necessary to provide a good approximation of the function and how to provide as good an approximation as possible with as few as possible (or a predefined number of) coefficients.

One thing worth discarding from the beginning is spectral methods. Taylor series are sensitive to small variations and overall lack a stable structure that preserves coefficients and/or degree dominance under approximation circumstances, an issue that grows the further away from 0 that we evaluate the series. From a practical point of view, it is not feasible, especially from sampled noisy data, to reliably reconstruct an exact and/or uniquely defined Taylor series, or to isolate the largest coefficients in the Taylor series. Taylor polynomials with very different coefficient distributions can be very similar pointwise, meaning that there is no way to define a preferred coefficient or degree from data points alone.

2.3.1 Bernstein polynomials

In certain ways the closest approach to spectral methods, Bernstein polynomials offer a promising approach to our problem. Work like [Wang and Ghosh, 2012, Ghosal et al., 2023] present explicit methods to perform shape-constrained regression using Bernstein polynomials from sampled data, with strong approximation and performance guarantees. Specifically, Bernstein basis poly-

nomials are a basis for the vector space of polynomials of a certain degree with important interpolation properties. [Wang and Ghosh, 2012] showed that polynomial regression of a certain degree with shape constraints can be solved by using an adequate matrix representing the shape constraints (monotonicity, convexity, etc.) on the discrete differences applied to the Bernstein basis. With a typical squared residual objective this can then be solved using linear least squares optimization.

However, this has a number of challenges for application to our problem. First, no matrix expression for absolute monotonicity can be found in the literature. Nonetheless, we believe it would be possible to express absolute monotonicity up to the maximum derivative that matters with adequate matrices, but these would likely increase the size of the optimization problem substantially, as it involves explicit constraints rather than absolute monotonicity by construction. More importantly, the method is constrained from inception to a maximum degree of polynomials. In absolutely monotonic fitting, especially when the underlying function is (hyper-)exponential, we may need very high degree polynomials to adequately represent the growth pattern, even for small samples. The Bernstein polynomial linear method requires a pre-determined maximum degree, and the complexity of the optimization problem increases with the degree (so large degrees are not feasible). Finally, this is inherently a nonparametric method that requires a quite large set of coefficients dependent on the degree. The basis is not set up in a way that allows removing some of the elements while preserving a significant proportion of fitting capabilities.

Another approach using Bernstein polynomials would be feasible if the function could be sampled at specific intervals, following Bernstein's original proof of the Weierstrass approximation theorem, but as discussed this is also infeasible for our problem.

2.3.2 Least squares Vandermonde interpolation

A simple yet important to compare to approach uses the Vandermonde matrix over the dataset up to a maximum degree to present a linear optimization problem to find a polynomial estimation. However, in order to enforce absolute monotonicity, non-negativity constraints need to be enforced on the matrix. This makes linear least squares unsuitable to solve the problem due to the severe restriction of the parameter space. Furthermore, the use of a custom loss function in our problem forces the use of nonlinear least squares optimization. Moreover, high degree polynomials are required to represent

absolutely monotonic functions well enough, increasing the complexity of the optimization problem. Perhaps much more importantly, Vandermonde interpolation is numerically unstable and badly represented for high degrees, with high sensitivity to noise or numeric imprecisions during optimization.

We implemented this method for comparison purposes even if it does not satisfy important conditions of our problem.

2.3.3 Convex splines

A popular family of shape-constrained methods discussed extensively in the literature use spline interpolation [Lu, 2015, Lebair and Shen, 2015, Nagahara and Martin, 2013, Zhang, 2004, He and Shi, 1998]. Similar to Bernstein polynomials, these methods use explicit constraints on an optimization problem to ensure monotonicity constraints. They have similar issues to Bernstein polynomials for application to our problem. Absolute monotonicity constraints are not discussed in the literature and would be even harder to implement for spline interpolation, if possible at all. Degree constraints are even harsher than Bernstein polynomial methods, with a degree of 3 used frequently in the literature, being entirely unsuitable for absolutely monotonic functions like (hyper-)exponentials. They are also nonparametric with the number of parameters being highly correlated with the quality of the approximation.

This method is more promising in terms of approximation capability and performance than other methods and relatively easy to implement, so we implemented it with degree 3 and convexity constraints for comparison, even if it does not satisfy all of the conditions of our problem.

2.3.4 Greedy lowballing monomial sums

To the best of our knowledge, there is no significant literature on the use of greedy algorithms for polynomial fitting. To a large degree, this makes sense: the general polynomial fitting problem is unstable and must be solved holistically for satisfactory solutions, lacking any significant stable structure that can be exploited to iteratively build approximations that consistently improve. Moreover, without further constraints, global solvers, including some of the ones discussed so far, work efficiently and effectively, so there is no reason for slower methods with fewer guarantees.

However, in the particular context of the problem presented in this paper, we argue that a greedy approach is attractive. From the point of view of

the (semi-)parametric goal, producing a polynomial one monomial at a time slowly approaching the target allows us to select a limited number of monomials to include in the final solution, reducing the quality of the approximation in a measured way in order to obtain a simpler solution expression. Perhaps more importantly, the absolutely monotonic constraints, together with the asymmetric loss favouring underestimation, lend themselves to an iterative *lowballing* approach, where we slowly add terms, always underestimating the solution by smaller and smaller amounts at each step.

We thus introduce the Greedy Lowballing Absolutely Monotonic Polynomial Fitting (GLAMPF) algorithm. This algorithm is highly specialised to this problem and has significant disadvantages, especially if used for more general tasks, compared to the other methods discussed in this section, but is particularly attractive to our problem in particular. We can summarize its key features as follows:

- The algorithm has two sets of parameters: *degrees* and *coefficients*, which are optimized simultaneously. Their number is not necessarily predetermined but can be limited with measured effects, making the method semi-parametric.

- At each iteration, we add a new monomial with a certain degree and coefficient, of the form $c_j x^{d_j}$, that minimizes the loss with respect to the remainder of the target. The same degree can be added multiple times. After each iteration, we subtract the current monomial from the target value at each point.

- Once an optimal degree and coefficient has been found for a given iteration, we reduce its contribution by a factor $\gamma < 1$ before subtracting it, to ensure the *lowballing*. Together with the asymmetric loss, this ensures that the target is always underestimated, but built towards in steps that best represent the structure of the remainder. That is, we do $^{j+1}y_i = {}^j y_i - \gamma c_j x_i^{d_j}$.

- γ is increased at each iteration, approaching 1 over time, to slowly converge to a full solution.

- We accumulate the coefficients and degrees, and once a certain threshold of approximation or a threshold for small coefficients are surpassed, the algorithm stops. We filter excessively small resulting coefficients and/or limit to a certain number of highest coefficients for an efficient parametric final expression. We note that $0^d = 0$ and $1^d = 1$ for all d,

which means that the coefficients of different degrees are comparable in scale when fitting in the $[0, 1]$ interval.

Among important disadvantages of this approach compared to other discussed ones, we include computational inefficiency due to the greedy nested optimization pattern and dependence on hyperparameter choice and data particularities in the overall result. However, unlike all the other discussed methods, this method simultaneously ensures absolute monotonicity by construction, behaves particularly well with asymmetric loss due to its inherent asymmetric nature, and produces semi-parametric outputs that can easily be limited to a small number of parameters with measured effects.

We implemented this method and consider it the one to best satisfy our particular problem, having potential for use in other variations, while having significant disadvantages outside of its scope.

2.4 Summary

Method	Representation	Monotonicity	Optimality	Hyperparameters	Noise	Loss	Sampling	Parametric	Performance	Complexity
Prony methods §2.1	⚠Exponential only	✓Absolutely monotonic, by construction	✓Yes	✓Irrelevant	✗Highly sensitive	✗Fixed target	✗Fixed	✓Parametric	✓Fast	●Medium
Non-linear least squares exponential sums §2.1	⚠Exponential only	✓Absolutely monotonic, by construction	⚠Local optima	⚠Sensitive	⚠Sensitive	●Flexible, but affected	✓Any	✓Parametric	●Moderate	✓Low
Constrained piecewise linear interpolation §2.2	✗Not differentiable	⚠Convex	✓Yes	✓Irrelevant	●Robust	✓Flexible	✓Any	✗Nonparametric	✓Fast	✓Low
Shape-constrained Bernstein polynomials §2.3.1	⚠Limited degree	⚠Explicit constraints, absolutely monotonic difficult	✓Yes	⚠Sensitive	●Robust	●Flexible, but affected	✓Any	✗Nonparametric	⚠Affected by degree and constraints	⚠Moderate
Precisely sampled Bernstein polynomials §2.3.1	●Selected degree	●Inherited from source function	✓Yes	✓Irrelevant	✗Highly sensitive	✗Fixed target	✗Fixed	✗Nonparametric	●Moderate	●Medium
Vandermonde interpolation §2.3.2	⚠Limited degree	✓Absolutely monotonic, by construction	⚠Local optima	⚠Sensitive	⚠Sensitive	●Flexible, but affected	✓Any	●Semi-parametric	⚠Suboptimal	✓Low
Convex splines §2.3.3	⚠Very limited degree	⚠Convex	✓Yes	⚠Sensitive	●Robust	●Flexible, but affected	✓Any	✗Nonparametric	⚠Affected by degree and constraints	●Medium
Greedy low-balling monomial sums §2.3.4	●Highly flexible, constrained by number of parameters	✓Absolutely monotonic, by construction	⚠No	⚠Sensitive	●Moderately robust	✓Specialised fit	✓Any	●Semi-parametric	⚠Inefficient	✓Low

As can be seen, GLAMPF is the only method that has good representation of the target function family and absolute monotonicity guarantees. The only other method with good qualities in these columns are precisely sampled Bernstein polynomials but these are completely infeasible due to the fixed sampling, nonparametric nature, and no space for an asymmetric loss function. In exchange, our method does not guarantee optimality, is

sensitive to hyperparameters, and is computationally inefficient. These are all issues that we are prepared to accept for our target problem.

Non-linear least squares exponential sums and Vandermonde interpolation with positivity constraints are the only two other methods with acceptable combinations of properties for our problem, and both have significant disadvantages with respect to GLAMPF.

3   Evaluation

Discuss functions that are being tested, evaluation set up.

4   Results and discussion

Show different functions and discuss pros and cons.

Should discuss tail issue with respect to noise in the data.

5   Conclusion

References

[Aguilera and Morin, 2008] Aguilera, N. E. and Morin, P. (2008). Approximating optimization problems over convex functions. *Numerische Mathematik*, 111:1–34.

[Bernstein, 1929] Bernstein, S. (1929). Sur les fonctions absolument monotones. *Acta Mathematica*, 52(1):1–66.

[Casanova, 2025] Casanova, J. (2025). Birating–iterative averaging on a bipartite graph of beat saber scores, player skills, and map difficulties. *arXiv preprint arXiv:2502.19742*.

[Chunaev and Safiullin, 2020] Chunaev, P. and Safiullin, I. (2020). Data fitting by exponential sums with equal weights. In Krzhizhanovskaya, V. V., Závodszky, G., Lees, M. H., Dongarra, J. J., Sloot, P. M. A., Brissos, S., and Teixeira, J., editors, *Computational Science – ICCS 2020*, pages 364–371, Cham. Springer International Publishing.

[Evans et al., 1980] Evans, J. W., Gragg, W. B., and LeVeque, R. J. (1980). On least squares exponential sum approximation with positive coefficients. *Mathematics of Computation*, 34(149):203–211.

[Ghosal et al., 2023] Ghosal, R., Ghosh, S., Urbanek, J., Schrack, J. A., and Zipunnikov, V. (2023). Shape-constrained estimation in functional regression with bernstein polynomials. *Computational Statistics & Data Analysis*, 178:107614.

[Groeneboom et al., 2001] Groeneboom, P., Jongbloed, G., and Wellner, J. A. (2001). Estimation of a convex function: characterizations and asymptotic theory. *The Annals of Statistics*, 29(6):1653–1698.

[Hannah and Dunson, 2013] Hannah, L. A. and Dunson, D. B. (2013). Multivariate convex regression with adaptive partitioning. *J. Mach. Learn. Res.*, 14(1):3261–3294.

[He and Shi, 1998] He, X. and Shi, P. (1998). Monotone b-spline smoothing. *Journal of the American Statistical Association*, 93(442):643–650.

[Holmström and Petersson, 2002] Holmström, K. and Petersson, J. (2002). A review of the parameter estimation problem of fitting positive exponential sums to empirical data. *Appl. Math. Comput.*, 126(1):31–61.

[Kammler, 1976] Kammler, D. W. (1976). Chebyshev approximation of completely monotonic functions by sums of exponentials. *SIAM Journal on Numerical Analysis*, 13(5):761–774.

[Keller and Plonka, 2021] Keller, I. and Plonka, G. (2021). Modifications of prony's method for the recovery and sparse approximation with generalized exponential sums. In Fasshauer, G. E., Neamtu, M., and Schumaker, L. L., editors, *Approximation Theory XVI*, pages 123–152, Cham. Springer International Publishing.

[Kopotun et al., 2011] Kopotun, K., Leviatan, D., Prymak, A., and Shevchuk, I. (2011). Uniform and pointwise shape preserving approximation by algebraic polynomials. *arXiv preprint arXiv:1109.0968*.

[Koumandos, 2014] Koumandos, S. (2014). On completely monotonic and related functions. In Rassias, T. M. and Pardalos, P. M., editors, *Mathematics Without Boundaries: Surveys in Pure Mathematics*, pages 285–321. Springer New York, New York, NY.

[Lachand-Robert and Oudet, 2005] Lachand-Robert, T. and Oudet, E. (2005). Minimizing within convex bodies using a convex hull method. *SIAM Journal on Optimization*, 16(2):368–379.

[Lebair and Shen, 2015] Lebair, T. M. and Shen, J. (2015). Uniform lipschitz property of nonnegative derivative constrained b-splines and applications to shape constrained estimation. *arXiv preprint arXiv:1510.05314*.

[Lu, 2015] Lu, M. (2015). Spline estimation of generalised monotonic regression. *Journal of Nonparametric Statistics*, 27(1):19–39.

[Nagahara and Martin, 2013] Nagahara, M. and Martin, C. F. (2013). Monotone smoothing splines using general linear systems. *Asian Journal of Control*, 15(2):461–468.

[Smith et al., 1976] Smith, M. R., Cohn-Sfetcu, S., and Buckmaster, H. A. (1976). Decomposition of multicomponent exponential decays by spectral analytic techniques. *Technometrics*, 18(4):467–482.

[Wang and Ghosh, 2012] Wang, J. and Ghosh, S. (2012). Shape restricted nonparametric regression with bernstein polynomials. *Computational Statistics & Data Analysis*, 56(9):2729–2741.

[Zhang, 2004] Zhang, J.-T. (2004). A simple and efficient monotone smoother using smoothing splines. *Journal of Nonparametric Statistics*, 16(5):779–796.