



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Exploring Anomalies in Astrophysical Data: a Minimum Enclosing Ball Approach

Optimization for Data Science

Granchelli, Cristian
ID: 2071977

Kelahan, Cameron
ID: 2071947

Orellana Rios, Joan
ID: 2081188

February 2024

Contents

1	Introduction	3
2	Minimum Enclosing Balls	3
2.1	Description of the problem	3
2.2	Optimality Conditions and Conclusions	4
3	Analysis of the algorithms	5
3.1	$(1 + \epsilon)$ -approximation to $\text{MEB}(\mathcal{A})$	5
3.2	The Frank-Wolfe algorithm and its variants	7
3.2.1	The Frank-Wolfe algorithm	7
3.2.2	Away-Steps Frank-Wolfe	7
3.2.3	Pairwise Frank-Wolfe	8
3.2.4	Convergence results	9
4	Line-search methods	9
4.1	Harmonic step-size	9
4.2	Armijo's rule	9
4.3	Exact line search method	10
5	Optimization in Astronomy and Astrophysics	10
5.1	Brief explanation	10
5.2	The Use of Optimization in Astronomy and Astrophysics	10
5.2.1	Image Reconstruction	11
5.2.2	Spectral Object Detection and Denoising	11
5.3	Photometry	11
5.3.1	Photometric Passbands	12

5.4	Spectroscopy	13
5.4.1	Redshift	13
6	Datasets	14
6.1	RR Lyrae Stars	14
6.2	Red and White Dwarf Stars	16
6.3	Quasars and Galaxies	17
7	Results and Discussion	18
7.1	Line search method selection	19
7.2	Anomaly Detection for RR Lyrae Stars	19
7.3	Anomaly Detection for White Dwarfs	21
7.4	Anomaly Detection for Quasars	23
7.5	Discussion	25
8	Conclusion	26

1 Introduction

Data analysis is found in every subject that deals with numbers, making it one of the most significant subjects of our time. Following suit, ensuring the data analysis performance is efficient given physical or temporal constraints is paramount to the process. For this reason, the optimization of data analysis algorithms is crucial to creating modern programs that can handle the ever-increasing complexity of data analysis problems in any field. This paper will define and explore the applications of several state-of-the-art optimization techniques to the Minimum Enclosing Ball (MEB) problem. These techniques include: the Frank-Wolfe (FW) algorithm and an adaptation of the FW algorithm referred to as the $(1 + \epsilon)$ approximation to the MEB described in Yildirim, 2008 [15]; the Pairwise FW Algorithm and the Away-Steps FW algorithm, both described in Lacoste-Julien, 2015 [7].

After explaining the main idea of the MEB problem and its different optimization approaches, the focus of this project will be finding efficient strategies to tackle computational challenges. This exploration will not only reveal the mathematical foundations of these algorithms, but also shed light on their practical applications by employing the listed optimization techniques to the MEB problem across three empirical astronomy and astrophysics datasets. Before diving into these details, the following sections will precisely define the problem and formulate it in a way that is suitable for optimization techniques.

2 Minimum Enclosing Balls

2.1 Description of the problem

The Minimum Enclosing Ball problem can be explained as follows: given a set of points $\mathcal{A} = a_1, \dots, a_n$ in d dimensions, the objective is to identify the smallest ball $B^n(c, \rho)$ that encloses all points of \mathcal{A} , satisfying the condition [15]:

$$\mathcal{A} \subseteq B^n(c, \rho) \tag{1}$$

where

$$B^n(c, \rho) = \{x \in \mathbb{R}^n \mid \|x - c\| \leq \rho\} \tag{2}$$

Here, $c \in \mathbb{R}^n$ and $\rho \in \mathbb{R}$ are the decision variables representing the center and radius of the minimum enclosing ball respectively. A sphere $B^n(c, \rho) \subset \mathbb{R}^n$ is considered an enclosing ball of \mathcal{A} if and only if condition (1) holds.

The problem of approximating the minimum enclosing ball of \mathcal{A} is denoted as $\text{MEB}(\mathcal{A})$, and it can be solved by addressing the following optimization problem:

$$\begin{aligned}
(\text{MEB}) \quad & \min_{c, \rho} \rho \\
\text{subject to} \quad & a_i - c \leq \rho, \quad i = 1, \dots, m
\end{aligned}$$

A second optimization problem, \mathcal{P}_2 , is obtained by defining $\gamma := \rho^2$ and squaring the constraints of MEB:

$$\begin{aligned}
(\mathcal{P}_2) \quad & \min_{c, \gamma} \gamma \\
\text{subject to} \quad & (a_i)^T a_i - 2(a_i)^T c + c^T c \leq \gamma, \quad i = 1, \dots, m
\end{aligned}$$

The Lagrangian dual, denoted as \mathcal{D} , is derived from \mathcal{P}_2 by introducing lagrange multipliers to each constraint:

$$\begin{aligned}
(\mathcal{D}) \quad & \max_u \Phi(u) := \sum_{i=1}^m u_i (a^i)^T a^i - \left(\sum_{i=1}^m u_i (a^i)^T a^i \right)^T \left(\sum_{i=1}^m a^i u_i \right) \\
\text{subject to} \quad & \sum_{i=1}^m u_i = 1, \quad u_i \geq 0, \quad i = 1, \dots, m.
\end{aligned}$$

where $u \in \mathbb{R}^m$ is the decision variable.

2.2 Optimality Conditions and Conclusions

By means of the Karush-Kuhn-Tucker (KKT) optimality conditions, it can be said that $(c_A, \gamma_A) \in \mathbb{R}^n \times \mathbb{R}$ is an optimal solution of \mathcal{P}_2 , given that it is a concave maximization problem with linear constraints, if and only if there exists $u_* \in \mathbb{R}^m$ such that:

$$\sum_{i=1}^m u_i^* = 1, \tag{3a}$$

$$c_A = \sum_{i=1}^m u_i^* a^i, \tag{3b}$$

$$(a^i)^T a^i - 2(a^i)^T c_A + (c_A)^T c_A \leq \gamma_A, \quad i = 1, \dots, m, \tag{3c}$$

$$u_i^* ((a^i)^T a^i - 2(a^i)^T c_A + (c_A)^T c_A - \gamma_A) = 0, \quad i = 1, \dots, m, \tag{3d}$$

$$u^* \geq 0. \tag{3e}$$

A straightforward manipulation of the optimality conditions reveals the relationship:

$$\gamma_A = \Phi(u^*) \quad (4)$$

This indicates that the vector u^* belonging to \mathbb{R}^m serves as an optimal solution for the dual problem (\mathcal{D}) , establishing strong duality between (\mathcal{P}_2) and (\mathcal{D}) . It's worth noting that the center c_A of the minimum enclosing ball of \mathcal{A} is expressed as a convex combination of the elements of \mathcal{A} through equation (3b). Moreover, as indicated by (3d), only the components of u^* corresponding to the points on the boundary of $\text{MEB}(\mathcal{A})$ can take positive values.

After a detailed examination, the outcome in Yildirim [15] results in a conclusion summarized by the following lemma:

Lemma: Let $A = \{a_1, \dots, a_m\}$. The existence and uniqueness of the minimum enclosing ball of \mathcal{A} are established. Denoting the optimal solution of the dual problem \mathcal{D} as $u^* \in \mathbb{R}^m$, we assert that $\text{MEB}(\mathcal{A}) = B_{c_A, \rho_A}$, where:

$$\begin{aligned} c_A &= \sum_{i=1}^m u_i^* a_i, \\ \rho_A &= \Phi(u^*). \end{aligned}$$

This outcome highlights the clear presence and distinctiveness of the minimum enclosing ball of \mathcal{A} , revealing the complex relationship between the best solution u^* and the important parameters c_A and ρ_A .

3 Analysis of the algorithms

3.1 $(1 + \epsilon)$ -approximation to $\text{MEB}(\mathcal{A})$

The first algorithm [15] is an application of the Frank-Wolfe algorithm to the minimum enclosing ball problem.

The algorithm begins by selecting two points, a^α and a^β , that are initially the furthest apart in the set \mathcal{A} . The weights u_i^0 are initialized such that all points have zero weight, except u_α^0 and u_β^0 , which are set to $1/2$ each. The set χ_0 , that represents the initial working core set, is initialized with the points a^α and a^β . The center c^0 is computed as a weighted sum of points based on the initialized weights. The initial value of γ^0 is calculated using the function $\Phi(u^0)$, which represents the objective value associated with the initial weights. The algorithm then identifies the index κ of the point furthest from the current center c^0 . The initial value of the convergence criterion δ_0 is computed, representing the ratio of the squared distance to the initial objective value. The iterative loop begins, aiming to refine the solution until a convergence criterion is met. Within each iteration, the step-size λ^k is computed based on the current value of δ_k . The weights u^k and the center c^k are updated accordingly, incorporating the step-size and the identified point κ . The set of points on the boundary, χ_k , is updated by including the newly identified point a^κ . The updated value of γ^k is computed using the function $\Phi(u^k)$. The index κ is identified again, this time considering the point furthest from the updated center

c^k . The value of δ_k is updated, representing the ratio of the squared distance to the current objective value. The loop continues until the convergence criterion is met, i.e., $\delta_k \leq (1 + \epsilon)^2 - 1$. Once the convergence is achieved, the final output is provided. This includes the converged center c^k , the working core set χ_k , the weights u^k , and the radius of the ball, $\sqrt{(1 + \delta_k)\gamma^k}$.

Algorithm 1 $(1 + \epsilon)$ -approximation to MEB(\mathcal{A})

Require: Input set of points $\mathcal{A} = \{a^1, \dots, a^m\} \subset \mathbb{R}^n, \epsilon > 0$;

- 1: $\alpha \leftarrow \arg \max_{i=1, \dots, m} \|a^i - a^1\|^2$; $\beta \leftarrow \arg \max_{i=1, \dots, m} \|a^i - a^\alpha\|^2$;
 - 2: $u_i^0 \leftarrow 0, \quad i = 1, \dots, m$;
 - 3: $u_\alpha^0 \leftarrow 1/2, \quad u_\beta^0 \leftarrow 1/2$;
 - 4: $\chi_0 \leftarrow \{a^\alpha, a^\beta\}$;
 - 5: $c^0 \leftarrow \sum_{i=1}^m u_i^0 a^i$;
 - 6: $\gamma^0 \leftarrow \Phi(u^0)$;
 - 7: $\kappa \leftarrow \arg \max_{i=1, \dots, m} \|a^i - c^0\|^2$;
 - 8: $\delta_0 \leftarrow (\|a^\kappa - c^0\|^2 / \gamma^0) - 1$;
 - 9: $k \leftarrow 0$;
 - 10: **while** $\delta_k > (1 + \epsilon)^2 - 1$ **do**
 - 11: **loop**
 - 12: $\lambda^k \leftarrow \delta_k / [2(1 + \delta_k)]$;
 - 13: $k \leftarrow k + 1$;
 - 14: $u^k \leftarrow (1 - \lambda^{k-1})u^{k-1} + \lambda^{k-1}e^\kappa$;
 - 15: $c^k \leftarrow (1 - \lambda^{k-1})c^{k-1} + \lambda^{k-1}a^\kappa$;
 - 16: $\chi_k \leftarrow \chi_{k-1} \cup \{a^\kappa\}$;
 - 17: $\gamma^k \leftarrow \Phi(u^k)$;
 - 18: $\kappa \leftarrow \arg \max_{i=1, \dots, m} \|a^i - c^k\|^2$;
 - 19: $\delta_k \leftarrow (\|a^\kappa - c^k\|^2 / \gamma^k) - 1$;
 - 20: **end loop**
 - 21: Output $c^k, \chi_k, u^k, \sqrt{(1 + \delta_k)\gamma^k}$.
-

At each iteration, the algorithm is actually building a trial ball with center c^k and radius $\sqrt{\gamma^k}$ which only match the MEB of \mathcal{A} if u^k is an optimal solution of the problem. That occurs when the stopping criterion is met. Otherwise, there will always be a point in \mathcal{A} not belonging to the ball. Notice how at each iteration the trial ball is expanded by a factor of $\sqrt{1 + \delta_k}$, since:

$$\phi(u^k) \leq \Phi(u^*) \leq (1 + \delta_k)\Phi(u^k) \quad (5)$$

The new center c^{k+1} is calculated by shifting the previous center towards its furthest point a^κ , which is added to the core set.

This algorithm is the adaptation of the Frank–Wolfe algorithm to the dual problem \mathcal{D} . At each iteration, the quadratic objective function $\Phi(u)$ of (\mathcal{D}) is linearized at the current feasible solution u_k . Since the feasible region of (\mathcal{D}) is the unit simplex, the unit vector e_κ , where κ is the index of the furthest point in A from c_k , solves the linearized subproblem. It is easy to verify that

$$\lambda^k = \arg \max_{\lambda \in [0,1]} \Phi((1 - \lambda)u^k + \lambda e^\kappa).$$

It is worth noting that the algorithm relies on the first-order approximation of the objective function Φ . Consequently, each iteration is relatively cheap, but the overall number of iterations

tends to be considerably higher compared to other algorithms that use second-order information, such as interior-point methods. Nevertheless, these general-purpose algorithms become computationally impractical for larger problems due to the significantly higher cost associated with each iteration.

From the analysis of the algorithm in Yildirim, it's worth noting that the algorithm computes a $(1+\epsilon)$ -approximation to the Minimum Enclosing Ball (MEB) in a relatively efficient manner. The algorithm's complexity is $O(mn/\epsilon)$ operations, and it returns an ϵ -core set of size $O(1/\epsilon)$.

3.2 The Frank-Wolfe algorithm and its variants

The Frank-Wolfe algorithm is one of the first algorithms developed to solve constrained convex optimization problems. It has been proven that its variants converge linearly for any strongly convex function optimized over a polytope, with a constant that only depends on the geometry of the polytope [7].

Setup: We consider the following problem:

$$\min_{x \in M} f(x) \quad M = \text{conv}(A) \quad LMO_A(r) \in \arg \min_{x \in A} \langle r, x \rangle$$

where $A \subseteq \mathbb{R}^n$ is a finite set of vectors called atoms, $f(x)$ is a μ -strongly convex function with L -Lipschitz continuous gradient over M .

3.2.1 The Frank-Wolfe algorithm

The Frank-Wolfe algorithm works by finding a feasible direction to move towards by minimizing a linear approximation of the objective function over the feasible set. The next iterate is obtained by doing a line-search on the function between the previous iterate and the search atom. An advantage of this approach is that it produces sparse iterates: at each iteration, the current point can be represented as a sparse convex combination of, at most, $t+1$ directions or “atoms”.

Algorithm 2 Frank-Wolfe Algorithm

- 1: Choose a point $x^{(1)} \in M$
 - 2: **for** $t \leftarrow 0$ to T **do**
 - 3: Set $\hat{x}^{(t)} = \text{Argmin}_{x \in M} \nabla f(x^{(t)})^T (x - x^{(t)})$
 - 4: Set $x^{(t+1)} = x^{(t)} + \alpha_t (\hat{x}^{(t)} - x^{(t)})$, with $\alpha_k \in (0, 1]$
-

Zig-Zagging Phenomenon When the optimal solution lies at the boundary of the feasible set, the convergence rate is slow (sublinear). This happens because the iterates start to zig-zag between the vertices of the face where the solution lies.

3.2.2 Away-Steps Frank-Wolfe

In order to overcome the zig-zagging problem, we can rely on a variant of the Frank-Wolfe algorithm called Away-Steps Frank-Wolfe. This variant is based on the idea of allowing us

to move away from the worst atoms. Specifically, the algorithm computes both the Frank-Wolfe direction and the Away-Steps direction, which maximizes the potential descent given by $g_t^A := \langle -\nabla f(x^{(t)}), x^{(t)} - v_t \rangle$, and then uses the one that gives the largest decrease in the objective function to compute the update. It's important to note that the search for the worst atom is conducted within the active set S_t , which is usually small in size. This makes the search operation substantially easier than the linear oracle LMO_A .

When using the away-steps direction, we need a parameter that regulates the maximum step-size to ensure that the next step will be in the feasible set. Computing the true maximum feasible step-size would require to know when we cross the boundary of the feasible set, which is generally not possible. Instead, the algorithm implements a conservative maximum step-size that does not require a more powerful oracle. This is the reason why the AS variant is required to maintain the set S_t .

Finally, the FW gap g_t^{FW} is an upper bound on the unknown sub-optimality, and can be used as stopping criterion:

$$g_t^{FW} := \langle -\nabla f(x^{(t)}), d_t^{FW} \rangle \geq \langle -\nabla f(x^{(t)}), x^* - x^{(t)} \rangle \geq f(x^{(t)}) - f(x^*)$$

Algorithm 3 Away-Steps Frank-Wolfe Algorithm

```

1: Let  $x^{(0)} \in A$ , and  $S^{(0)} := \{x^{(0)}\}$ 
2: for  $t \leftarrow 0$  to  $T$  do
3:   Let  $s_t := LMO_A(\nabla f(x^{(t)}))$  and  $d_t^{FW} := s_t - x^{(t)}$ 
4:   Let  $v_t \in \arg \max_{v \in S^{(t)}} \langle \nabla f(x^{(t)}), v \rangle$  and  $d_t^A := x^{(t)} - v_t$ 
5:   if  $g_t^{FW} := \langle -\nabla f(x^{(t)}), d_t^{FW} \rangle \leq \epsilon$  then return  $x^{(t)}$ 
6:   if  $\langle -\nabla f(x^{(t)}), d_t^{FW} \rangle \geq \langle -\nabla f(x^{(t)}), d_t^A \rangle$  then
7:      $d_t := d_t^{FW}$  and  $\gamma_{max} := 1$ 
8:   else
9:      $d_t := d_t^A$  and  $\gamma_{max} := \alpha_{v_t} / (1 - \alpha_{v_t})$ 
10:  Line-search:  $\gamma_t \in \arg \min_{\gamma \in [0, \gamma_{max}]} f(x^{(t)} + \gamma d_t)$ 
11:  Update  $x^{(t+1)} := x^{(t)} + \gamma_t d_t$ 
12:  Update  $S^{(t+1)} := \{v \in A \text{ s.t. } \alpha_v^{(t+1)} > 0\}$ 

```

3.2.3 Pairwise Frank-Wolfe

The Pairwise Frank-Wolfe algorithm modifies the Frank-Wolfe step: instead of moving towards a new vertex, it adjusts the weights between two vertices of the current solution.

The Pairwise step is obtained by moving weight from the away atom to the Frank-Wolfe atom, keeping all the others unchanged. The convergence rate guarantee for the Pairwise Frank-Wolfe (PFW) algorithm is not as strong as for other variants. Despite this, the algorithm works well in practice, often outperforming the Away-Steps variant, especially in the case of sparse solutions. Pairwise FW effectively maintains the sparsity of the solution by only adjusting the weights between two vertices. This allows it to avoid the introduction of additional vertices into the solution.

The main difference between the Pairwise and the Away-Steps algorithms is that the former

only moves mass onto FW atoms, while the latter moves it back uniformly onto all other active atoms.

Algorithm 4 Pairwise Frank-Wolfe Algorithm

- 1: as in the Away-Steps algorithm, except replacing line 6 to 10 by: $d_t = d_t^{PWF} := s_t - v_t$,
and $\gamma_{max} := \alpha_{v_t}$
-

3.2.4 Convergence results

The FW variants converge at a linear rate. For the AFW and PFW we define a ‘drop step;’ a step after which the active set shrinks. For the PFW, we define a ‘swap step;’ a step where $\gamma_t = \gamma_{max}$, but the active set remains of the same size. Moreover, we define a ‘good step’ each step that is not a drop step or a swap step. Under some conditions, f has L-Lipschitz gradient and it is μ -strongly convex over M, the sub-optimality of the iterates of the FW variants decreases geometrically at each good step. Specifically, let $k(t)$ be the number of good steps up to iteration t ; $k(t) \geq \frac{t}{2}$ for AFW, and $k(t) \geq \frac{t}{3|A|!+1}$ for PFW. Furthermore, the duality gap also reduces linearly, if f has L-Lipschitz gradient over the diameter of M.

4 Line-search methods

The variants of the Frank-Wolfe algorithm require conducting a line search during each iteration to determine an appropriate step-size. There are multiple methods to obtain a suitable step-size, α .

4.1 Harmonic step-size

A simple and straightforward way to obtain a decreasing step-size is to use the Harmonic step-size method. This method simply sets the step-size $\alpha = \frac{2}{t+1}$, where t is the number of iterations. The step-size in this case depends solely on the iteration.

4.2 Armijo’s rule

Armijo’s rule is an inexact line search method that aims to ensure sufficient decrease in the objective function, while avoiding overly small step-sizes. The process starts by choosing an initial value Δ for the step-size, then we need to set parameter $\delta \in (0, 1)$, which is the reduction factor used to iteratively decrease the step-size during the line search process, if the sufficient decrease condition is not met with the current step-size, the step-size is multiplied by δ to reduce it. This process is repeated until the sufficient decrease condition is satisfied. Parameter $\gamma \in (0, 1/2)$ determines what is considered a sufficient decrease in the objective function for a given step-size. We set $\Delta = 0.9$, $\delta = 0.5$, $\gamma = 0.1$.

4.3 Exact line search method

The exact line search method finds the step-size that minimizes the objective function in a direction, by solving the problem $\alpha^* = \operatorname{argmin}_{\alpha > 0} f(x + \alpha d)$. In our case the direction was given by either the Away-Steps direction or the Frank-Wolfe direction for the Away-Steps FW algorithm, depending on which one was chosen and the Pairwise direction for the Pairwise algorithm. Since the objective function was quadratic, the value for α was obtained by computing $-\frac{d^T \nabla f(x)}{d^T Q d}$, where $Q = 2AA^T$ and d is the direction.

5 Optimization in Astronomy and Astrophysics

For the purpose of exemplifying these algorithms in real-world scenarios, each algorithm will be applied to the MEB problem across three datasets within the field of astronomy. Modern-day and near-future telescopes and observatories create massive amounts of data that need to be processed and analyzed. The observations that led to the first image of a black hole taken by the Event Horizon Telescope produced five petabytes of data over the course of only a few days. Understanding how to utilize optimization techniques within the fields of Astronomy and Astrophysics will become a necessity to perform scientific research in a reasonable time frame.

5.1 Brief explanation

Astronomy is a natural science that studies astronomical objects of the sky, such as stars, planets, their natural satellites, asteroids, comets, and more. Observational technology, including telescopes and observatories, play a crucial role in capturing data from these celestial entities. Moreover, astronomy delves into the study of various astronomical phenomena related to these objects, such as supernovas, quasars, and other cosmic events.

Astrophysics, a specialized field within astronomy, goes beyond simple observation. It focuses on the application of physics principles to understand the behaviour of those celestial objects. This involves an in-depth exploration of the physical properties exhibited by them, offering insights into their composition, structure, and behavior.

5.2 The Use of Optimization in Astronomy and Astrophysics

In the field of astrophysics, gathering information is frequently hindered by different disturbances that distort the data obtained through observations. These disturbances include atmospheric effects, limitations in telescope resolution, and instrumental noise. Additionally, the data may not directly reveal the desired information, such as the internal structure of pulsating stars. To estimate parameters like age, radius, and effective temperature, it is often necessary to infer characteristics through the combination and analysis of observable features and compare empirical data with numerical models.

Optimization proves to be a powerful tool in overcoming these challenges encountered in

various observational practices in astrophysics. Given the complexity of extracting astronomical data, where supervision is crucial, employing objectives and constraints becomes an effective strategy. Consequently, optimization plays a central role in addressing numerous astronomical challenges. [2] In this report, our main objective is to utilize both spectral and photometric information to detect and distinguish celestial objects. We explore the optimization methods used on spectral and photometric data with the goal of improving the efficiency, accuracy, and dependability in identifying different astronomical entities.

5.2.1 Image Reconstruction

Astronomical images obtained through telescopes often suffer from blurring, requiring deconvolution for better resolution. This process involves solving optimization problems with positivity constraints using algorithms like ISRA and Richardson-Lucy.

Radio-astronomy utilizes algorithms like CLEAN for image reconstruction, and recent advancements involve convex optimization techniques with sparsity regularization for improved results. [2]

5.2.2 Spectral Object Detection and Denoising

Observations in astronomy often involve measuring the full spectrum of objects. Spectral denoising, crucial for high-resolution sources, is formulated as an optimization problem. Techniques like sparse modeling in the Discrete Cosine Transform basis are employed, addressed through optimization with proximal methods. [2]

5.3 Photometry

Photometry is a branch of astronomy that involves the measurement of the brightness of celestial objects across the electromagnetic spectrum without considering specific wavelengths. It is a fundamental technique used to quantify the amount of light emitted or reflected by stars, galaxies, and other astronomical objects. Photometric measurements are crucial for understanding the physical properties, composition, and behavior of celestial bodies. These measurements are often obtained using specialized instruments and standardized photometric systems, allowing astronomers to compare and analyze the brightness of objects in a consistent and meaningful way [1].

Photometric systems are standardized methods used to measure the brightness of celestial objects at different wavelengths. They consist of specific passbands, which are defined wavelength ranges where measurements are made, and are collected by applying filters to isolate the desired range (figure 1). The measurements obtained through photometric systems allow astronomers to determine the flux or brightness of celestial objects in a standardized and systematic manner. Different photometric systems may have varying passbands and calibration methods, but they all serve the purpose of quantifying the brightness of astronomical objects across the electromagnetic spectrum. In our case, we will partially be focusing on the use of

the photometric spectrum of celestial objects for classification, applying the aforementioned algorithms. As variables for this algorithm, we will use differences between photometric passband values, known as color indices.

5.3.1 Photometric Passbands

The datasets used for this study are extracted from the Sloan Digital Sky Survey (SDSS) database. All the data extracted from this database follows the UGRIZ system. The letters u, g, r, i, and z are used to denote different photometric passbands for this specific photometric system:

- **Ultraviolet Passband:** U-band
- **Visible Passbands:** G-band (green), R-band (red)
- **Near-infrared Passbands:** I-band, Z-band

Effective Wavelength midpoints, denoted as λ_{eff} , refer to the central or average wavelengths of the passbands used in photometric systems. These midpoints represent the effective or characteristic wavelength at which the measurements are made within a specific passband.

Band	Effective Wavelength Midpoint (λ_{eff})
U	365 nm
G	464 nm
R	658 nm
I	806 nm
Z	900 nm

Table 1: Photometric Bands and Effective Wavelength Midpoints. [12]

Figure 1 represents the spectral response of the passbands used in the Sloan Digital Sky Survey (SDSS) photometric system. It shows the sensitivity of the u, g, r, i, and z passbands to different wavelengths of light. The normalized response indicates how efficiently each passband detects light at specific wavelengths, providing a visual representation of the system's spectral sensitivity.

Differences between passbands, such as (u-g) and (g-r), are used for object classification in astronomy because they provide valuable information about the spectral energy distribution and color properties of celestial objects. Figure 2 demonstrates the information provided by these color indices. The color indices are calculated by taking the differences between the flux or magnitudes measured in different passbands, essentially identifying steep and flat slopes of the objects spectrum that locate the peak of the observed spectrum. This offer insights into the temperature, metallicity, and other physical characteristics of stars, galaxies, and other astronomical objects. For example, the (u-g) color index can be used to distinguish between different types of stars, such as hot, blue stars and cooler, redder stars. Similarly, the (g-r) color index can provide information about the metallicity and age of stars and galaxies.

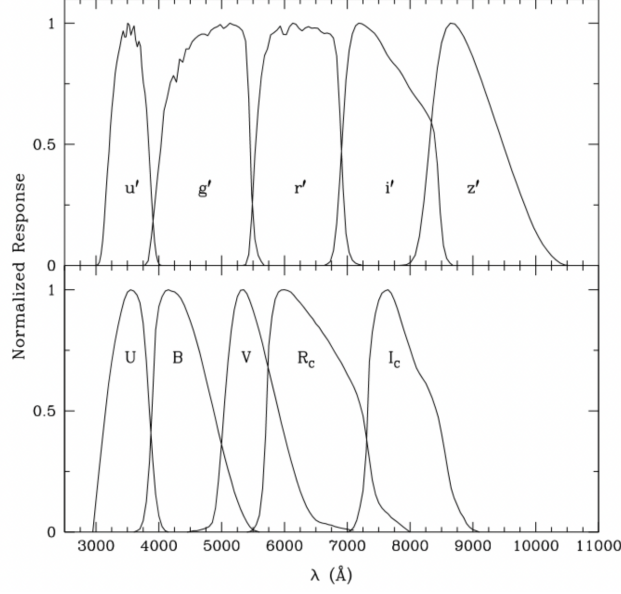


Figure 1: Spectral response of the passbands.

5.4 Spectroscopy

In comparison with photometry, spectroscopy measures the intensity of light as a function of the independent wavelength, allowing for precise analysis of spectral features found in spectra. Elements absorb light at specific wavelengths, causing spectra to act as fingerprints for celestial objects, offering insight to their composition, temperature, and even motion. For ease, the SDSS database provides the spectrum projected onto the filters of the UGRIZ photometric system, offering a snapshot of the spectra without having to read all measured wavelengths. This information is then used in similar fashion to the photometry outlined above, creating estimated color-indices normally only available from photometric observations.

5.4.1 Redshift

On top of the UGRIZ system utilized in all of the datasets, the Galaxy and Quasar dataset utilizes ‘redshift’ as a variable. Redshift describes the ‘shifting’ of an object’s spectrum towards longer (redder) wavelengths. This phenomenon is normally caused by the ‘Doppler Effect,’ a shift in the wavelength of a wave in relation to an observer because the source is moving towards or away from the observer. A classic example of this is the compression and expansion of sound waves as an ambulance is approaching and then driving away from you, manifesting in the different pitches you hear. In the case of galaxies and quasars that are extremely far away, the redshift is mostly caused by the expansion of the Universe. Faraway objects are moving away from us, causing observations performed on or around Earth of these distant objects to produce flux values ‘shifted’ towards the red-end of the electromagnetic spectrum (see figure 3). By nature of the expansion of the Universe, the farther away an object is, the older it is. Consequently, the older objects have been affected by the increasing rate of expansion of the Universe for a longer time, meaning its relative motion away from us will be greater than nearby objects, and the more pronounced the redshift will be. This makes redshift an important distance metric across the Universe.

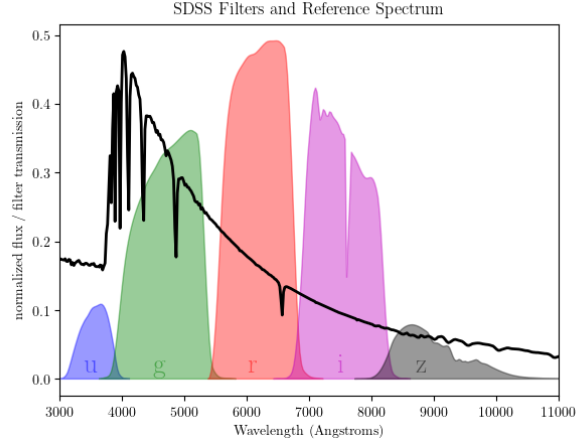


Figure 2: UGRIZ passbands overlaid on the spectrum of standard star Vega [11]. In this example, calculating the (g-r) color index by subtracting the peak flux value of the r passband from the peak flux value of the g passband yields a larger (steeper) and negative slope between these wavelengths than compared to the (i-z) color index, giving insight to where the peak of the spectrum sits.

6 Datasets

The chosen datasets revolve around important, sought-after celestial objects including the RR Lyrae Stars dataset from the astroML [11] library (originally taken from the SDSS database) utilized by Canu et al. 2016 [2] for their implementation of a modified MEB, a dataset of red and white dwarf star observations from the Sloan Digital Sky Survey: Data Release 18 [4], and a collection of observed galaxies and quasars from the Sloan Digital Sky Survey: Data Release 18 [4]. In this paper, we show how optimization algorithms can be applied to improve efficiency of methods that detect these important objects.

6.1 RR Lyrae Stars

RR Lyrae stars serve a great purpose in astronomical research, improving our understanding of both the Milky Way galaxy and the Universe. RR Lyrae stars differ from most other star classifications due to the considerable variability of their brightness over short time periods (hours to days, see figure 4). To standardize the brightness measurements of stars, astronomers use two main measurements: apparent magnitude (m) and absolute magnitude (M). Apparent magnitude measures the brightness of an object as seen from Earth, while absolute magnitude standardizes this brightness measurement for all objects by calculating the brightness of said object from a distance of 10 parsecs. For most objects, apparent magnitude is the only magnitude measurement astronomers have the ability to collect, unless the object happens to be 10 parsecs away. Thanks to extensive observations of these stars over the last century, astronomers discovered a strong relationship between the period of variation of an RR Lyrae star and its intrinsic luminosity. Because of this, astronomers can use the apparent magnitude measurements to calculate the absolute magnitude using the formula

$$M = a * \log_{10}(P) + b$$

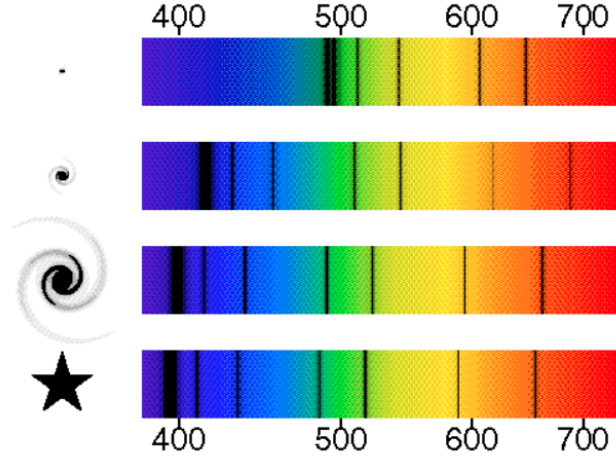


Figure 3: Example of how spectral features (e.g. the black absorption lines) are affected by redshift. The ‘star’ spectrum (the constant used for comparison; also often laboratory measurements) is the reference spectrum. The ascending spectra of the galaxies represent galaxies that are moving away from the observer at faster and faster rates, causing a larger red-shift in the spectral features [14].

where P is the pulsation period of the star and a and b are coefficients determined from observations and analysis of RR Lyrae stars. Obtaining both magnitude measurements grants the ability to measure the distance to the star using the distance modulus

$$m - M = 5 * \log_{10}(d) - 5$$

rearranged as

$$d = 10^{(m-M+5)/5}$$

Normally, measuring absolute magnitude and distance are not so simple, making RR Lyrae stars an important ‘standard candle’ for astronomers to use as accurate distance reference points as they map out the structure of the galaxy and the Universe.

The training parameters used for our experiments on the RR Lyrae dataset include the (u-g) and (g-r) color indices of the UGRIZ photometric system (see figure 7). The dataset consist of 4517 standard star and 483 RR Lyrae star data points, totaling 5,000 examples.

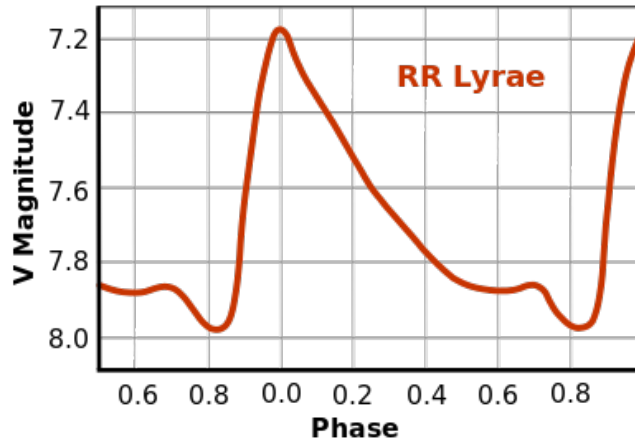


Figure 4: Example of the repeating variability of RR Lyrae stars over a fixed period of time ($P = 0.643$ days) [13]

6.2 Red and White Dwarf Stars

One specific subset of the SDSS dataset focuses on red and white dwarfs, two distinct types of stars with unique characteristics. Red and white dwarfs play crucial roles in our understanding of stellar evolution and contribute valuable insight to the field of astronomy.

Red Dwarf Stars

Red dwarfs are low-mass stars that make up a significant portion of the stellar population. They are known for their relatively cool temperatures and extended lifespans [9]. Despite their commonality, red dwarfs are often challenging to observe due to their faintness. The SDSS dataset on red dwarfs contains information on their spectral characteristics, luminosities, temperatures, and their positions in the sky.

For red dwarfs, in the UGRIZ system, the color index that corresponds to the peak emission would be the ‘r-i’ or ‘i-z’ color indices [3]. This is because the ‘r’ band corresponds to the red part of the visible spectrum, and the ‘i’ and ‘z’ bands correspond to near-infrared parts of the spectrum. Since red dwarfs emit the most in the red and near-infrared parts of the spectrum, as shown in figure 5, the ‘r-i’ and ‘i-z’ color indices would capture this peak emission for red dwarfs in the UGRIZ system.

Two sub-classes of red dwarfs exist in this dataset: K and M type red dwarf stars. The nomenclature stems from the Morgan–Keenan stellar classification system, binning stars by their temperature into one of the following groups: O, B, A, F, G, K, M; with O being the hottest stars and M being the coolest stars. For reference, the Sun is a G type star with a temperature of just under 6,000 Kelvin. K and M stars are much cooler with temperatures around 3,500 and 3,000 Kelvin respectively. The inherent differences between K and M type red dwarfs are noticeable in the data, presenting as partially separate vertical and horizontal clusters when considering certain parameters, as shown in the Results and Discussion section of this paper (see figure 10).

White Dwarf Stars

White dwarfs, on the other hand, represent the endpoint of stellar evolution for stars with low to intermediate masses. These stars have exhausted their nuclear fuel and undergone gravitational collapse, resulting in a highly dense, Earth-sized remnant. White dwarfs exhibit high surface temperatures, and their study provides valuable insights into the late stages of stellar evolution.

In the UGRIZ system, the color index that corresponds to the peak emission of a white dwarf would be the ‘u-g’ color index [5]. This is because the ‘u’ band corresponds to the ultraviolet part of the spectrum, and the ‘g’ band corresponds to the green part of the visible spectrum, as shown in figure 1. Since white dwarfs are hotter stars, they emit the most in the bluer part of the spectrum, the ‘u-g’ color index would capture this peak emission for white dwarfs in the UGRIZ system.

The training parameters used for our experiments on the red and white dwarf dataset include the (u-g), (g-r), (r-i), and (i-z) color index projections of the stellar spectra (see figure 10 for the parameter grid plots). The final dataset consists of about 2,000 red dwarf and 2,000 white dwarf examples directly sampled from the SDSS: Data Release 18 dataset with a random

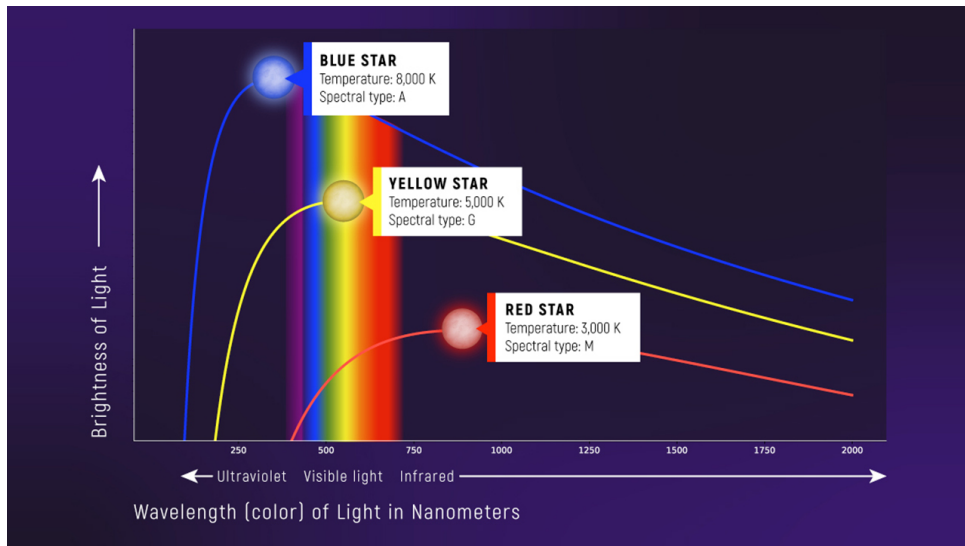


Figure 5: Example of a 3 stellar blackbody radiation plots for stars of different temperatures showing how their spectra peak at different wavelengths depending on their temperature. Red dwarfs have temperatures around 3000 K, meaning their emission peaks in the red and infrared parts of the spectrum. White dwarfs have a wide range of temperatures, anywhere between 4,000 and 150,000 Kelvin, meaning their peak wavelength is usually found in the ultraviolet [6].

sample seed of 42.

6.3 Quasars and Galaxies

Galaxies are some of the most luminous objects in the sky, being comprised of billions of stars. Much like stars, there are different categories of galaxies. At the highest level, galaxies can be classified as ‘normal’ or ‘active’, referring to the activity found at the center, or nucleus, of the galaxy. Much like the Milky Way, many large galaxies contain a black hole in the center of their nuclei. If matter is *actively* falling into the black hole, a vast amount of energy is output during the violent accretion process, which can be detected and measured by telescopes. These environments are classified as Active Galactic Nuclei (AGN). Some of these AGN’s were found to be extremely luminous, but paradoxically small, creating a new sub-classification of AGN’s called quasars. These objects are theorized to be ‘early’ galaxies whose nucleus is absorbing orders of magnitude more matter than most AGN’s. Their extreme brightness allows modern observatories to investigate their properties, giving astronomers the chance to study the supermassive black holes found at their centers, test theories of general relativity, examine the interstellar medium between the quasar source and the telescope, galactic evolution, and the history of the Universe in general. [10]

Inherent differences exist between normal galaxies and standard AGN’s versus quasars. Quasars are smaller and similar to the appearance of a star in photometric observations (figure 6 left image), leading to their original name: ‘quasi-stars’. The SDSS database provides the measured galactic radius of the observed objects, which we can use as a distinguishable metric to help identify quasars. The amount of activity between non-active (normal) galaxies and quasars differ greatly, and the same is true between quasars and standard AGN’s. Much like with the previous datasets, this can be examined using the color indices provided by the UGRIZ system. As seen in figure 6’s right image, quasar galaxies have a much higher redshift than

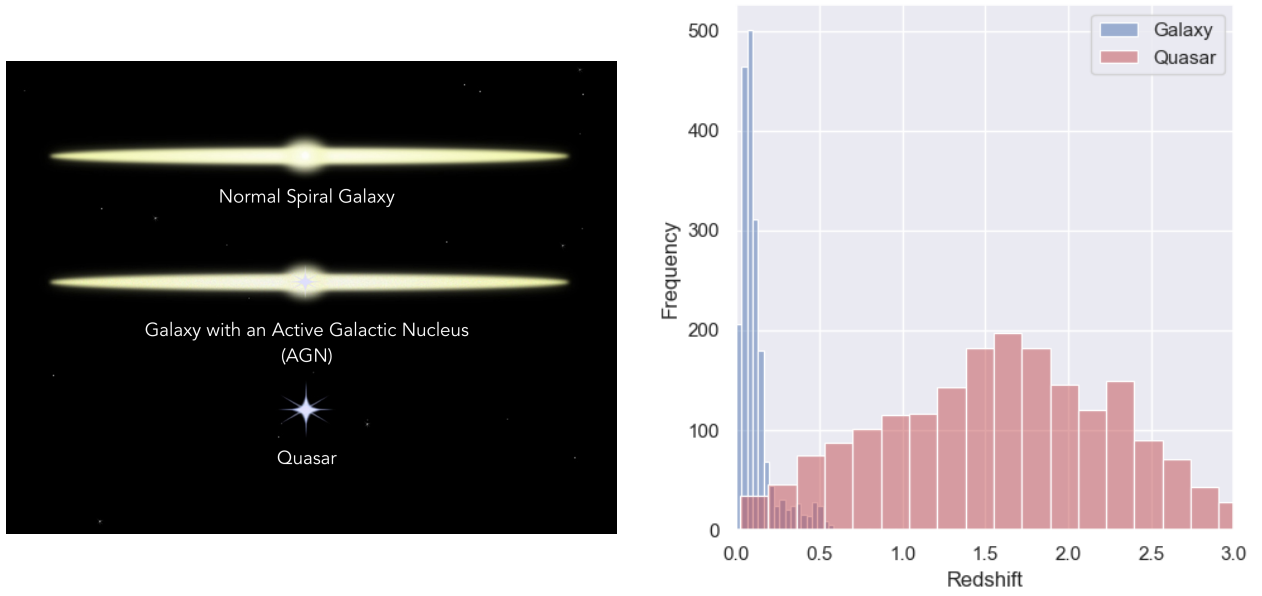


Figure 6: Left Image: Size comparison of quasars to normal galaxies and AGN galaxies. [8] Right Image: A histogram of redshift values between quasar and non-quasar Galaxies

non-quasar galaxies. This is a natural consequence of being able to detect quasar AGN’s from much farther distances, thanks to their immense energy outputs, causing them to experience a greater redshift. The identification of quasars in massive observational surveys, such as the SDSS, is of great use to the field of astronomy.

The training parameters used for our experiments on the galaxy and quasar dataset include the (u-r), (r-z), (u-z) color index projections over their spectra, and the redshift values of the objects. With the introduction of the redshift parameter, this dataset was scaled in order to compare the values of features with different units (color indices vs. redshift). Also, we removed any quasar data points whose redshift value was farther than 0.2 standard deviations away from the mean of the standardized value, in order to more accurately capture the redshift feature space of quasars (reference the histogram in figure 6). The final dataset consists of 2,000 quasar and 2,000 non-quasar galaxy examples directly sampled from the SDSS: Data Release 18 dataset with a random sample seed of 42.

7 Results and Discussion

In the experimental setup, we initially computed the Minimum Enclosing Ball (MEB) on all the entries of the positive class (RR Lyrae/White Dwarfs/Quasars). After that, we evaluated the accuracy of the computed MEB on the complete dataset.

For clarity, we will refer to the three algorithms in the following manner:

- Algorithm 1: $(1 + \epsilon)$ Approximation to the MEB
- Algorithm 2: Away-Steps Frank-Wolfe
- Algorithm 3: Pairwise Frank-Wolfe

7.1 Line search method selection

In order to select a suitable step-size for algorithm 2 and 3, we conducted some tests by running the algorithms using all three line search methods on the RR Lyrae Stars dataset and on its standardized version. During the tests we kept track of the number of iterations and the computational time. We set the maximum number of iterations to 1000 and the stopping condition $\epsilon = 0.001$. The results show that the Exact Line Search method is faster than the other methods for both algorithms and on both versions of the dataset. The experiment has also shown that the Armijo method failed to converge in both algorithms and both versions of the dataset, reaching the maximum number of iterations.

Method	Dataset	Iterations	CPU Time [ms]
Harmonic	Original	97	570
	Standardized	1000	5060
Armijo	Original	1000	52900
	Standardized	1000	35480
Exact	Original	8	69
	Standardized	81	570

Table 2: Line search selection results for the Away-Steps algorithm

Method	Dataset	Iterations	CPU Time [ms]
Harmonic	Original	41	240
	Standardized	110	590
Armijo	Original	1000	53000
	Standardized	1000	38260
Exact	Original	4	40
	Standardized	43	300

Table 3: Line search selection results for the Pairwise algorithm

7.2 Anomaly Detection for RR Lyrae Stars

Algorithm	ϵ	Radius	Core Set Size	CPU Time [ms]	Iterations
1	0.01	0.219	4	132.6	95
	0.001	0.218	4	1325.9	996
	0.0001	0.218	4	>1315.1	>1000
2	0.01	0.212	3	18.02	6
	0.001	0.218	2	39.27	8
	0.0001	0.218	2	37.08	8
3	0.01	0.218	2	48.06	4
	0.001	0.218	2	17.22	4
	0.0001	0.218	2	17.58	4

Table 4: Training Metrics for RR Lyrae Anomaly Detection

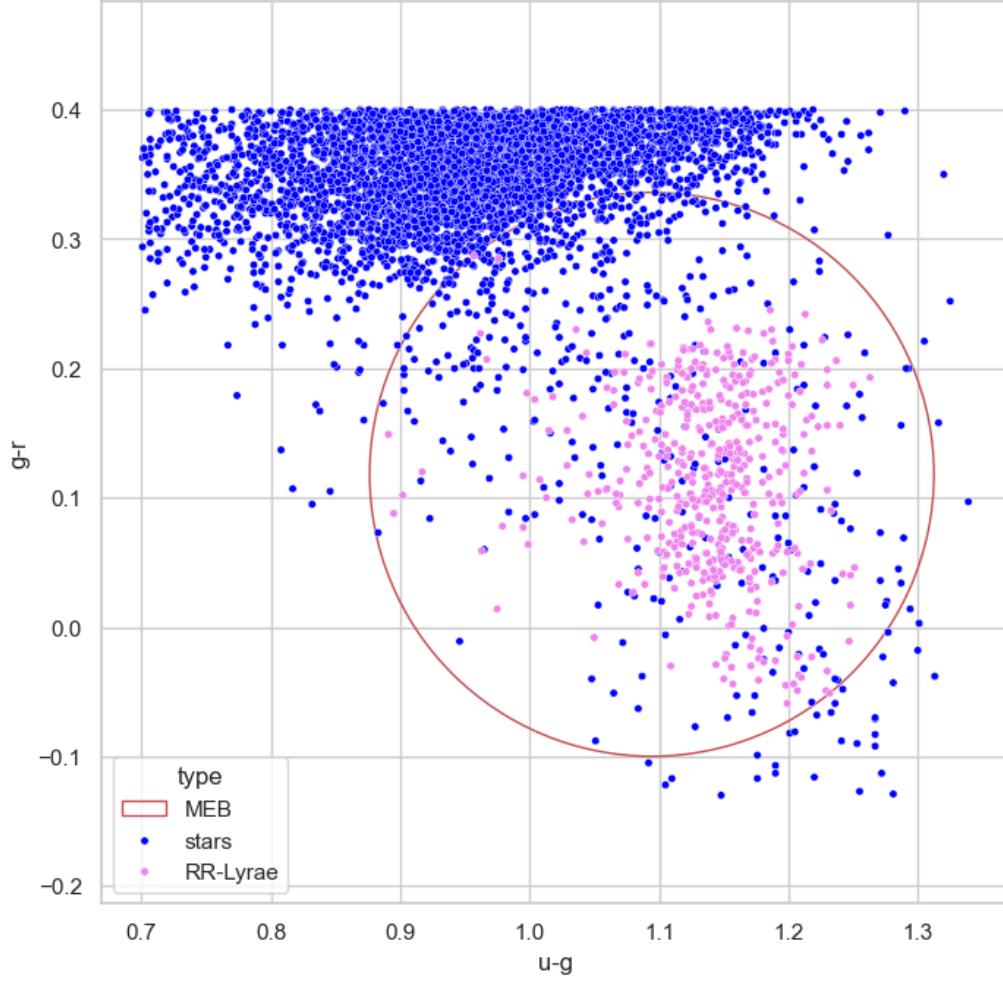


Figure 7: Scatterplot for RR Lyrae Anomaly Detection (Algorithm 1, $\epsilon = 0.0001$)

Metric	Algorithm 1			Algorithm 2			Algorithm 3		
ϵ	0.01	0.001	0.0001	0.01	0.001	0.0001	0.01	0.001	0.0001
TP	483	482	482	478	481	481	481	481	481
FN	0	1	1	5	2	2	2	2	2
TN	4056	4055	4055	4145	4059	4059	4059	4059	4059
FP	461	462	462	372	458	458	458	458	458
Precision	0.512	0.511	0.511	0.562	0.512	0.512	0.512	0.512	0.512
Recall	1.000	0.998	0.998	0.990	0.996	0.996	0.996	0.996	0.996
F1 Score	0.677	0.676	0.676	0.717	0.677	0.677	0.677	0.677	0.677

Table 5: Accuracy Metrics for RR Lyrae Anomaly Detection

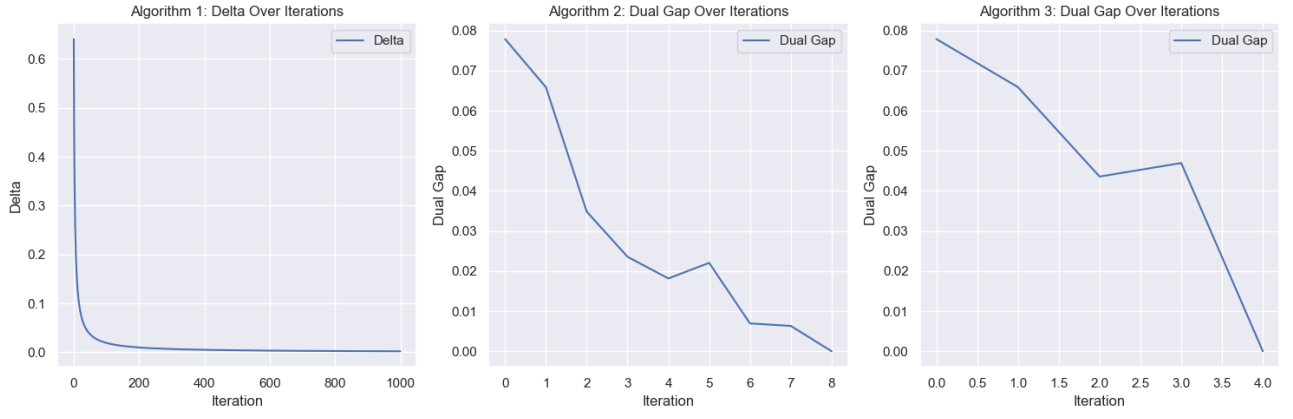


Figure 8: Delta/Dual gap over iteration for the three algorithms on the RR Lyrae dataset

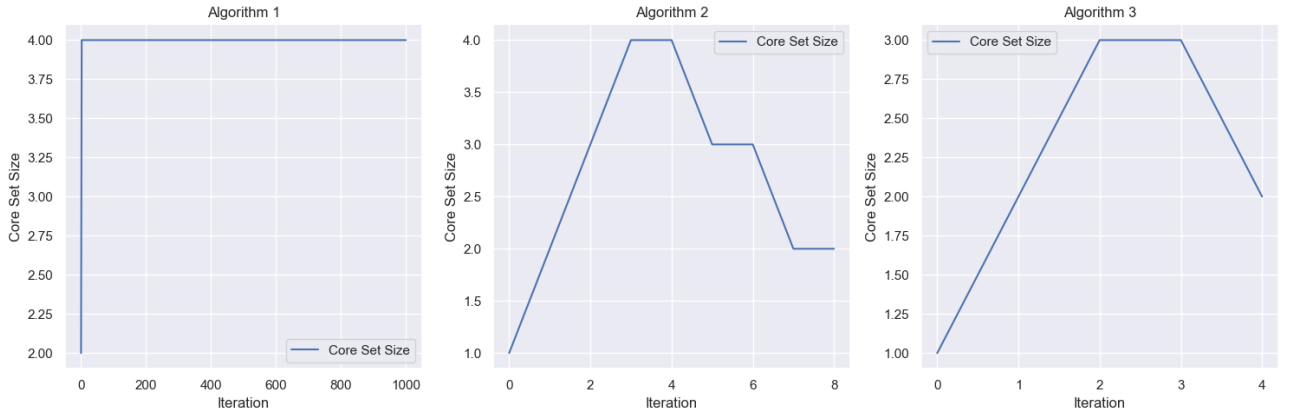


Figure 9: Core Set Size over Iterations for the three algorithms on the RR Lyrae dataset

7.3 Anomaly Detection for White Dwarfs

Algorithm	ϵ	Radius	Core Set Size	CPU Time [ms]	Iterations
1	0.01	325.1	4	36.46	5
	0.001	322.6	4	58.21	11
	0.0001	322.3	4	85.75	17
2	0.01	322.3	3	753.2	25
	0.001	322.3	3	891.1	28
	0.0001	322.3	2	941.6	31
3	0.01	322.3	3	686.2	25
	0.001	322.3	3	743.0	26
	0.0001	322.3	3	858.5	27

Table 6: Training Metrics for White Dwarfs Anomaly Detection

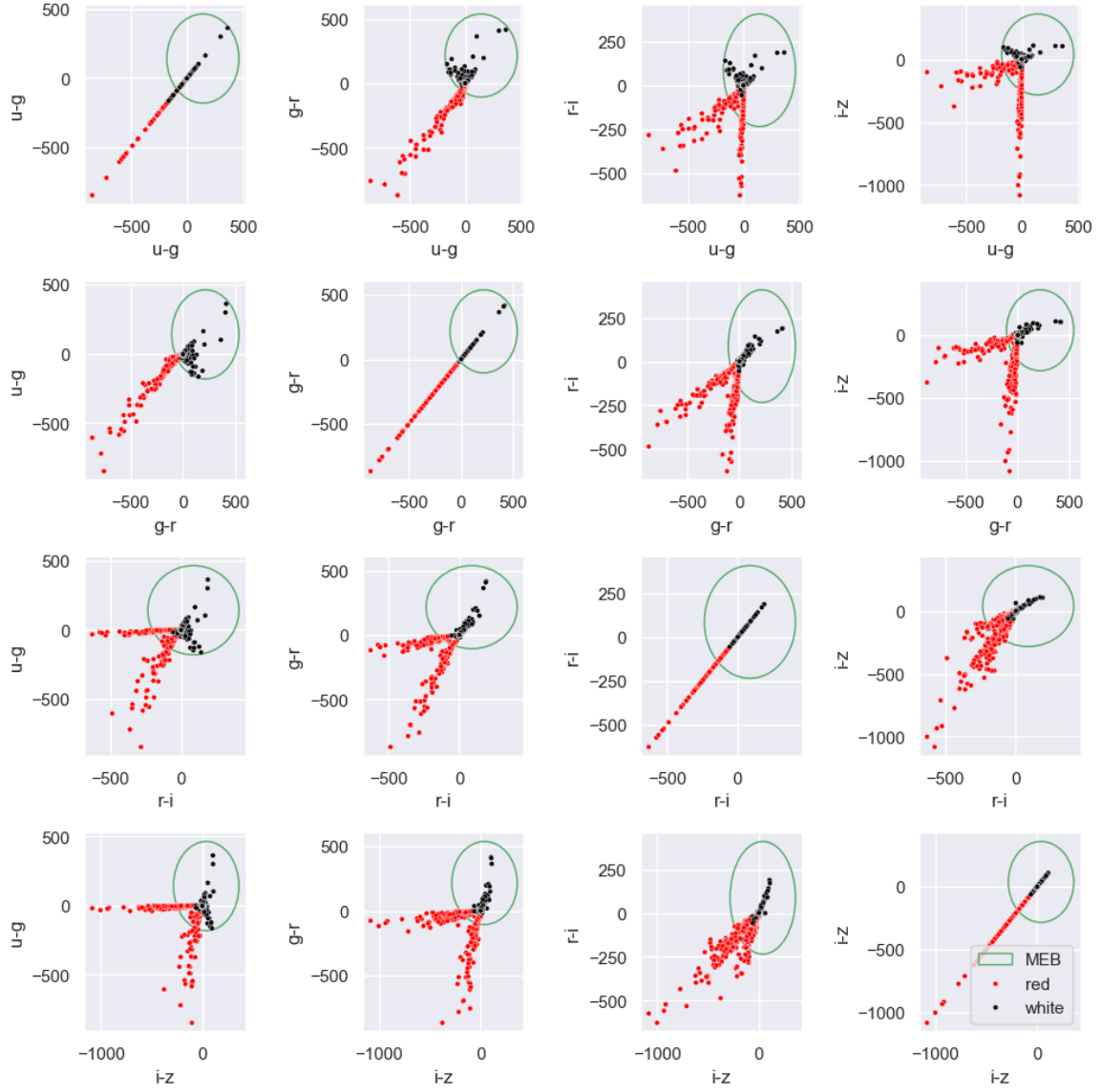


Figure 10: Scatterplot Matrix for White Dwarfs Anomaly Detection (Algorithm 1, $\epsilon = 0.0001$)

Metric	Algorithm 1			Algorithm 2			Algorithm 3		
ϵ	0.01	0.001	0.0001	0.01	0.001	0.0001	0.01	0.001	0.0001
TP	1999	1999	1999	1998	1999	1998	1999	1998	1999
FN	1	1	1	2	1	2	1	2	1
TN	381	423	426	427	427	427	427	427	427
FP	1581	1539	1536	1535	1535	1535	1535	1535	1535
Precision	0.558	0.565	0.565	0.566	0.566	0.566	0.566	0.566	0.566
Recall	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
F1 Score	0.716	0.722	0.722	0.722	0.722	0.722	0.722	0.722	0.722

Table 7: Accuracy Metrics for White Dwarfs Anomaly Detection

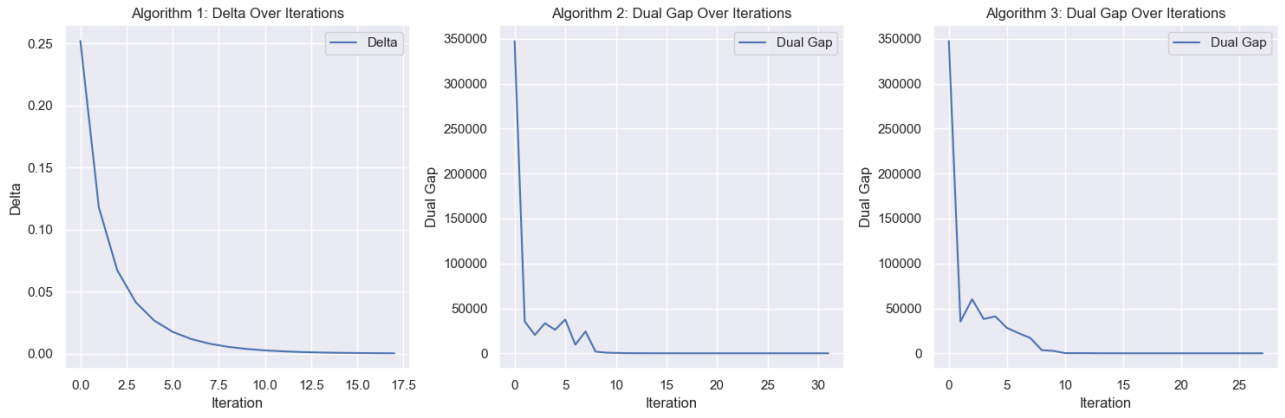


Figure 11: Delta/Dual gap over iteration for the three algorithms on the White Dwarfs dataset

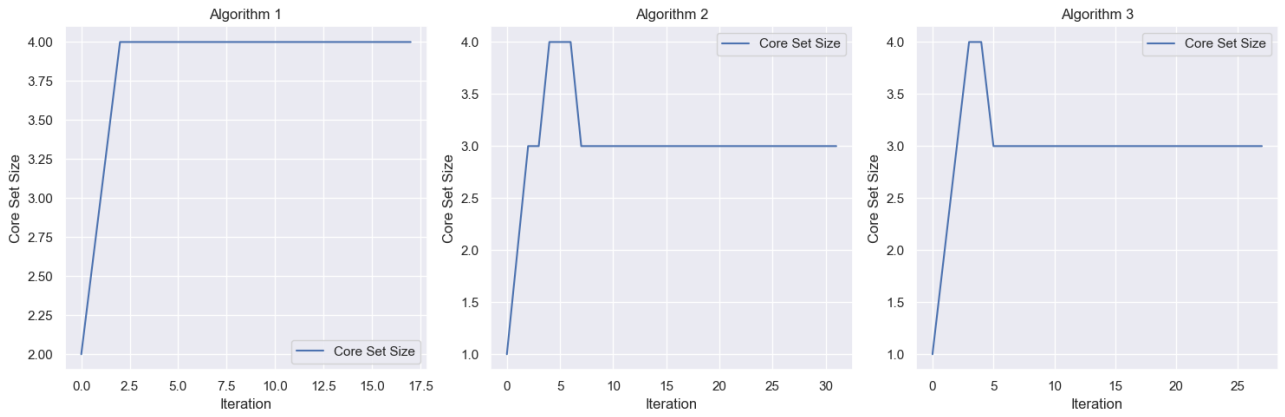


Figure 12: Core Set Size over Iterations for the three algorithms on the White Dwarfs dataset

7.4 Anomaly Detection for Quasars

Algorithm	ϵ	Radius	Core Set Size	CPU Time [ms]	Iterations
1	0.01	2.757	4	62.97	16
	0.001	2.734	4	173.5	48
	0.0001	2.732	4	273.5	84
2	0.01	2.731	3	373.9	23
	0.001	2.732	3	496.2	31
	0.0001	2.732	3	592.7	41
3	0.01	2.732	3	228.6	16
	0.001	2.732	3	200.6	16
	0.0001	2.732	3	275.1	20

Table 8: Training Metrics for Quasars Anomaly Detection

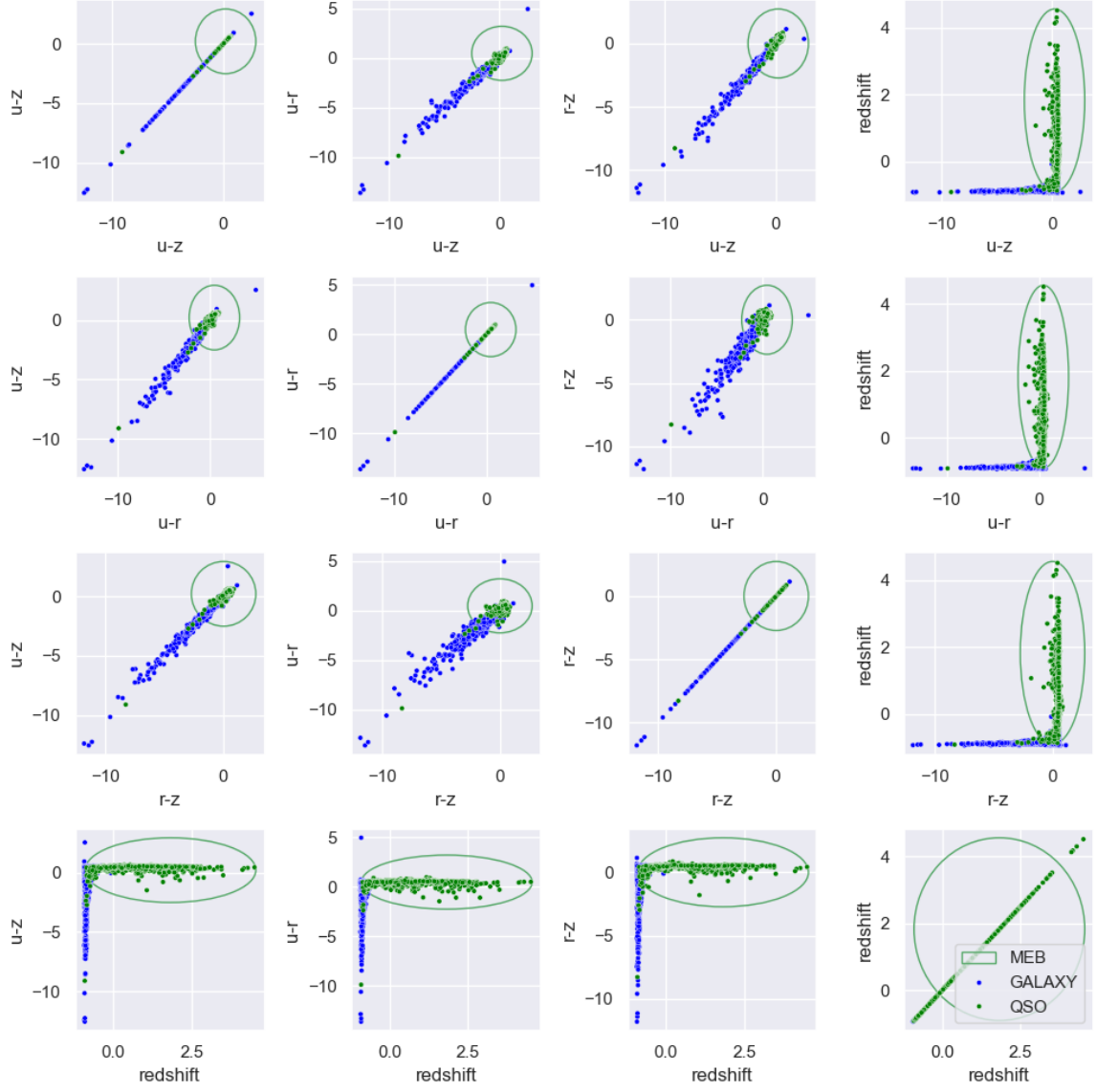


Figure 13: Scatterplot Matrix for Quasars Anomaly Detection (Algorithm 1, $\epsilon = 0.0001$)

Metric	Algorithm 1			Algorithm 2			Algorithm 3		
ϵ	0.01	0.001	0.0001	0.01	0.001	0.0001	0.01	0.001	0.0001
TP	1983	1982	1981	1980	1980	1980	1981	1981	1981
FN	17	18	19	20	20	20	19	19	19
TN	1319	1375	1379	1379	1379	1380	1380	1380	1380
FP	681	625	621	621	621	620	620	620	620
Precision	0.744	0.761	0.761	0.761	0.761	0.762	0.762	0.762	0.762
Recall	0.991	0.991	0.991	0.990	0.990	0.990	0.991	0.991	0.991
F1 Score	0.850	0.860	0.861	0.861	0.861	0.861	0.861	0.861	0.861

Table 9: Accuracy Metrics for the Quasars Anomaly Detection

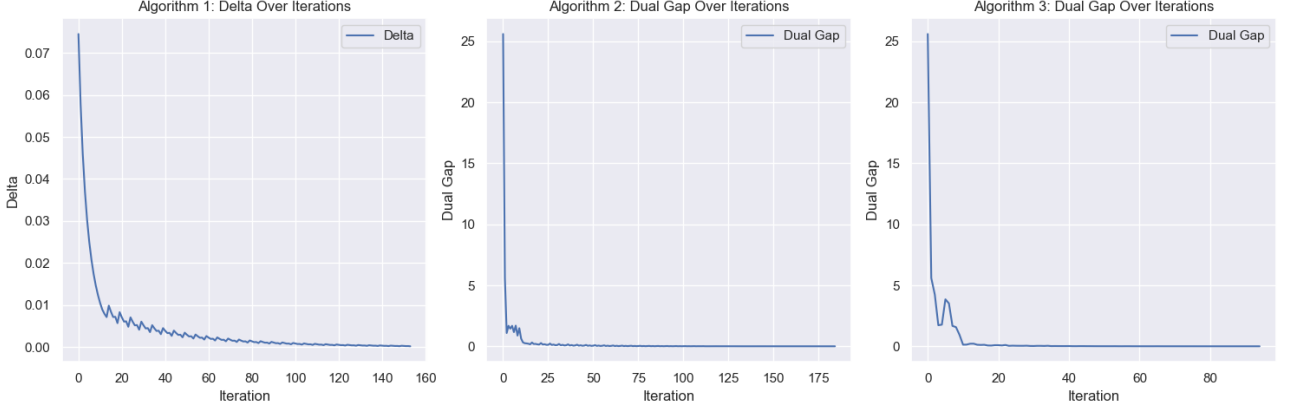


Figure 14: Delta/Dual gap over iteration for the three algorithms on the Quasars dataset

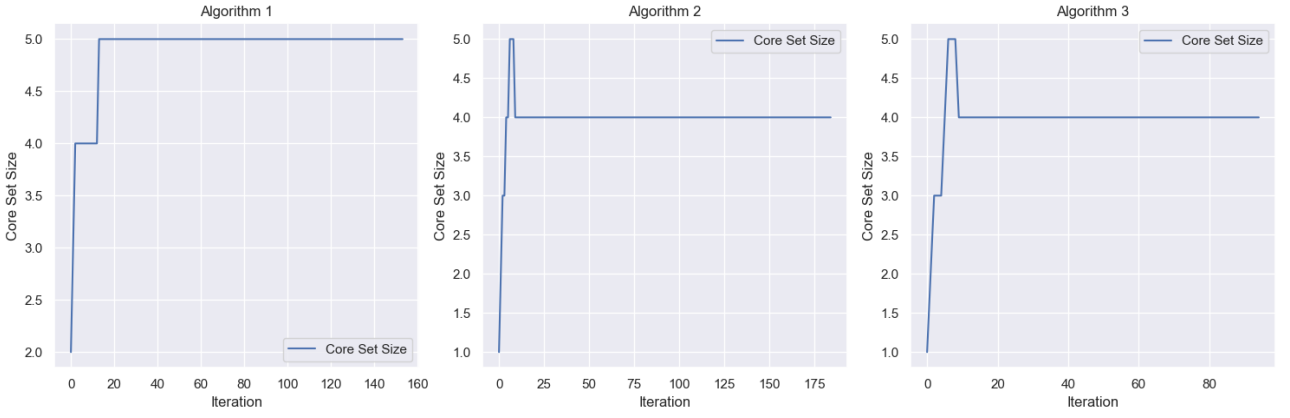


Figure 15: Core Set Size over Iterations for the three algorithms on the Quasars dataset

7.5 Discussion

- The experiments on the three datasets show that, as expected, the three algorithms gave very similar results in terms of precision, recall and F1-score. This is not surprising since all the algorithms aimed to optimize the same function and converged.
- It is interesting to notice that the FW variants outperformed the first algorithm in terms of iterations and computational time on the first dataset, but the opposite is true for the other two datasets. This might be related to the fact that the first dataset only contains two variables, while the other two contain four variables.
- Surprisingly, in the RR Lyrae dataset, for both the first and the second algorithm, the obtained results were better in terms of precision for bigger values of epsilon, which exemplifies the fact that minimizing the objective function of the MEB dual problem does not necessarily lead to better accuracy metrics.
- In some cases, when doing parameter selection for the datasets, increasing the number of parameters for the model to train worsened the performance of the model. One surprising example of this was the addition of radius as a parameter of galaxies and quasars. As shown in Figure 6, quasar galaxies are naturally ‘smaller’ than non-quasar galaxies, but galaxies come in all shapes and sizes. There were enough non-quasar galaxies with similar radius values to the quasar values to make the use of the radius variable useless, and even

harmful. This highlights the importance of choosing quality variables over a high quantity of variables when performing data analysis tasks.

- In Figure 14, the first algorithm demonstrates the distinctive zig-zagging pattern, which is a characteristic behavior of the simple Frank-Wolfe algorithm. Moreover, we observe that both the second and third algorithms mitigate this zig-zagging behavior. This improved convergence is important to mention as it contributes to the algorithms' efficiency in optimizing the objective function over iterations.
- For the second and third datasets, we can see how both the second and third algorithm behave erratically on the first few iterations before reaching stability around the 15th iteration. This might indicate that the Frank-Wolfe variants might be more sensitive to initial conditions. When comparing this to the core set size evolution over iterations, we can see how the size rapidly increases at first and then stabilizes on what is the final optimal value.
- By comparing the coreset size plots, we can notice that the first algorithm only adds vertices to the coreset, until it finds the correct number of vertices, while the other two algorithms add vertices in the beginning, but then remove some of them. The removal of the vertices in excess seems to be related to the noisy behaviour of the FW variants in the first few iterations.

8 Conclusion

In this project, we explored how optimization algorithms can be used in astronomy and astrophysics. Our main focus was on solving the Minimum Enclosing Ball (MEB) problem, which is important for anomaly detection and understanding celestial object distribution in large datasets. We compared three variants of the Frank-Wolfe algorithm (Simple Frank-Wolfe, Away Step Frank-Wolfe, and Pairwise Frank-Wolfe algorithm) and analyzed their performance on different astronomical datasets.

Our analysis revealed interesting behaviors of the optimization algorithms on different datasets. We observed zig-zagging behavior in Algorithm 1, indicating the need for improvements to enhance its convergence. Algorithms 2 and 3 showed erratic behavior initially but stabilized after the 15th iteration. This raises questions about the adaptability and stability of these algorithms in astronomy.

Our findings may have significant implications for astronomy and astrophysics. The optimization algorithms showed good performance in addressing anomaly detection challenges, particularly in computing the Minimum Enclosing Ball. The observed stability and convergence patterns provide valuable insights for implementing these algorithms in analyzing massive astronomical datasets. With the continuous generation of massive data from telescopes and observatories, optimization will continue to play a crucial role in processing and extracting valuable insights in the fields of Astronomy and Astrophysics.

References

- [1] Michael S. Bessell. Standard photometric systems. *Annu. Rev. Astron. Astrophys.*, 43:293–336, 2005.
- [2] Stéphane Canu, Rémi Flamary, and David Mary. Introduction to optimization with applications in astronomy and astrophysics. <https://hal.archives-ouvertes.fr/hal-01346134>, 2016.
- [3] M. S. Catalan, M. J. Sarna, C. M. Jomaron, and Robert Connon Smith. Spectroscopy of the white dwarf-red dwarf binary re j1629+780. *Monthly Notices of the Royal Astronomical Society*, 275(1):153–161, July 1995.
- [4] Almeida A. et al. SDSS data release 18, 2023.
- [5] Jesse L. Greenstein and Mildred S. Matthews. Studies of the white dwarfs. i. broad features in white dwarf spectra. *Mount Wilson and Palomar Observatories*, February 1957. Received on February 14, 1957.
- [6] Space Telescope Science Institute. Stellar black bodies, 2021.
- [7] Simon Lacoste-Julien and Martin Jaggi. On the global linear convergence of frank-wolfe optimization variants, 2015.
- [8] Chandra Xray Observatory. Quasar and galaxy size comparison, 2012.
- [9] Donald E. Osterbrock. The internal structure of red dwarf stars. *Princeton University Observatory*, May 1953. Received on May 10, 1953.
- [10] Michael A. Seeds and Dana E. Backman. *Foundations of Astronomy*. Cengage, 14 edition, 2019.
- [11] J.T. Vanderplas, A.J. Connolly, Ž. Ivezić, and A. Gray. Introduction to astroml: Machine learning for astrophysics. In *Conference on Intelligent Data Understanding (CIDU)*, pages 47 –54, oct. 2012.
- [12] Wikipedia. Photometric system. https://en.wikipedia.org/wiki/Photometric_system, 2023. Accessed: January 30, 2024.
- [13] Wikipedia. Rr lyrae period-luminosity relationship. https://en.wikipedia.org/wiki/RR_Lyrae_variable, 2023. Accessed: January 31, 2024.
- [14] Edward L. Wright. Doppler shift, 2002.
- [15] E. Alper Yildirim. Two algorithms for the minimum enclosing ball problem. *SIAM Journal on Optimization*, 19(3):1368–1391, 2008.