

1 We deeply thank reviewers for their insightful and constructive comments, which greatly helped us improve the work.  
2 As all 3 reviewers pointed out, **the main novelty and contribution of this work are that we found that an adapted**  
3 **interior point algorithm can be much more efficient than traditionally expected both theoretically and computa-**  
4 **tionally in solving the Wasserstein Barycenter problem. The creativity lies in reducing the complexity of solving**  
5 **the normal equations in each iteration by fully exploiting its matrix's block diagonal structure. This finding is**  
6 **potentially useful in further speeding up many state-of-art interior point algorithms that mainly focus on re-**  
7 **ducing the iteration numbers.** Compared to popular regularization approaches such as Sinkhorn-based methods, this  
8 algorithm may gain a good balance between accuracy and speed.

9 Some common issues from reviewers are addressed here: we conducted more experiments to compare with Solomon et  
10 al.'s convolutional Sinkhorn algorithm(will keep trying more experiments!), and the comparison result is similar to that  
11 of IBP. For reproducibility, codes are on GitLab now. We will post the link in the revision.

12 **Response to comments of Reviewer 1:** Thank you for very insightful inputs!

13 Q1: Our methods for solving normal equations is parallelizable on the  $N$  axis, too. In SLRM/DLRM, the major cost is  
14 computing  $A_{ii}^{-1}$  for each  $i = 1, \dots, N$  independently(in Algorithm 2 line 1), which can be done in parallel for all  $i$ 's.

15 Q2: It's a crucial observation! Thanks for raising it. In Thm. 4.3: the complexity of one iteration in our IPM is linear  
16 w.r.t.  $N$  and  $m$  (or  $m_i$ ). This is because we directly build up  $\bar{A}D\bar{A}^T$  and the complexity of one IPM iteration is the  
17 same as the complexity of SLRM/DLRM. We'll explain it more in the paper. Thus, in figure 4, if we fix 2 parameters

18 among  $N$ ,  $m$ ,  $m_i$  and change the other one, it'll show a linear rate. When varying  $m$ , the pattern is similar to figure 4.

19 Q3: We use 2 columns to demonstrate the linear relation between runtime and  $N(m')$ . We regard Gurobi's solution  
20 accurate and omit it in the middle 2 columns. In the right column, we compare with Gurobi to display MAAIPM's low  
21 memory occupancy advantages over traditional second order methods, though it does take more memory than first order  
22 methods by nature. It is great advice to display all methods. We'll add it in the paper.

23 Q4: Thanks for this important advice! We have tested the method and it indeed improved the results from Sinkhorn's  
24 side! We will add it in the paper.

25 Improvement suggestions: Thanks for detailed instructions! We have followed each point carefully and revised the  
26 paper accordingly and will include them in the revision.

27 **Response to Reviewer 2:** Thanks! Your concerns helped us strengthen some important issues of the paper!

28 Q1: "However, I am concerned...run time for the previous best interior point algorithm... reported to be  $O(N^3m^4)$ ..."

29 ANS: Thanks for raising up this issue! It's the most important theoretical part and we should have explained it better(will  
30 revise!). The paper focuses on reducing the complexity of solving the normal equation in one iteration. For a standard  
31 LP, let  $\bar{m}$ ,  $\bar{n}$  be the number of constraints and variables. Then computing  $ADA^T$  requires  $O(\bar{m}^2\bar{n})$  flops, while solving  
32 a system with the coefficient matrix  $ADA^T$  requires  $O(\bar{m}^3)$  flops. So a single IPM iteration takes  $O(\bar{m}^2\bar{n})$ . If  $m = m_i$ ,  
33 the LP has  $Nm^2 + m$  variables and  $2Nm + 1$  constraints. Thus  $O(N^3m^4)$  directly comes from formulating and  
34 solving the normal equations in one iteration via direct matrix multiplication and Cholesky decomposition. Possibly  
35  $O(N^3m^4)$  can be improved by techniques such as fast matrix multiplication, but not too much. We'll add more related  
36 results(especially overall complexity) and look into the valuable references you provide.

37 Q2: "I am confused why the NeurIPS paper by Cuturi is cited for Iterative Bregman Projection....."

38 ANS: Thank you for pointing this out! It is a mistake and we'll correct it.

39 Q3: "I think the evaluation of SLRM and DLRM must be done against more recent solvers for normal equations....."

40 ANS: We totally agree! Though the computing time of normal equations in a single iteration can't be exacted accurately  
41 from commercial solvers(told by solver developers), we'll try our best to figure out.

42 Q4: "I think it doesn't make much sense to compare an interior point method with first order methods....."

43 ANS: Thanks for very meaningful comments! We did compare our algorithm with best commercial solvers using  
44 interior points methods(Gurobi/MOSEK, and CPLEX later). We also compared with mainstream methods in this area-  
45 even though they are totally different types of methods- to show its competitiveness.

46 The comments on improving the clarity and the fast matrix inversion are greatly appreciated. They can improve the  
47 theoretical complexity of our SLRM/DLRM. The revision will be made accordingly!

48 **Response to Reviewer 3:** Thank you for sharp observation and kind advice!

49 Q1: "The paragraphs from 132-148 ..." ANS: Thanks for pointing out the problem! Our writing here isn't clear.  
50 MAAIPM cannot optimally solve the free support case. We'll rewrite this more carefully.

51 Q2: "In lines 237-239 ..." ANS: Thanks for great suggestion! We have finished the experiment and will add it in  
52 the paper. The result shows that our memory usage is higher than popular first order methods, but it's usually within  
53 constant times of them. Essentially our advantage lies in the balance of accuracy and speed.

54 Q3: "In addition to Gurobi ..." ANS: We thank the referee for raising an excellent point. Mosek was tested and is much  
55 faster than Gurobi, though it is still roughly 3 times slower than MAAIPM and occupies more memory than MAAIPM  
56 for large problems. We will include it in the paper and will test CPLEX later. MAAIPM possesses not only computation  
57 speed advantage, but also computation memory efficiency.

58 Q4: "The authors can also compare ..." ANS: Thanks for helpful suggestions! We've been testing newer algorithms and  
59 will add those in the revision besides Solomon et al.'s convolutional Sinkhorn algorithm.