Safe Screening for ℓ_2-norm Penalized Unbalanced Optimal Transport

Anonymous Author
Anonymous Institution

Abstract

Unbalanced optimal transport (UOT) has attracted a surge of research interest in the optimal transport (OT) theory. Recently, a mutual connection between the UOT problem and the Lasso problem has been revealed. Safe screening saves computational time by freezing the zero elements in the sparse solution of the Lasso problem. To this end, we present a proposal of the first safe screening technique to the ℓ_2-norm penalized UOT problem. Specifically, addressing a specific structure of the constraint matrix in the UOT problem, a new two-hyperplane safe region construction method and a shifting projection method are proposed. Numerical evaluations on the Gaussian distributions and the MNIST dataset demonstrate the extraordinary effectiveness of our proposed screening method compared with the state-of-the-art methods in the Lasso problem without significantly increasing the computational burden.

1 INTRODUCTION

Optimal transport (OT) has a long history in mathematics and has recently become popular in machine learning and statistical learning due to its excellence in measuring the distance between two probability measures. It has outperformed traditional methods in many different areas such as domain adaptation (Courty, 2017), generative models (Arjovsky et al., 2017), graph machine learning (Maretic et al., 2019) and natural language processing (Chen et al., 2019). The popularity of OT is attributed to the introduction of the Sinkhorn's algorithm (Sinkhorn, 1974) for the entropy-regularized Kantorovich formulation problem (Cuturi, 2013), which alleviates the computational burden in large-scale problems.

Addressing one limitation that the standard OT problem handles only *balanced* samples, the unbalanced optimal transport (UOT) has been proposed envisioning a wider range of applications with *unbalanced* samples (Caffarelli and McCann, 2010; Chizat et al., 2017). This application fields include, for example, computational biology (Schiebinger et al., 2019), machine learning (Janati et al., 2019) and deep learning (Yang and Uhler, 2019). Mathematically, UOT replaces the equality constraints with penalty functions on the marginal distributions with a divergence including KL divergence (Liero et al., 2018), and ℓ_1-norm (Caffarelli and McCann, 2010) and ℓ_2-norm distances (Benamou, Jean-David, 2003). The KL-penalized UOT with an entropy regularizer can be solved by the Sinkhorn algorithm (Pham et al., 2020). It is fast, scalable, and differentiable, but suffers from instability (Schmitzer, 2016), larger errors, and a lack of sparsity in solution compared with other regularizers (Blondel et al., 2018). In contrast, ℓ_2-norm regularized UOT not only has lower errors but also brings a sparse solution. This attracted the attention of researchers to develop faster optimization algorithms (Blondel et al., 2018; Nguyen et al., 2022).

Despite the success in the development of fast and efficient optimization algorithms for UOT, the computational burden remains a bottleneck for large-scale applications. This paper tackles this issue from a different direction independent of optimization algorithms. Recently, Chapel et al. (2021) have suggested a mutual connection between the UOT problem and the Lasso-like problem (Tibshirani, 1996; Efron et al., 2004) and the non-negative matrix factorization problem (Lee and Seung, 2000). This motivates us to adapt *safe screening* (Ghaoui et al., 2010) in these fields to accelerate the computation of the UOT problem. In fact, the OT and UOT problems expect the solution to be sparse due to the effectiveness of their optimal transport cost and thus share a similar idea with the Lasso problem. Safe screening in the Lasso problem is a promising technique to speed up optimization com-

Preliminary work. Under review by AISTATS 2022. Do not distribute.

putations by exploiting the sparsity of the solution. It eliminates elements in the solution that are guaranteed to be zero without solving the optimization problem. This does not affect the final solution. However, this straightforward application is not trivial. In fact, because the cost matrix of the UOT problem has a wide range of values, the projection step in the conventional Lasso-type screening method deteriorates the efficiency when applied to the UOT problem, and fail as long as zero cost exists. Furthermore, while the Lasso-like problem has a dense *constraint matrix*, that of the UOT problem is extremely sparse and has a unique transport matrix structure. This would benefit the design of a specialized safe screening region.

This paper presents a proposal of a new safe screening method designated to the UOT problem. To this end, we first derive a dual formulation of the vectorized Lasso-like UOT problem. Then, particularly addressing the structure of the constraint matrix, a safe screening method is derived, which includes a new projection method and a safe screening region construction method. These are suitable for the UOT problem and could largely improve the effects.

Contributions.:

- To the best of our knowledge, this work is the first dynamic screening method designated to the UOT problem, which is based on the Lasso-like formulation of the UOT problem and its dual formulation. Our proposed method is independent of optimization algorithms which could be combined will any solvers.
- We propose a projection method by leveraging the specific structure of the UOT problem, and it significantly reduces the errors in the projection process. We also propose a cruciform two hyperplane method that can adjust the safe screening region for every primal element to achieve a better screening effect without adding computational burden.
- Numerical evaluation reveals their effectiveness in terms of projection distances and screening ratios compared with several methods including state-of-the-art methods.

The paper is organized as follows. **Section 2** presents preliminary descriptions of optimal transport and unbalanced optimal transport, as well as safe screening methods in the Lasso-like problem. In **Section 3**, our proposed safe screening method for the UOT problem is detailed, where the region construction method for the safe screening and the projection method are proposed. **Section 4** shows numerical experiments.

2 PRELIMINARIES

$\mathbb{R}^n$ denotes n-dimensional Euclidean space, and $\mathbb{R}^n_+$ denotes the set of vectors in which all elements are non-negative. $\mathbb{R}^{m\times n}$ represents the set of $m\times n$ matrices. Also, $\mathbb{R}^{m\times n}_+$ stands for the set of $m\times n$ matrices in which all elements are non-negative. We present vectors as bold lower-case letters $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \dots$ and matrices as bold-face upper-case letters $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$. The i-th element of $\boldsymbol{a}$ and the element at the (i,j) position of $\mathbf{A}$ are represented respectively as a_i and $A_{i,j}$. The i-th column of $\mathbf{A}$ is represented as $\boldsymbol{a}_i$. In addition, $\mathbb{1}_n \in \mathbb{R}^n$ is the n-dimensional vector in which all the elements are one. For $\boldsymbol{x}$ and $\boldsymbol{y}$ of the same size, $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}^T\boldsymbol{y}$ is the Euclidean dot-product between vectors. For two matrices of the same size $\mathbf{A}$ and $\mathbf{B}$, $\langle \mathbf{A}, \mathbf{B}\rangle = \mathrm{tr}(\mathbf{A}^T\mathbf{B})$ is the Frobenius dot-product. We use $\|\boldsymbol{a}\|_2$ and $\|\boldsymbol{a}\|_1$ to represent the ℓ_2-norm and ℓ_1-norm of $\boldsymbol{a}$, respectively. D_ϕ is the Bregman divergence with the strictly convex and differentiable function ϕ, i.e., $D_\phi(\boldsymbol{a}, \boldsymbol{b}) = \sum_i d_\phi(a_i, b_i) = \sum_i [\phi(a_i) - \phi(b_i) - \phi'(a_i)(a_i - b_i)]$. In addition, we suggests a vectorization for $\mathbf{A} \in \mathbb{R}^{m\times n}$ as a lowercase letters $\boldsymbol{a} \in \mathbb{R}^{mn}$ and $\boldsymbol{a} = \mathrm{vec}(\mathbf{A}) = [\mathbf{A}_{1,1}, \mathbf{A}_{1,2}, \cdots, \mathbf{A}_{m,n-1}, \mathbf{A}_{m,n}]$.

2.1 Optimal Transport and Unbalanced Optimal Transport

Optimal Transport (OT): Given two discrete probability measures $\boldsymbol{a} \in \mathbb{R}^m$ and $\boldsymbol{b} \in \mathbb{R}^n$, the standard OT problem seeks a corresponding *transport matrix* $\mathbf{T} \in \mathbb{R}^{m\times n}_+$ that minimizes the total transport cost (Kantorovich, 1942). This is a linear programming (LP) problem formulated as

$$\begin{aligned} \mathrm{OT}(\boldsymbol{a}, \boldsymbol{b}) \quad &:= \quad \min_{\mathbf{T}\in\mathbb{R}^{m\times n}_+} \langle \mathbf{C}, \mathbf{T}\rangle \\ \text{subject to} \quad & \quad \mathbf{T}\mathbb{1}_n = \boldsymbol{a}, \mathbf{T}^T\mathbb{1}_m = \boldsymbol{b}, \end{aligned} \tag{1}$$

where $\mathbf{C} \in {\mathbb{R}_+}^{m\times n}$ is the *cost matrix*. The constraints in (1) are so-called *mass-conservation constraints* or *marginal constraints*, and assume $\|\boldsymbol{a}\|_1 = \|\boldsymbol{b}\|_1$. Thus, the solution $\hat{\boldsymbol{t}}$ does not exist when $\|\boldsymbol{a}\|_1 \neq \|\boldsymbol{b}\|_1$. The obtained OT matrix $\mathbf{T}^*$ brings powerful distances as $\mathcal{W}_p = \langle \mathbf{T}^*, \mathbf{C}\rangle^{\frac{1}{p}}$, which is known as the p-th order *Wasserstein distance* (Villani, 2008).

As $\boldsymbol{t} = \mathrm{vec}(\mathbf{T}) \in \mathbb{R}^{mn}$ and $\boldsymbol{c} = \mathrm{vec}(\mathbf{C}) \in \mathbb{R}^{mn}$, we reformulate Eq.(1) in a vector format as (Chapel et al., 2021)

$$\begin{aligned} \mathrm{OT}(\boldsymbol{a}, \boldsymbol{b}) \quad &:= \quad \min_{\boldsymbol{t}\in\mathbb{R}^{mn}_+} \boldsymbol{c}^T\boldsymbol{t} \\ \text{subject to} \quad & \quad \mathbf{N}\boldsymbol{t} = \boldsymbol{a}, \mathbf{M}\boldsymbol{t} = \boldsymbol{b}, \end{aligned} \tag{2}$$

where $\mathbf{N} \in \mathbb{R}^{m\times mn}$ and $\mathbf{M} \in \mathbb{R}^{n\times mn}$ are two matrices composed of "0" and "1". $\mathbf{N}$ and $\mathbf{M}$ in case of $m =$

$n = 3$ are given, respectively, as

$$\mathbf{N} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix},$$
$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Note that $\mathbf{N}$ and $\mathbf{M}$ both have a specific structure, where each column has only one single non-zero element equal to 1.

Unbalanced Optimal Transport (UOT): The marginal constraints in (1) may exacerbate degradation of the performance of some applications where weights need not be strictly preserved. In contrast, the UOT problem *relaxes* them by replacing the equality constraints with penalty functions on the marginal distributions with a divergence (Caffarelli and McCann, 2010; Chizat et al., 2017). Formally, defining $\boldsymbol{y} = [\boldsymbol{a}, \boldsymbol{b}]^T \in \mathbb{R}^{mn}$ and the concatenation *constraint matrix* $\mathbf{X} = [\mathbf{M}^T, \mathbf{N}^T]^T \in \mathbb{R}^{(m+n)\times mn}$, the UOT problem can be formulated introducing a penalty function for the probability measures as (Chapel et al., 2021)

$$\mathrm{UOT}(\boldsymbol{a}, \boldsymbol{b}) := \min_{\boldsymbol{t} \in \mathbb{R}_+^{mn}} \boldsymbol{c}^T \boldsymbol{t} + D_\phi(\mathbf{X}\boldsymbol{t}, \boldsymbol{y}). \tag{3}$$

It is worth mentioning that this function is convex due to the convexity of the Bregman divergence.

Relationship with Lasso-like problem: Addressing that $\boldsymbol{c}$ and $\boldsymbol{t}$ in (3) are nonnegative, the term $\boldsymbol{c}^T\boldsymbol{t}$ is represented as $\boldsymbol{c}^T\boldsymbol{t} = \sum_i c_i t_i = \sum_i c_i |t_i|$. Addition, setting $D_\phi(\mathbf{X}\boldsymbol{t}, \boldsymbol{y}) = \frac{1}{2}\|\mathbf{X}\boldsymbol{t} - \boldsymbol{y}\|_2^2$, we find that the UOT problem in (3) is equivalent to a weighted ℓ_1-norm regularized Lasso-like problem. It is, however, that $\mathbf{X} = [\mathbf{M}^T, \mathbf{N}^T]^T$ in (3) substantially differs from that of the Lasso problem. More concretely, the former $\mathbf{X}$ has a specific structure and has only two non-zero elements equal to 1 in each row whereas the latter $\mathbf{X}$ in Lasso problem is non-structured and dense (Chapel et al., 2021).

2.2 Safe Screening in Lasso-like problem

Solutions to many large-scale optimization problems tend to be sparse, and a large amount of computation is wasted on updating zero elements during the optimization process. *Safe screening* is a well-known effective technique in the Lasso problem and the SVM problem (Ogawa et al., 2013), where the ℓ_1-norm regularizer leads to a sparse solution (Ghaoui et al., 2010). It can pre-select solutions that must be zero theoretically and freeze them before optimization calculations, thus saving optimization time. Many safe screening methods have been proposed in this past decade (Liu et al., 2014; Wang et al., 2015). Recent dynamic screening methods efficiently drop variable elements, which include Dynamic Screening (Bonnefoy et al., 2015), Gap Safe screening (Ndiaye et al., 2017) and Dynamic Sasvi (Yamada and Yamada, 2021). Hereinafter, we briefly elaborate on the framework proposed in (Yamada and Yamada, 2021) to introduce the dynamic screening technique for an optimization problem:

$$\min_{\boldsymbol{t}} \{f(\boldsymbol{t}) := g(\boldsymbol{t}) + h(\mathbf{X}\boldsymbol{t})\}, \tag{4}$$

where g and h are proper convex functions, and where $\boldsymbol{t}$ is the primal optimization variable, and $\mathbf{X}$ is a constant matrix. For this problem, the Fenchel-Rockafellar Duality yields the dual problem as presented below:

Theorem 1 (Fenchel-Rockafellar Duality (Rockafellar and Wets, 1998))**.** *Consider the problem (4). We assume that all the assumptions are satisfied. Then we have the following:*

$$\min_{\boldsymbol{t}} g(\boldsymbol{t}) + h(\mathbf{X}\boldsymbol{t}) = \max_{\boldsymbol{\theta}} -h^*(-\boldsymbol{\theta}) - g^*(\mathbf{X}^T\boldsymbol{\theta}), \tag{5}$$

where g^* and h^* are the convex conjugate of g and h, respectively.

Because the primal function h is always convex, the dual function h^* is concave. Assuming h^* is an L-strongly concave problem, one can design an area for any feasible $\tilde{\boldsymbol{\theta}}$ by the strongly concave property:

Theorem 2 (L-strongly concave (Yamada and Yamada, 2021, Theorem 5))**.** *Considering problem in Eq.(4), if function h and g are both convex, for $\forall\ \tilde{\boldsymbol{\theta}}$ and satisfied the constraints on the dual problem, we have the following area constructed by its L-strongly concave property:*

$$\mathcal{R} := \boldsymbol{\theta} \in \left\{ \frac{L}{2}\|\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}\|_2^2 + h^*(-\tilde{\boldsymbol{\theta}}) \leq h^*(-\boldsymbol{\theta}) \right\}.$$

Theorem 3 (Circle constraint for Lasso-like problem (Yamada and Yamada, 2021, Theorem 8))**.** *Consider the Lasso-like problem given as $\min g(\boldsymbol{t}) + \frac{1}{2}\|\mathbf{X}\boldsymbol{t} - \boldsymbol{y}\|_2^2$, where g is a norm function. We also consider Fenchel-Rockafellar Duality by Eq.(5). We assume that $\tilde{\boldsymbol{\theta}}$ is inside the domain of the dual problem. Then, $\mathcal{R}$ in* **Theorem 2** *is equivalent to the following circle constraint $\mathcal{R}^C$:*

$$\mathcal{R}^C := \{\boldsymbol{\theta} \mid (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^T(\boldsymbol{\theta} - \boldsymbol{y}) \leq 0\}. \tag{6}$$

$\boldsymbol{\theta}$ and $\boldsymbol{y}$ are the endpoints of the diameter of this circle, and $\hat{\boldsymbol{\theta}} \in \mathcal{R}^C$.

As the dual optimal solution $\hat{\boldsymbol{\theta}}$ satisfies the **Theorem 2** as a corollary of L-strongly concave property. The area $\mathcal{R}^C$ is not empty.

3 UOT SAFE SCREENING

This section presents a proposal of a new safe screening method designated to the UOT problem. For this purpose, after deriving the dual formulation of the Lasso-like formulated UOT problem in Eq.(3), the dual feasible area, denoted as $\mathcal{R}^D$, is derived. Then, a new construction method for safe screening region, denoted as $\mathcal{R}^S$, is derived, which is composed of the circle constraint $\mathcal{R}^C$ in (6) and a relaxed region of $\mathcal{R}^D$ with two hyperplances. Finally, we derive a new shifting projection method to perform our safe screening based on $\mathcal{R}^S$. Concrete proofs of lemma and theorems are provided in the supplementary file.

3.1 Dual formulation and Feasible Region

Dual formulation of UOT: We consider the case where D_ϕ is a quadratic divergence in (3). Then, the UOT problem has the form of $h(\mathbf{X}\boldsymbol{t}) = \frac{1}{2}\|\mathbf{X}\boldsymbol{t} - \boldsymbol{y}\|_2^2$ and $g(\boldsymbol{\theta}) = \lambda \boldsymbol{c}^T\boldsymbol{t}$, where $t_i \geq 0$ for $i \in [mn]$ in Eq.(4). Therefore, we obtain the dual form by **Theorem 1**:

Lemma 4 (Dual form of the UOT problem)**.** *For the UOT problem in Eq.(3) with a quadratic divergence D_ϕ, we obtain the following dual problem:*

$$\begin{aligned} -h^*(-\boldsymbol{\theta}) - g^*(\mathbf{X}^T\boldsymbol{\theta}) &= -\frac{1}{2}\|\boldsymbol{\theta}\|_2^2 + \boldsymbol{y}^T\boldsymbol{\theta} \\ \textit{subject to} \qquad \boldsymbol{x}_p^T\boldsymbol{\theta} - \lambda c_p &\leq 0, \quad \forall p \in [mn], \end{aligned} \tag{7}$$

where $\boldsymbol{x}_p$ corresponds to the p-th column of $\mathbf{X}$.

Note that the strongly concave coefficient L for the dual function f is 1.

Dual feasible region $\mathcal{R}^D$: The constraint in Eq.(7) specifies a dual variable feasible area, denoted as $\mathcal{R}^D$, and the optimal dual solution $\hat{\boldsymbol{\theta}} \in \mathcal{R}^D$ because of **Theorem 2**. According to the KKT condition of the Fenchel-Rockafellar Duality in **Theorem 1**, there is a connection between the optimal primal solution $\hat{\boldsymbol{t}}$ and dual solution $\hat{\boldsymbol{\theta}}$ as presented below:

Theorem 5. *(From the KKT condition) For the primal optimal solution $\hat{\boldsymbol{t}}$ and the dual optimal solution $\hat{\boldsymbol{\theta}}$, we have the following relationship:*

$$\boldsymbol{x}_p^T\hat{\boldsymbol{\theta}} - \lambda c_p \begin{cases} < 0 & \implies \quad \hat{t}_p = 0 \\ = 0 & \implies \quad \hat{t}_p \geq 0. \end{cases} \tag{8}$$

The proof of **Theorem 5** is given in the supplementary material. It should be noted that the left-hand side term of Eq.(8), $\boldsymbol{x}_p^T\hat{\boldsymbol{\theta}}$, is different from that in the standard Lasso problem, $\|\boldsymbol{x}_p^T\hat{\boldsymbol{\theta}}\|_1$, due to $t_p > 0$ in the UOT problem. In addition, c_p appears in the left-hand side, which has a big impact on the projection method in the UOT problem as will be described later.

Eq.(8) indicates a potential method to screen the primal variable $\boldsymbol{t}$. However, since we do not know the optimal dual solution $\hat{\boldsymbol{\theta}}$ directly, this paper seeks a safe screening region for screening, denoted as $\mathcal{R}^S$, as small as possible that contains the $\hat{\boldsymbol{t}}$. In other words, if

$$\max_{\boldsymbol{\theta} \in \mathcal{R}^S} \quad \boldsymbol{x}_p^T\boldsymbol{\theta} - \lambda c_p < 0, \tag{9}$$

we assert $\boldsymbol{x}_p^T\hat{\boldsymbol{\theta}} - \lambda c_p < 0$, which means the corresponding $\hat{t}_p = 0$, and can be screened out.

For this purpose, we focus on the special structure of $\mathbf{X}$ in the UOT problem, i.e., the p-th column of $\mathbf{X}$ has only two nonzero elements equal to 1. Defining $\boldsymbol{\theta} = [\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T]^T$, where $\boldsymbol{\alpha} \in \mathbb{R}^m$ and $\boldsymbol{\beta} \in \mathbb{R}^n$, and $(u, v) = (p \mid n, v = p \mod n)$, we rewrite $\boldsymbol{x}_p^T\hat{\boldsymbol{\theta}} - \lambda c_p < 0$ as

$$\alpha_u + \beta_v - \lambda c_p < 0. \tag{10}$$

Because nm inequalities can be reorganized as a $m \times n$ matrix like Fig. 2, every element of the matrix corresponds to each inequality. If we obtain $\tilde{\boldsymbol{\theta}}$ in $\mathcal{R}^D$, combining $\mathcal{R}^D$ with $\mathcal{R}^C$ in **Theorem 3**, we could use $\mathcal{R}^C \cap \mathcal{R}^D$ as the area $\mathcal{R}^S$ in Eq.(9). It is, however, that two difficulties arise: (i) maximizing the left side of (10) on the $\mathcal{R}^D$ is time consuming, and (ii) we need a $\tilde{\boldsymbol{\theta}} \in \mathcal{R}^D$ first to construct the $\mathcal{R}^C$. We address two issues below:

3.2 two-hyperplane Safe Screening Region

Because $\mathcal{R}^D$ forms a $(n+m)$-polytope with nm facets, it is time-consuming to maximize on the intersection of a hyper-ball $\mathcal{R}^C$ and a polytope with the Simplex method or the Lagrangian method for every primal element t_p. One simple remedy for dealing with the problem is to *relax* the polytope constraint into a set of hyper-plane constraints. Thus, considering the specific structure of $\mathbf{X}$ in the UOT problem, our proposed method relaxes the polytope into two hyperplanes by combining the nm dual constraints with a positive weight $t_p \geq 0$. It should be noted that, not like (Xiang et al., 2017), we design two specific hyperplanes for every single element in the UOT problem, and maximize (10) on them without an extra computational burden. If the constraint is relevant to the dual elements α_u and β_v, which would largely influence the screening of t_p, we add it to one group, and the rest of the constraints are added to another.

More concretely, we divide the polytope constraints into two groups represented by two index sets of I_p and I_p^C for every p, and then two-hyperplanes regions given by these two groups construct $\mathcal{R}_p^S$. This new $\mathcal{R}_p^S$ is smaller than the relaxation in (Yamada and Yamada, 2021). Moreover, benefiting from the structure

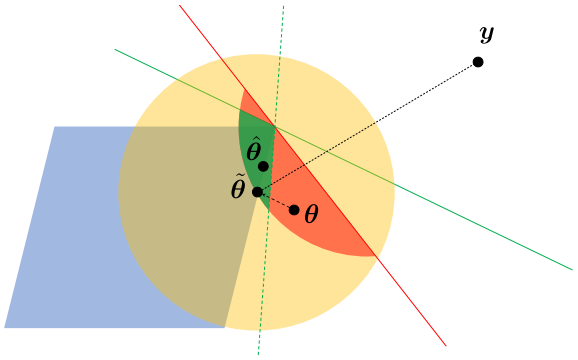

Figure 1: Comparision of Gap region (yellow), Dynamic Sasvi's region (red) and CTP method's region (green), the blue part is the $\mathcal{R}^D$, the blue region is the dual feasible area $\mathcal{R}^D$, the green line and the dash green line represent two planes constructed by I_p and I_p^C.

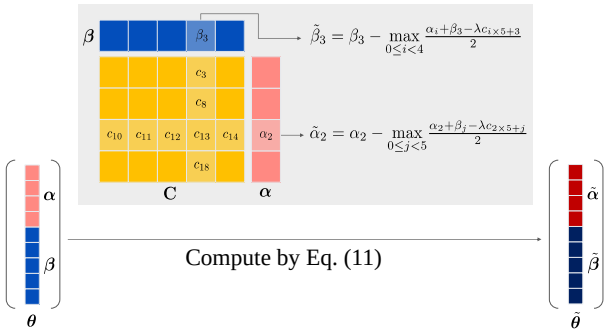

Figure 2: Example of Shifting Projection on a 4×5 matrix. $\boldsymbol{\theta}$ has been projected to the $\mathcal{R}^D$ as $\hat{\boldsymbol{\theta}}$, mn constraints can be rearranged as a yellow matrix, and the light yellow blocks indicate the constraints that directly affects the corresponding elements.

of $\boldsymbol{x}_p^T\boldsymbol{\theta} = \alpha_u + \beta_v$, we can put $m+n$ constraints that are directly connected with the optimization direction α_u and β_v together, to alleviate the relaxation influence caused by the other secondary variables.

Theorem 6 (Cruciform two-hyperplane safe screening region for UOT)**.** *For every primal variable $t_l \ge 0$, $l \in [mn]$, let $I_p = \{i \mid 0 \le i < nm, u = i \mid n \vee v = i \bmod n\}$, and $I_p^C = \{i \mid 0 \le i < nm, i \notin I_p\}$. Then, we construct a specific area $\mathcal{R}_p^S$ as presented below:*

$$\mathcal{R}_p^S(\boldsymbol{\theta}, \boldsymbol{t}) := \left\{ \boldsymbol{\theta} \;\middle|\; \begin{array}{l} \sum_{l\in I_p} (\boldsymbol{x}_l^T\boldsymbol{\theta} - \lambda c_l) t_l \le 0, \\ \sum_{l\in I_p^C} (\boldsymbol{x}_l^T\boldsymbol{\theta} - \lambda c_l) t_l \le 0, \\ (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^T(\boldsymbol{\theta} - \boldsymbol{y}) \le 0 \end{array} \right\}. \quad (11)$$

Consequently, as illustrated in Figure 1, the green region is $\mathcal{R}_p^S$, constructed by the circle and two hyperplanes. We have $\hat{\boldsymbol{\theta}} \in \mathcal{R}^C \cap \mathcal{R}^D \subset \mathcal{R}_p^S$ (The proof is given in the Appendix). We designate this as *cruciform two-hyperplane screening (CTP)* in this paper. As $\mathcal{R}^C$ is a circle, we assume that its center is θ_o. Then, rewriting $\boldsymbol{\theta} = \boldsymbol{\theta}_o + \boldsymbol{q}$ and $r = \frac{\|\boldsymbol{y}-\boldsymbol{\theta}\|_2}{2}$, the Lagrangian function of the optimization problem of (9) is given as

$$\begin{aligned} \min_{\boldsymbol{q}} \max_{\eta,\mu,\nu\ge 0} L(\boldsymbol{q},\eta,\mu,\nu) = \min_{\boldsymbol{q}} \max_{\eta,\mu,\nu\ge 0} & -\boldsymbol{x}_p^T\boldsymbol{\theta}_o \\ & - \boldsymbol{q}_u - \boldsymbol{q}_{u+v+1} + \eta(\boldsymbol{q}^T\boldsymbol{q} - r^2) \\ & + \mu \left((\sum_{l\in I_p} \boldsymbol{x}_l)^T(\boldsymbol{q} + \boldsymbol{\theta}_o) - \lambda \sum_{l\in I_p} c_l t_l) \right) \\ & + \nu \left((\sum_{l\in I_p^C} \boldsymbol{x}_l)^T(\boldsymbol{q} + \boldsymbol{\theta}_o) - \lambda \sum_{l\in I_p^C} c_l t_l) \right), \end{aligned} \quad (12)$$

where η, μ, ν (≥ 0) are Lagrangian multipliers. This problem can be solved efficiently by the Lagrangian method in constant time. Its computational process is detailed in the supplementary material.

3.3 Shifting Projection

Before we start to screening with the safe screening region $\mathcal{R}^S$ that is constructed by **Theorem 6**, we first have to find a dual variable $\tilde{\boldsymbol{\theta}} \in \mathcal{R}^D$. Although there exists a relationship between the primal variable and dual variable that $\boldsymbol{\theta} = \boldsymbol{y} - \mathbf{X}\boldsymbol{t}$, inevitably, we often get $\boldsymbol{\theta} \notin \mathcal{R}^D$. This requires us to project $\boldsymbol{\theta}$ onto $\mathcal{R}^D$. Dynamic screening gets a more accurate screening outcome because the optimization algorithm gradually provides $\boldsymbol{t}^k \to \hat{\boldsymbol{t}}$, and also the dual variable $\boldsymbol{\theta}^k \to \hat{\boldsymbol{\theta}}$. If $\tilde{\boldsymbol{\theta}} \to \hat{\boldsymbol{\theta}}$, the safe screening region $\mathcal{R}^C \cap \mathcal{R}^D$ would get smaller and has a better screening outcome. As $\mathcal{R}^D$ is a polytope combined with nm hyperplanes, the cost of a real projection is expensive even if one uses a fast Linear programming (LP) algorithm or the Dykstra's projection algorithm. To this issue, in the Standard Lasso problem, its dual constraints corresponding to Eq.(7) are $\|\boldsymbol{x}_p\hat{\boldsymbol{\theta}}\|_1 \le \lambda$, where $\boldsymbol{x}_p$ is almost dense and non-structured. Thus, one always uses a simple shrinking method to obtain a $\tilde{\boldsymbol{\theta}} \in \mathcal{R}^D$ as $\tilde{\boldsymbol{\theta}} = \frac{\boldsymbol{\theta}}{\max(1, \|\mathbf{X}^T\boldsymbol{\theta}/\lambda\|_\infty)}$ (Ndiaye et al., 2017, Proposition 11) (Yamada and Yamada, 2021, Theorem 11).

One simple extension of this into the UOT problem could be

$$\tilde{\boldsymbol{\theta}} = \frac{\boldsymbol{\theta}}{\max(1, \|\frac{\mathbf{X}^T\boldsymbol{\theta}}{\lambda \boldsymbol{c}^T}\|_\infty)},$$

where $(\frac{\cdot}{\cdot})$ is the element-wise division. However, this projection pushes $\boldsymbol{\theta}$ far away from the optimal solution $\hat{\boldsymbol{\theta}}$ because it suffers from the influence of a small c_p. Also, it will degenerate when one of the costs is $c_p = 0$

and disables the screening process. Therefore, noting that the UOT problem only allows $t_p \geq 0$ and the $\boldsymbol{x}_p$ only consists of two non-zero elements, we can adapt a more efficient projection method. Specifically, as Figure 2 illustrates, we shift every single α_u, β_v to $\tilde{\alpha}_u$, $\tilde{\beta}_v$ with a half of the maximum difference in all of its constraints. We guarantee that our shifting projection ensures that the projected values are in $\mathcal{R}^D$ as stated below:

Theorem 7 (Shifting projection for UOT)**.** *For any $\boldsymbol{\theta} = [\boldsymbol{\alpha}^T, \boldsymbol{\beta}^T]^T$, we define the projection operator $\tilde{\boldsymbol{\theta}} =$ Projection$(\boldsymbol{\theta})$, where*

$$\begin{cases} \tilde{\alpha}_u & = \alpha_u - \max\limits_{0 \leq j < n} \dfrac{\alpha_u + \beta_j - \lambda c_{un+j}}{2} \\ \tilde{\beta}_v & = \beta_v - \max\limits_{0 \leq i < m} \dfrac{\alpha_i + \beta_v - \lambda c_{in+v}}{2}. \end{cases} \tag{13}$$

Then, the calculated $\tilde{\boldsymbol{\theta}} = [\tilde{\boldsymbol{\alpha}}^T, \tilde{\boldsymbol{\beta}}^T]^T \in \mathcal{R}^D$.

The proof is in the supplementary material. Note that this cheap projection method can get good accuracy with $O(knm)$ computational time.

3.4 Optimization with Safe Screening

The purpose of screening methods is to remove variable elements that are ignored in the optimization process in the adopted optimization algorithm before actual optimization process. For this, we introduce *screening mask vector*, denoted as $\boldsymbol{s} \in \mathbb{R}^{mn}$, for the transport vector $\boldsymbol{t}$. $s_p = 0$ tells the optimizer that its corresponding element t_p should be screened out. We summarize a specific algorithm for ℓ_2-norm penalized UOT problem to show the entire optimization process as in **Algorithm 1**. The update$(\boldsymbol{t}, \boldsymbol{s})$ operator in the algorithm indicates that the optimizer updates $\boldsymbol{t}^k$ into $\boldsymbol{t}^{k+1}$ except the variables t_p with $s_p = 0$. $w \in \mathbb{N}$ specifies the period of the screening to update the mask $\boldsymbol{s}$ because the change in the screening ratio is actually small if we update $\boldsymbol{s}$ every time. The period w is decided by a user, and the faster converging the optimizer is, the smaller period r is. It should be emphasized that the proposed screening method is *independent* of the adopted optimization algorithm.

3.5 Computational Cost Analysis

Our proposed method needs to construct different plans for every single primal element t_p. However, thanks to the special structure of the $\mathbf{X}$, for every single t_p, the data required for maximizing (by the Lagrangian method in supplementary material) can be summarized as some specific sum, which can be computed together and reused for other elements that have the same $t \mod m$ or $t \mid m$. This helps us to preserve the whole optimization complexity to $O(kmn)$, where k is a constant.

Algorithm 1 UOT Optimization with Safe Screening

Input: $\boldsymbol{t}^0, \boldsymbol{s} \in \mathbb{R}^{nm}, s_p = 1 \ \forall p \in [mn], w$
Output: $\boldsymbol{t}^K$

```
for k = 0 to K − 1 do
  if (k mod w) = 0 then
    θ̃^k = Projection(θ^k)                                   Eq.(13)
    for p = 0 to mn − 1 do
      R^S ← R_p^S(θ̃^k, t^k)                                 Eq.(12)
      s ← s_p = 0 if max_{θ∈R^S} x_p^T θ^k < λ c_p           Eq.(9)
    end for
  end if
  t^{k+1} = update(t^k, s)        (Optimization process)
end for
return t^K
```

4 EXPERIMENTS

In this section, we conduct three experiments on randomly generated gaussian distributions and the MNIST dataset. We first validate the effectiveness of our proposed projection method by measuring the distance between the dual variable and the projected point. Next, the proposed two-plane screening (CPT) is compared with some state-of-the-art methods. We finally apply our safe screening method on popular l_2-norm penalized UOT problem solvers including FISTA, BFGS, Lasso (celer), Multiplicative update, and Regularization path to show its speed-up ratios.

4.1 Projection Method Comparison

To show the effectiveness of our projection method compared with conventional ones in the Lasso problem, we evaluate the projection distance and screening ratio with randomly generated Gaussian measures by two projection methods. Because the standard Lasso projection method would degenerate when $c_p = 0$, we add small constants $\epsilon = 0.01, 0.001$ on the cost matrix for both methods. Shifting projection method moves the dual less than the Lasso method and ensure it is inside the $\mathcal{R}^D$, and the order of magnitude is ignorable. It ensures the projection step would not harm the approximation process for the computational dual solution $\boldsymbol{\theta}^k$ to the optimum $\hat{\boldsymbol{\theta}}$. In the experiment, we set the $\lambda = \frac{\|\mathbf{X}^T y\|}{100}$ and test for 10 different pairs. We use the FISTA solver for solving the L_2-norm penalized UOT problems. The results of the projection distance and the screening ratio are shown in Figures 3(a) and 3(b), respectively. From Figures 3(a), we find .....

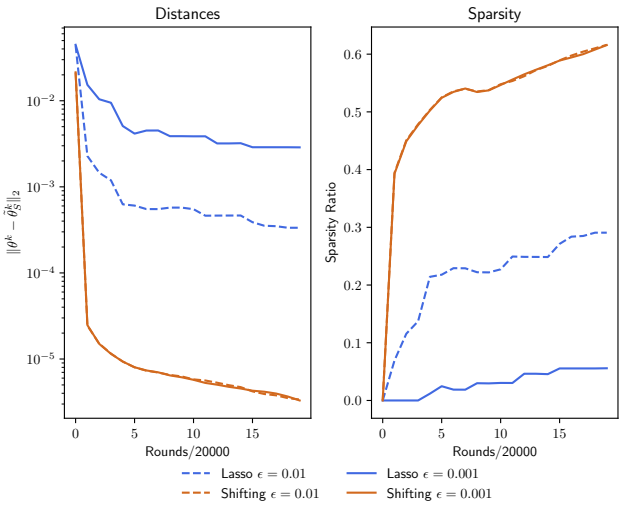

Figure 3: Projection Method Comparison

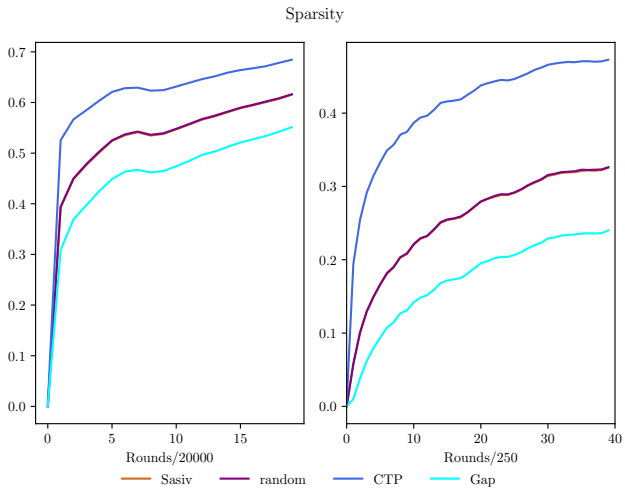

Figure 4: Screening ratio of dividing method

4.2 Comparing with Other Screening Methods

We compared the proposed screening ratio with three different methods, including our Cross Two-Plane method (CTP), a Random Two-Plane method (RTP), Dynamic Sasvi method, and Gap method. All the methods would use our projection method as it has been proven better than the standard Lasso projection method. It is clear that the CTP still outperform all the other method, CTP can start screening at a very early stage. The RTP just randomly divides the constraints and combines them into two planes, RTP's advantage is ignorable than the Sasvi algorithm, which indicates that our achievements can not be simply attributed to the choice of more planes, The effectiveness is highly connected with the specific structure of the UOT problem, which decrease the influence of other variables and make us able to divide the constraints according to the variables.

4.3 Speed up Ratio

We applied the CTP in different UOT optimizers like FISTA, BFGS, Lasso(celer), Multiplicative update, and Regularization path method, and use them to solve the UOT problem between random number figure pairs in the MNIST dataset, we compared the consuming time for different accuracy and showed in the table, Our Screening method can get a good speed up ratio for almost every optimizer, and performs especially good on Multiplicative update algorithms

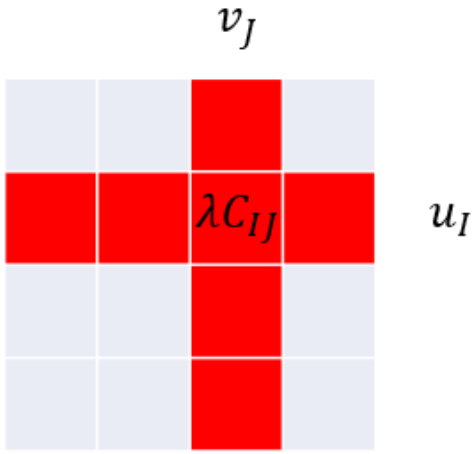

Figure 5: speed up ratio for different solver

5 CONCLUSION

This paper introduced the Screening method in the Lasso community to the UOT problem and demonstrates that the screening method has extra potential to make a difference in the UOT community due to its specific structure. We provide a better projection method for the UOT screening process and design a Cross Two-Plane safe Screening region for it, which accomplishes a good performance and shows the potential to exploit the structure of the UOT problem. We believe these approaches can be extended to other regularized and penalized UOT problems like KL penalized UOT (Dantas et al., 2021) problem and even entropic UOT problem. we are going to combine our method with the Sinkhorn algorithm to alleviate its

drawbacks on computational time and sparsity.

References

Arjovsky, M., Chintala, S., and Bottou, L. (2017). Wasserstein generative adversarial networks. In *ICML*.

Benamou, Jean-David (2003). Numerical resolution of an "unbalanced" mass transport problem. *ESAIM: M2AN*, 37(5):851–868.

Blondel, M., Seguy, V., and Rolet, A. (2018). Smooth and sparse optimal transport. In *AISTATS*.

Bonnefoy, A., Emiya, V., Ralaivola, L., and Gribonval, R. (2015). Dynamic screening: Accelerating first-order algorithms for the lasso and group-lasso. *IEEE Transactions on Signal Processing*, 63(19):5121–5132.

Caffarelli, L. A. and McCann, R. J. (2010). Free boundaries in optimal transport and Monge-Ampère obstacle problems. *Annals of Mathematics*, 171(2):673–730.

Chapel, L., Flamary, R., Wu, H., Févotte, C., and Gasso, G. (2021). Unbalanced optimal transport through non-negative penalized linear regression. In *NeurIPS*.

Chen, L., Zhang, Y., Zhang, R., Tao, C., Gan, Z., Zhang, H., Li, B., Shen, D., Chen, C., and Carin, L. (2019). Improving sequence-to-sequence learning via optimal transport. In *ICLR*.

Chizat, L., Peyré, G., Schmitzer, B., and Vialard, F.-X. (2017). Scaling algorithms for unbalanced transport problems. *arXiv preprint: arXiv:1607.05816*.

Courty, N. (2017). Optimal transport for domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(9):1853–1865.

Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *NeurIPS*.

Dantas, C. F., Soubies, E., and Févotte, C. (2021). Safe screening for sparse regression with the Kullback-Leibler divergence. In *ICASSP*.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. *Annals of Statistics*, 32:407–489.

Ghaoui, L. E., Viallon, V., and Rabbani, T. (2010). Safe feature elimination for the lasso and sparse supervised learning problems. *arXiv preprint arXiv:1009.4219*.

Janati, H., Cuturi, M., and Gramfort, A. (2019). Wasserstein regularization for sparse multi-task regression. In *AISTATS*.

Kantorovich, L. (1942). On the transfer of masses. *Dokl. Akad. Nauk*, 37(2):227–229.

Lee, D. D. and Seung, H. S. (2000). Algorithms for non-negative matrix factorization. In *NIPS*.

Liero, M., Mielke, A., and Savaré, G. (2018). Optimal entropy-transport problems and a new hellinger–kantorovich distance between positive measures. *Inventiones mathematicae*, 211(3):969–1117.

Liu, J., Zhao, Z., Wang, J., and Ye, J. (2014). Safe screening with variational inequalities and its application to lasso. In *ICML*.

Maretic, H. P., Gheche, M. E., Chierchia, G., and Frossard, P. (2019). GOT: An optimal transport framework for graph comparison. In *NeurIPS*.

Ndiaye, E., Fercoq, O., Alex, re Gramfort, and Salmon, J. (2017). Gap safe screening rules for sparsity enforcing penalties. *Journal of Machine Learning Research*, 18(128):1–33.

Nguyen, Q. M., Nguyen, H. H., Zhou, Y., and Nguyen, L. M. (2022). On unbalanced optimal transport: Gradient methods, sparsity and approximation error. *arXiv preprint arXiv:2202.03618*.

Ogawa, K., Suzuki, Y., and Takeuchi, I. (2013). Safe screening of non-support vectors in pathwise SVM computation. In *ICML*.

Pham, K., Le, K., Ho, N., Pham, T., and Bui, H. (2020). On unbalanced optimal transport: An analysis of Sinkhorn algorithm. In *ICML*.

Rockafellar, R. T. and Wets, R. J. B. (1998). *Variational Analysis*. Springer.

Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A., Gould, J., Liu, S., Lin, S., Berube, P., Lee, L., Chen, J., Brumbaugh, J., Rigollet, P., Hochedlinger, K., Jaenisch, R., Regev, A., and Lander, E. S. (2019). Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*, 176(4):928–943.e22.

Schmitzer, B. (2016). Stabilized sparse scaling algorithms for entropy regularized transport problems. *CoRR*, abs/1610.06519.

Sinkhorn, R. (1974). Diagonal equivalence to matrices with prescribed row and column sums. ii. *Proceedings of the American Mathematical Society*, 45(2):195–198.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Villani, C. (2008). *Optimal Transport: Old And New*. Springer.

Wang, J., Wonka, P., and Ye, J. (2015). Lasso screening rules via dual polytope projection. *Journal of Machine Learning Research*, 16:1063–1101.

Xiang, Z. J., Wang, Y., and Ramadge, P. J. (2017). Screening tests for lasso problems. *IEEE Transac-*

tions on Pattern Analysis and Machine Intelligence, 39(5):1008–1027.

Yamada, H. and Yamada, M. (2021). Dynamic sasvi: Strong safe screening for norm-regularized least squares. In *NeurIPS*.

Yang, K. D. and Uhler, C. (2019). Scalable unbalanced optimal transport using generative adversarial networks. In *ICLR*.

Supplementary Material: Safe Screening for ℓ_2-norm Penalized Unbalanced Optimal Transport

A PROOFS

A.1 Proof of Theorem 5

From the (Yamada and Yamada, 2021, Proposition 3) we know that:

$$\mathbf{X}\hat{\boldsymbol{t}} \in \partial h^*(-\hat{\boldsymbol{\theta}}) \wedge \hat{\boldsymbol{t}} \in \partial g^*(\mathbf{X}^T\hat{\boldsymbol{\theta}}) \tag{14}$$

where ∂ is the subgradient, then we know that $\mathbf{X}^T\hat{\boldsymbol{\theta}} \in \partial g(\hat{\boldsymbol{t}})$, compute and you can get the **Theorem 5**.

A.2 Proof of Theorem 7

For any $p \in 0, 1, ..., nm-1$, then we can compute that:

$$\begin{aligned}
\boldsymbol{x}_p^T\tilde{\boldsymbol{\theta}} &= \tilde{\boldsymbol{\alpha}}_u + \tilde{\boldsymbol{\beta}}_v \\
&= \boldsymbol{\alpha}_u + \boldsymbol{\beta}_v - \max_{0\le j<n} \frac{\boldsymbol{\alpha}_u + \boldsymbol{\beta}_j - \lambda \boldsymbol{c}_{un+j}}{2} - \max_{0\le i<m} \frac{\boldsymbol{\alpha}_i + \boldsymbol{\beta}_v - \lambda \boldsymbol{c}_{in+v}}{2} \\
&\le \boldsymbol{\alpha}_u + \boldsymbol{\beta}_v - \frac{\boldsymbol{\alpha}_u + \boldsymbol{\beta}_v - \lambda \boldsymbol{c}_p}{2} - \frac{\boldsymbol{\alpha}_u + \boldsymbol{\beta}_v - \lambda \boldsymbol{c}_p}{2} \\
&\le \lambda \boldsymbol{c}_p
\end{aligned} \tag{15}$$

For $\forall p$, we have $\tilde{\boldsymbol{\theta}} \in \mathcal{R}^D$

A.3 Proof of Theorem 6

We Generalize the problem as

$$\max_{\boldsymbol{\theta}\in\mathcal{R}_I^S} \boldsymbol{\theta}_{I_1} + \boldsymbol{\theta}_{I_2} \tag{16}$$

Considering the center of the circle as $\boldsymbol{\theta}^o$, we define $\boldsymbol{\theta} = \boldsymbol{\theta}^o + q$, as $\boldsymbol{\theta}^o_{I_1} + \boldsymbol{\theta}^o_{I_2}$ is a constant, the problem is equal to $\min_{\boldsymbol{\theta}\in\mathcal{R}_I^S} -(\boldsymbol{q}_{I_1} + \boldsymbol{q}_{I_2})$, we compute the Lagrangrian function of later:

$$\min_{\boldsymbol{q}} \max_{\eta,\mu,\nu\ge 0} L(\boldsymbol{q}, \eta, \mu, \nu) = \min_{\boldsymbol{q}} \max_{\eta,\mu,\nu\ge 0} -\boldsymbol{q}_{I_1} - \boldsymbol{q}_{I_2} + \eta(\boldsymbol{q}^T\boldsymbol{q} - r^2) + \mu(a^T\boldsymbol{q} - e_a) + \nu(b^T\boldsymbol{q} - e_b) \tag{17}$$

$$\frac{\partial L}{\partial \boldsymbol{q}_i} = \begin{cases} -1 + 2\eta\boldsymbol{q}_i + \mu a_i + \nu b_i & i = I_1, I_2 \\ 2\eta\boldsymbol{q}_i + \mu a_i + \nu b_i & i \ne I_1, I_2 \end{cases} \tag{18}$$

$$\boldsymbol{q}_i{}^* = \begin{cases} \dfrac{1 - \mu a_i - \nu b_i}{2\eta} & i = I_1, I_2 \\ -\dfrac{\mu a_i + \nu b_i}{2\eta} & i \ne I_1, I_2 \end{cases} \tag{19}$$

We can get the Lagrangian dual problem:

$$\max_{\eta,\mu,\nu\ge 0} L(\eta, \mu, \nu) = \max_{\eta,\mu,\nu\ge 0} \frac{\mu a_{I_1} + \nu b_{I_1} - 1}{2\eta} + \frac{\mu a_{I_2} + \nu b_{I_2} - 1}{2\eta} + \eta(\boldsymbol{q}^{*T}\boldsymbol{q}^* - r^2) + \mu(a^T\boldsymbol{q}^* - e_a) + \nu(b^T\boldsymbol{q}^* - e_b) \tag{20}$$

From the KKT optimum condition, we know that if

$$\begin{aligned}\eta(\boldsymbol{q}^{*T}\boldsymbol{q}^* - r^2) &= 0\\ \mu(a^T\boldsymbol{q}^* - e_a) &= 0\\ \nu(b^T\boldsymbol{q}^* - e_b) &= 0\end{aligned} \tag{21}$$

We set η^*, μ^*, ν^* as the solution of the equations, which is also the solution of the dual problem. Firstly, we assume that $\eta^*, \mu^*, \nu^* \neq 0$, then the solution is equal to compute the following equations:

$$\begin{aligned}&(1-\mu a_{I_1} - \nu b_{I_1})^2 + (1-\mu a_{I_2} - \nu b_{I_2})^2 + \sum_{i\neq I_1, I_2}^{m+n}(a_i\mu + b_i\nu)^2 - 4\eta^2 r^2 = 0\\ &a_{I_1} - \mu a_{I_1}^2 - \nu b_{I_1}a_{I_1} + a_{I_2} - \mu a_{I_2}^2 - \nu b_{I_2}a_{I_2} - \sum_{i\neq I_1, I_2}^{m}(a_i^2\mu + b_i a_i \nu) - 2\eta e_a = 0\\ &b_{I_1} - \nu b_{I_1}^2 - \mu b_{I_1}a_{I_1} + b_{I_2} - \nu b_{I_2}^2 - \mu b_{I_2}a_{I_2} - \sum_{i\neq I_1, I_2}^{m}(b_i^2\nu + b_i a_i \mu) - 2\eta e_b = 0\end{aligned} \tag{22}$$

Rearranged as:

$$\begin{aligned}&2 - 2\mu(a_{I_1} + a_{I_2}) - 2\nu(b_{I_1} + b_{I_2}) + \|a\|^2\mu^2 + \|b\|^2\nu^2 + 2\mu\nu a^T b - 4\eta^2 r^2 = 0\\ &(a_{I_1} + a_{I_1}) - \|a\|^2\mu - a^T b\nu - 2\eta e_a = 0\\ &(b_{I_1} + b_{I_2}) - \|b\|^2\nu - a^T b\mu - 2\eta e_b = 0\end{aligned} \tag{23}$$

we have

$$\begin{aligned}\mu &= \frac{2(e_b a^T b - e_a\|b\|^2)\eta + (a_{I_1} + a_{I_2})\|b\|^2 - (b_{I_1} + b_{I_2})(a^T b)}{\|a\|^2\|b\|^2 - a^T b}\\ \nu &= \frac{2(e_a a^T b - e_b\|a\|^2)\eta + (b_{I_1} + b_{I_2})\|a\|^2 - (a_{I_1} + a_{I_2})(a^T b)}{\|a\|^2\|b\|^2 - a^T b}\end{aligned} \tag{24}$$

set it as:

$$\begin{aligned}\mu &= s_1\eta + s_2\\ \nu &= u_1\eta + u_2\end{aligned} \tag{25}$$

Then we can solve the η as a quadratic equation:

$$\begin{aligned}0 &= a\eta^2 + b\eta + c\\ a &= 4r^2 - s_1^2\|a\|^2 - u_1^2\|b\|^2 - 2s_1u_1a^T b\\ b &= 2(a_{I_1} + a_{I_2})s_1 + 2(b_{I_1} + b_{I_2})u_1 - 2s_1s_2\|a\|^2 - 2u_1u_2\|b\|^2 - 2(s_1u_2 + s_2u_1)a^T b\\ c &= 2(a_{I_1} + a_{I_2})s_2 + 2(b_{I_1} + b_{I_2})u_2 - s_2^2\|a\|_1 - u_2^2\|b\|_1 - 2s_2u_2a^T b - 2\end{aligned} \tag{26}$$

Then we can put it back into 25 and get μ, ν.

If the solution satisfied the constraints $\eta^*, \mu^*, \nu^* > 0$, then it is the solution. However, if one of the dual variables is less than 0, the problem would degenerate into a simpler question.

If only η^* is larger than 0, $\min_{\boldsymbol{\theta}\in\mathcal{R}_I^S} -(\boldsymbol{q}_{I_1} + \boldsymbol{q}_{I_2}) = -\sqrt{2}r$

If only μ^* or ν^* is less than 0, we are optimizing on a sphere cap, the solution can be found in (Yamada and Yamada, 2021, supplementary material B)

if only $\eta^* \leq 0$: As the sphere is inactivated, the problem gets maximum at every point of the intersection of two planes.

$$\min_{\boldsymbol{q}}\max_{\mu,\nu\geq 0} L(\boldsymbol{q}, \mu, \nu) = \min_{\boldsymbol{q}}\max_{\mu,\nu\geq 0} -\boldsymbol{q}_{I_1} - \boldsymbol{q}_{I_2} + \mu(a^T\boldsymbol{q} - e_a) + \nu(b^T\boldsymbol{q} - e_b) \tag{27}$$

To have a solution, the equations satisfied

$$\frac{\partial L}{\partial q} = \begin{cases} -1 + \mu a_i + \nu b_i = 0 & i = I_1, I_2 \\ -\mu a_i - \nu b_i & = 0 i \neq I_1, I_2 \end{cases} \tag{28}$$

As the equation satisfied, we can just set $\boldsymbol{q}_i^* = 0, i \neq I_1, I_2$, then we compute the

$$\min_{\boldsymbol{\theta} \in \mathcal{R}_I^S} -(\boldsymbol{q}_{I_1} + \boldsymbol{q}_{I_2}) = \frac{a_{I_2} e_b - b_{I_2} e_a - a_{I_1} e_b + b_{I_1} e_a}{a_{I_1} b_{I_2} - a_{I_2} b_{I_1}} \tag{29}$$