MATCHING UNDER METRIC PRIVACY

Anonymous authors
Paper under double-blind review

ABSTRACT

We study minimum-total-distance matching between public and private points under local metric privacy. We compare perturbing original points, distance vectors, and a projected representation that preserves distances to the public candidates. We then correct the projected representation using one additional private distance. Its expected additional matching cost is $O(n\sqrt{m\log(2n)}/\varepsilon)$, where $m \leq \min\{n, d\}$, without distributional or boundedness assumptions on the private points. The correction controls errors in relative costs rather than individual locations. Two constructions reverse the ranking of projected and distance-vector reports, showing why reconstruction error alone does not explain matching quality. Synthetic experiments examine these comparisons and the tradeoff between public approximation and privacy noise.

1 INTRODUCTION

Pairing two sets of points to minimize total distance is a common problem in allocation and data analysis. Distance-based matching assigns workers to nearby tasks in spatial crowdsourcing (To et al., 2014), aligns source and target samples in domain adaptation (Courty et al., 2017), and pairs comparable units in randomized experiments to improve the precision of average treatment effect estimates (Bai, 2022). We study minimum-total-distance bipartite matching.

The features used for matching can be sensitive. Differential privacy limits how much an individual's record can affect the distribution of released outputs (Dwork & Roth, 2014). Under its usual replacement-based definition, the same privacy bound applies whether a record changes slightly or is replaced by a distant point. Metric privacy instead scales the bound on distinguishability with the distance between inputs, requiring closer inputs to produce more similar output distributions (Chatzikokolakis et al., 2013).

We consider data sanitized locally before downstream use for matching. The sanitized data can then serve as input to matching algorithms without additional privacy loss. A useful private representation can therefore support existing matching algorithms without requiring a separate privacy mechanism for each solver.

A standard approach perturbs each point with noise calibrated to the underlying metric and then computes a matching from the perturbed data (Chatzikokolakis et al., 2013; Hardt & Talwar, 2010). In Euclidean space, this approach minimizes mean squared reconstruction error among input-independent additive mechanisms satisfying the corresponding privacy constraint (Koufogiannis et al., 2015). Minimizing reconstruction error, however, does not imply minimizing matching error, which depends on preserving comparisons between competing assignments. This mismatch becomes particularly consequential in high dimensions: we show that direct perturbation can systematically distort those comparisons and produce large additional matching costs.

Optimizing the perturbation for utility offers a potential remedy. Bordenabe et al. (2014) construct metric-private mechanisms that minimize expected service loss for a specified prior over a finite location set, while Imola et al. (2022) develop more scalable mechanisms for minimizing individual distortion on finite metric spaces. These objectives evaluate each perturbed point separately. Matching loss instead depends on how all perturbed points are paired under the one-to-one constraint, so their utility guarantees do not directly address it. Their finite-domain formulations also leave open how to obtain matching guarantees for unrestricted Euclidean inputs.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

An opportunity arises when one set of points is public – for example, when workers are assigned to publicly announced inspection sites, or sensitive target data are matched to a public source dataset. These known candidates specify which distance comparisons the sanitized data need to support. Their arrangement can make a much smaller representation sufficient, particularly when they lie near a low-dimensional subspace or form tight clusters (Narayanan & Nelson, 2019). This suggests using the geometry of the public candidate set to guide what information is preserved before adding noise.

Two observations guide our approach. First, shifting all of a point's matching costs equally leaves the optimal assignment unchanged. Controlling errors in relative costs can therefore be more useful than reconstructing every feature accurately. Second, when many pairings have similar costs, the optimizer can accommodate noise without paying much additional true cost. We use these observations to design a new private representation and to explain when directly perturbing the distances used for matching is effective. Our analysis and experiments establish their accuracy gains, as well as how their relative performance depends on the point configuration.

The contributions of our research are as follows.

- We design a locally private representation that supplements a compressed noisy point with one private distance to a selected public candidate, enabling more accurate comparisons between matching costs. For n points in $\mathbb{R}^d$ and privacy rate ε, its expected additional total cost is $O\Big(n\sqrt{\min\{n,d\}\log n}/\varepsilon\Big)$, improving the dimension dependence of the standard perturbation bound without distributional or boundedness assumptions on private points.
- We establish complementary examples in which coordinate-based and distance-based representations reverse their performance ranking. Experiments on uniformly distributed, clustered, and low-dimensional point sets examine these differences. They show when simplifying the public candidate set improves accuracy and when the additional private distance is not worth its privacy cost.

Appendix H reviews related work in metric privacy, private optimization, geometric representations, and location-private assignment.

2 Problem Setup and Preliminaries

Matching and privacy. Let $Q = (q_i)_{i=1}^n$ be public and $R = (r_j)_{j=1}^n$ private, with points in $\mathbb{R}^d$. For a permutation $\sigma \in S_n$, define

$$C_{ij} = \|q_i - r_j\|_2, \qquad u(\sigma) = \sum_i C_{i,\sigma(i)}, \qquad \mathrm{OPT} = \min_{\sigma\in S_n} u(\sigma).$$

Write $D(R,R') = \sum_j \|r_j - r'_j\|_2$. A mechanism M_Q satisfies εD-metric privacy (Chatzikokolakis et al., 2013) if

$$\Pr[M_Q(R) \in A] \le e^{\varepsilon D(R,R')} \Pr[M_Q(R') \in A] \tag{1}$$

for all R, R' and measurable events A. We require a local implementation: each individual independently releases $Y_j = K_{j,Q}(r_j)$, using only their own record and public information, with

$$\Pr[K_{j,Q}(x) \in A] \le e^{\varepsilon\|x-x'\|_2} \Pr[K_{j,Q}(x') \in A]. \tag{2}$$

This holds for all x, x' and measurable A. The objective is to minimize $\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT}$ under this guarantee. Independent local releases satisfy (1); post-processing them into a matching costs no additional privacy. For adaptive releases from one individual, their privacy rates add. Appendix G states these composition properties. All choices based only on Q are public, and all ties use a fixed public rule.

Noise calibration. We use norm-Laplace noise (Chatzikokolakis et al., 2013; Hardt & Talwar, 2010).

Proposition 1 (Lipschitz calibration)**.** *If $h_Q : \mathbb{R}^d \to \mathbb{R}^s$ obeys $\|h_Q(x) - h_Q(x')\|_B \le L\|x - x'\|_2$ for a norm $\|\cdot\|_B$ and $L > 0$, then $h_Q(x) + Z$, with density $p_Z(z) \propto \exp(-\varepsilon\|z\|_B/L)$, satisfies local ε-metric privacy.*

At rate λ, the magnitude has distribution $\|Z\|_B \sim \mathrm{Gamma}(s, \text{rate } \lambda)$ and mean s/λ (Awan & Slavković, 2021). In every mechanism, the server forms $\widehat{C}$ and returns a minimum-cost assignment for this matrix. Pseudocode for the first three mechanisms is in Appendix G.

2.1 Laplace Perturbation of Original Points

Each individual releases $\widetilde{r}_j = r_j + Z_j$, where $p_Z(z) \propto e^{-\varepsilon\|z\|_2}$. The server uses $\widehat{C}_{ij} = \|q_i - \widetilde{r}_j\|_2$.

Theorem 2 (Original-point perturbation)**.** *For every $Q, R \in (\mathbb{R}^d)^n$ and $\varepsilon > 0$, Algorithm 2 satisfies local ε-metric privacy, and its joint releases and output matchings satisfy εD-metric privacy. Moreover,*

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le \frac{2nd}{\varepsilon}. \tag{3}$$

The proof is in Appendix A. The dependence on d motivates smaller representations.

2.2 Distance-Vector Representation

Matching uses distances between private points and public candidates. We can therefore replace each private point by its distance vector

$$\Phi_Q(x) = \big(\|q_1 - x\|_2, \ldots, \|q_n - x\|_2\big).$$

This representation contains all the costs needed for matching and has n coordinates regardless of the original dimension d. Perturbing it directly yields an error bound with no dependence on d.

For $p \in [1, \infty]$, individual j releases $\widetilde{c}_j = \Phi_Q(r_j) + Z_j$ with density $f_p(z) \propto \exp(-\varepsilon\|z\|_p / n^{1/p})$, and the server sets $\widehat{C}_{ij} = [\widetilde{c}_j]_i$.

The server uses the perturbed costs directly, including any negative entries.

Theorem 3 (Distance-vector perturbation)**.** *For every $Q, R \in (\mathbb{R}^d)^n$, $\varepsilon > 0$, and $p \in [1, \infty]$, Algorithm 3 satisfies local ε-metric privacy, and its joint releases and output matchings satisfy εD-metric privacy. Moreover,*

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le 2n\,\mathbb{E}\|Z_1\|_\infty \le \frac{2n^{2+1/p}}{\varepsilon}. \tag{4}$$

Within this family, $p = \infty$ minimizes the noise-based upper bound $2n\,\mathbb{E}\|Z_1\|_\infty$, giving

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le \frac{2n^2}{\varepsilon}. \tag{5}$$

The proof appears in Appendix A. Thus, $p = \infty$ gives the strongest guarantee within this family, and we use this choice henceforth. Its bound improves on the $2nd/\varepsilon$ guarantee for original-point perturbation when $n < d$.

The cube-mixture form of this noise, used in the comparison proofs, is given in Appendix D.1.

3 Projection-Based Private Matching

The public affine span permits a representation that preserves all matching costs while reducing the dimension of the noise.

3.1 Projected Representation

Assume without loss of generality that $q_1 = 0$. Let

$$W = \mathrm{span}\{q_1, \ldots, q_n\}, \qquad k = \dim W,$$

and denote the orthogonal projections onto W and $W^\perp$ by P_W and $P_{W^\perp}$.

For any private point x, its distance to a public candidate satisfies

$$\|x - q_i\|_2^2 = \|P_W x - q_i\|_2^2 + \|P_{W^\perp} x\|_2^2.$$

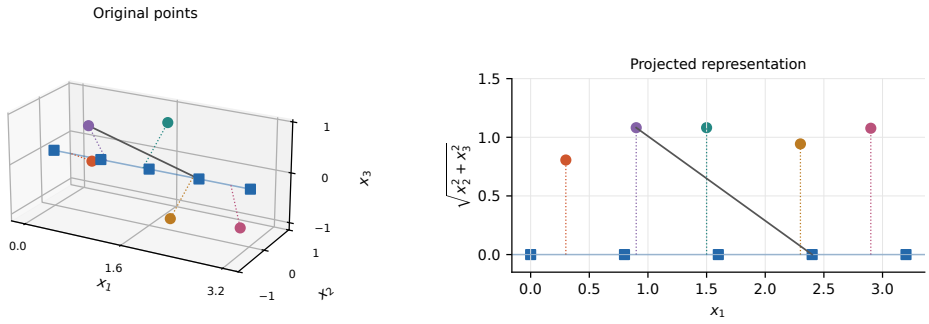

Figure 1: Original and projected coordinates. Squares are public points; matching colors identify the same private point. The transformation preserves every public-to-private distance, including the highlighted segment.

Thus, the component outside W affects every matching cost only through its magnitude. We can retain the projection onto W and replace all remaining coordinates by this single scalar.

Set $m = \min\{k+1, d\}$. Writing projected vectors in a fixed orthonormal basis of W, define $f_Q : \mathbb{R}^d \to \mathbb{R}^m$ by

$$f_Q(x) = \begin{cases} \big(P_W x, \|P_{W^\perp} x\|_2\big), & m < d, \\ x, & m = d. \end{cases} \tag{6}$$

When the construction would not reduce dimension, we retain the original representation. The subspace and its coordinates depend only on Q.

It preserves every public-to-private distance: $\|f_Q(x) - f_Q(q_i)\|_2 = \|x - q_i\|_2$. This construction is also used for exact terminal dimensionality reduction in Narayanan & Nelson (2019).

Individual j releases $Y_j = f_Q(r_j) + Z_j$, with $p_Z(z) \propto e^{-\varepsilon\|z\|_2}$ in $\mathbb{R}^m$. The server uses $\widehat{C}_{ij} = \|f_Q(q_i) - Y_j\|_2$.

Theorem 4 (Projected representation)**.** *For every $Q, R \in (\mathbb{R}^d)^n$ and $\varepsilon > 0$, Algorithm 4 satisfies local ε-metric privacy, and its joint releases and output matchings satisfy εD-metric privacy. Moreover,*

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le \frac{2nm}{\varepsilon} \le \frac{2n\min\{n,d\}}{\varepsilon}. \tag{7}$$

The proof is in Appendix B. If the public points lie on a line, two coordinates suffice regardless of the private points.

Figure 1 illustrates the representation when the public points lie on the first coordinate axis in three dimensions. It retains a private point's coordinate along that line and its distance to the line: $(x_1, x_2, x_3) \mapsto (x_1, \sqrt{x_2^2 + x_3^2})$. Every distance to a public candidate is preserved, although distances between private points need not be. The illustration shows the transformation before noise is added.

3.2 Corrected Projected Representation

Projection reduces the dimension in which noise is added, but direct distance reconstruction still carries error shared across all candidates. We remove this shared error by subtracting squared distances before taking square roots. Subtraction also removes the true distance to one selected candidate; one additional private scalar restores this missing information.

Write $z_j = f_Q(r_j)$, $p_i = f_Q(q_i)$, and $Y_j = z_j + Z_j$. Every squared distance contains $\|Z_j\|_2^2$, whose effect after taking square roots depends on the true distance. For a selected candidate q_{a_j}, subtracting squared distances gives

$$\|Y_j - p_i\|_2^2 - \|Y_j - p_{a_j}\|_2^2 = \|z_j - p_i\|_2^2 - \|z_j - p_{a_j}\|_2^2 + 2\langle Z_j, p_{a_j} - p_i\rangle.$$

Both the squared noise magnitude and the private point's contribution to the noise inner product cancel. The remaining noise depends on a direction between public candidates.

Adding t_j^2, where $t_j = \|z_j - p_{a_j}\|_2 = \|r_j - q_{a_j}\|_2$, restores the desired squared distance apart from the inner-product error. To protect this scalar, release

$$T_j = \max\{0, t_j + \eta_j\}, \qquad p_s(t) \propto e^{-\varepsilon_s|t|}, \quad t \in \mathbb{R}, \tag{8}$$

where η_j has density p_s and is independent of Z_j. The server reconstructs

$$\widehat{C}_{ij} = \sqrt{T_j^2 + \|Y_j - p_i\|_2^2 - \|Y_j - p_{a_j}\|_2^2}. \tag{9}$$

Choosing a_j nearest to Y_j makes the difference of squared distances nonnegative, so the reconstructed costs are always well-defined. The server can recover a_j from Y_j and Q, using a fixed public rule for ties.

Algorithm 1 Matching with a corrected projected representation

Require: Public points Q, private points R, privacy parameters $\varepsilon_y, \varepsilon_s > 0$

1: Construct f_Q from Q and set $p_i = f_Q(q_i)$ for all i.
2: **for** each individual j, independently **do**
3: Set $z_j = f_Q(r_j)$.
4: Draw $Z_j \in \mathbb{R}^m$ with density $p_y(z) \propto e^{-\varepsilon_y\|z\|_2}$ and set $Y_j = z_j + Z_j$.
5: Choose $a_j \in \arg\min_i \|Y_j - p_i\|_2$, using a fixed public rule for ties.
6: Independently draw $\eta_j \in \mathbb{R}$ with density $p_s(t) \propto e^{-\varepsilon_s|t|}$.
7: Set $T_j = \max\{0, \|z_j - p_{a_j}\|_2 + \eta_j\}$ and release (Y_j, T_j).
8: **end for**
9: Recover each a_j from Y_j and compute $\widehat{C}_{ij} = \sqrt{T_j^2 + \|Y_j - p_i\|_2^2 - \|Y_j - p_{a_j}\|_2^2}$.
10: Find $\widehat{\sigma} \in \arg\min_{\sigma \in S_n} \sum_{i=1}^n \widehat{C}_{i,\sigma(i)}$.
11: **return** $\widehat{\sigma}$.

Theorem 5 (Corrected projected representation)**.** *For every public point set Q and $\varepsilon > 0$, there exist $\varepsilon_y, \varepsilon_s > 0$, depending only on m, n, ε, with $\varepsilon_y + \varepsilon_s = \varepsilon$, such that, for every private point set R, Algorithm 1 satisfies local ε-metric privacy, and its joint releases and output matchings satisfy εD-metric privacy. Moreover,*

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} = O\left(\frac{n\sqrt{m\log(2n)}}{\varepsilon}\right). \tag{10}$$

The proof and privacy split are in Appendix C. The correction replaces the factor m by $\sqrt{m\log(2n)}$ using one additional scalar. Since $m \le \min\{n, d\}$, it improves both baseline bounds when m is large relative to $\log n$. No proximity assumption on the private points is required.

4 Two geometries distinguish the mechanisms

Two constructions compare the mechanisms at the same public affine dimension. Rates concern total excess as even $n = d$ tends to infinity, with point sets scaled as n/ε; they do not describe bounded uniform data.

Proposition 6. *There are two families with the same n, d, ε and public affine dimension $n-1$ having the following expected excess costs:*

	Projected representation	*Distance vectors*	*Corrected projected representation*
Two populations	$\Theta(n^2/\varepsilon)$	$O(n/\varepsilon)$	$O(n^{3/2}\sqrt{\log n}/\varepsilon)$
Unique partners	$o(n^2/\varepsilon)$	$\Theta(n^2/\varepsilon)$	$O(n^{3/2}\sqrt{\log n}/\varepsilon)$

For both families, the full projected representation and original-point mechanism have the same cost distribution.

Two populations. Public candidates lie near two locations separated by order n/ε. Both private populations prefer the same side, but one benefits more. Noise in the projected representation reverses their relative preferences, so the wrong population receives those candidates and the loss is $\Theta(n^2/\varepsilon)$. Distance vectors exploit the repeated private locations; the correction instead controls noise projections. The failure persists under independent Gaussian perturbations with fixed coordinate variance. Replacing public points by their two centers permits a two-dimensional report and an $O(n/\varepsilon)$ bound for the unperturbed example. The full construction and proofs are in Appendix D.

Unique partners. Set $q_i = r_i = (n/(4\sqrt{2}\varepsilon))e_i$. Correct partners cost zero and incorrect partners cost $n/(4\varepsilon)$. Distance noise favors wrong transpositions with constant probability in linearly many disjoint pairs. Projected noise has small directional components relative to the separation, so the identity is recovered with probability tending to one. The numerical comparison tests the ordering at finite sizes, not these asymptotic limits.

5 Numerical results

We compare the four mechanisms at the same privacy rate per individual. Assignments are evaluated on the original points. Let RAND be the expected total cost of a uniformly random assignment. We report *relative matching cost*,

$$\frac{\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT}}{\mathrm{RAND} - \mathrm{OPT}}. \tag{11}$$

Zero means optimal matching, one means random-matching performance, and lower is better; values above one are retained. For each geometry, we average noise draws before dividing by its own $\mathrm{RAND} - \mathrm{OPT}$. Randomized studies average these ratios over 20 geometries with 50 draws each. Bands are pointwise 95% intervals over geometry means, except for fixed configurations, where they quantify noise uncertainty. All vertical axes are linear. Generators, seeds, and interval formulas are in Appendix F.

Original dimension and uniform points. Figure 2 (left) embeds the same 128 public and private four-dimensional Gaussian points in $d = 8, 32, 128, 512$ by appending zeros; $\varepsilon = 64$. Its horizontal axis is d on a logarithmic scale. True costs are unchanged. The projected mechanism uses five coordinates and stays near 0.0062, while original-point perturbation increases from 0.0092 to 0.284. The corrected mechanism also avoids dependence on d, but its extra scalar does not improve performance here.

Figure 2 (right) fixes 20 public and private points sampled independently from $[0, 1]^{50}$. The horizontal axis is ε on a logarithmic scale; larger values permit less noise. Each rate uses 100 draws. At $\varepsilon = 16$, relative costs are 0.302 for the projected representation, 0.346 for its correction, 0.463 for original points, and 0.622 for distance vectors. A stronger uniform bound need not improve every instance. Two additional configurations are in Figure 6.

Retained public dimension. With $n = d = 128$, generate public points with residual variation of size a outside four leading directions. Their full affine dimension is 127. For each specified r, project Q onto its centered principal subspace, then apply the projected mechanism to the original private points, retaining both subspace coordinates and distance to the subspace. Evaluate the assignment against the original Q, R. This trades public approximation error (Appendix E) against noise dimension. Private points are independent and have heterogeneous residual variation.

In Figure 3, points and privacy rate are fixed within each panel. Horizontal positions are discrete evaluated ranks, not equally spaced numerical increments. At $r = 0$, a fixed assignment needs no report. The dashed line uses all 127 directions; the dotted line adds no noise and is a nonprivate diagnostic. For $a = 0.02$, retaining four directions reduces relative cost from 0.0575 to 0.0046 at $\varepsilon = 256$, and from 0.338 to 0.0450 at $\varepsilon = 64$. With $a = 0.2$ and $\varepsilon = 256$, dimension four instead gives 0.0693, compared with 0.0616 at full dimension and 0.0404 at the best measured intermediate dimension. Discarding too many directions can therefore hurt. Figure 5 gives additional rates.

Contrasting point arrangements. Figure 4 evaluates Section 4 at $n = d \in \{32, 64, 128, 256\}$ and $\varepsilon = 1$, with a logarithmic horizontal axis. The spatial scale grows with n. The two populations re-

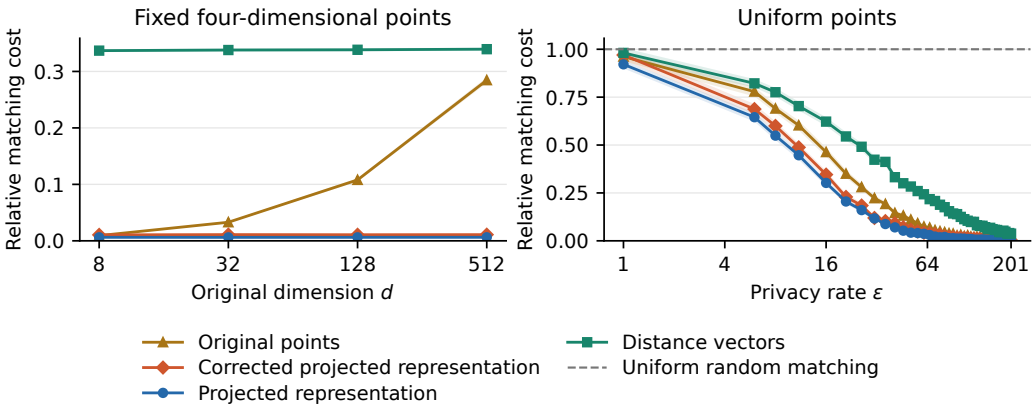

Figure 2: Relative matching cost versus original dimension (left) and privacy rate (right). Lower is better. Projection avoids the cost of unused coordinates and improves performance on the uniform configuration.

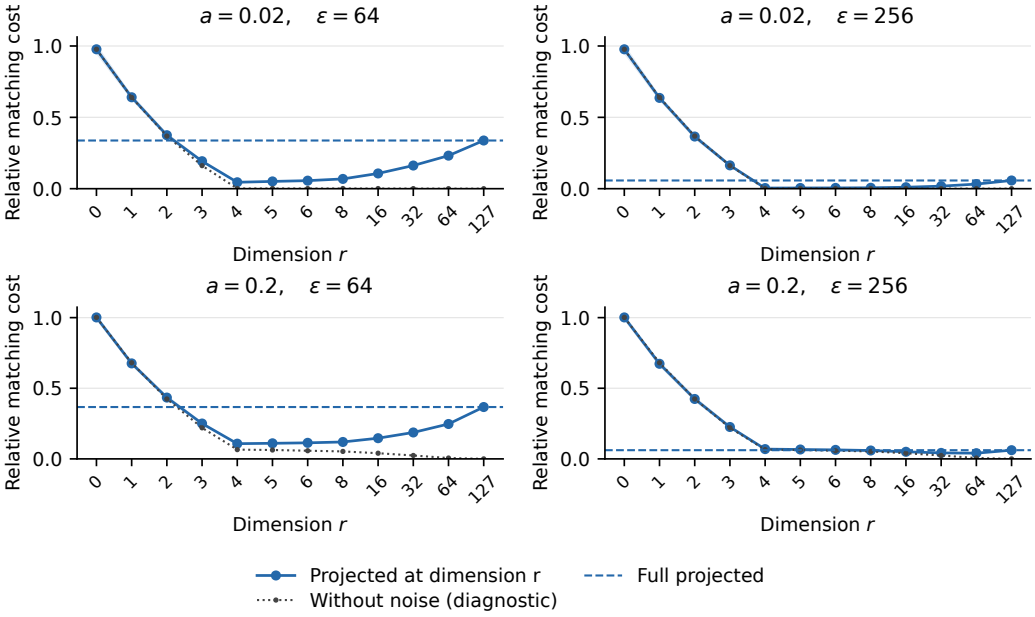

Figure 3: Relative matching cost versus retained dimension r. Lower is better. Four directions work well with small residual variation; more can help when it is larger. The dotted curve is a nonprivate diagnostic.

ceive independent Gaussian perturbations with coordinate standard deviation 0.1; the unique-partner simplex is deterministic. Both use 50 noise draws. Original-point and projected mechanisms have identical reconstructed-cost laws here and reuse samples, so one curve represents both. At $n = 256$, projected, corrected, and distance-vector relative costs are 1.253, 0.628, and 0.205 for two populations, versus 0.483, 0.478, and 0.880 for unique partners. The ranking reverses, although these sizes do not establish the asymptotic rates.

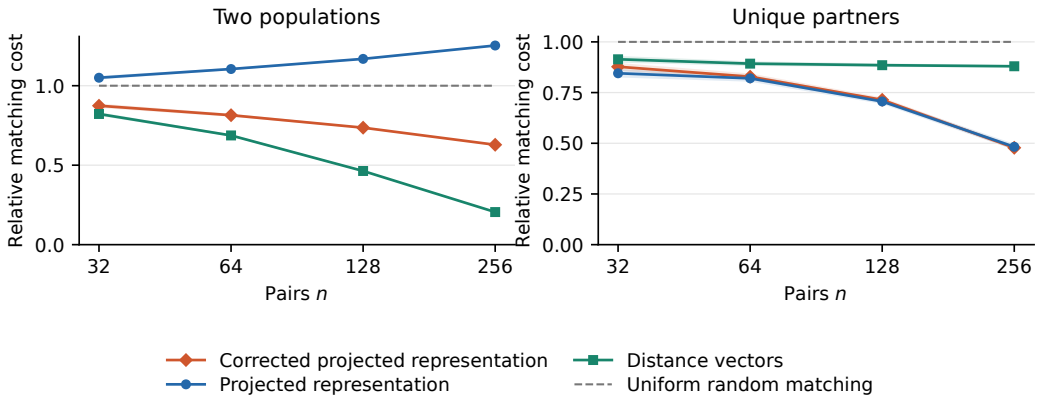

Figure 4: Relative matching cost versus number of pairs n. Lower is better; one is random matching. Distance vectors are better for two populations, while projected mechanisms are better for unique partners.

6 Discussion

We show how public geometry can guide the design of representations for locally private matching. Projection preserves the distances needed for assignment while reducing the dimension of the private report. The corrected representation then uses one additional private distance to control errors in relative matching costs, yielding a stronger utility guarantee without distributional or boundedness assumptions on the private points. This establishes a direct connection between the information preserved before perturbation and the accuracy of the resulting assignment.

The complementary constructions and numerical experiments explain how these representations behave across different geometries. They identify configurations favoring distance reports or projected reports and demonstrate when approximating public points improves the tradeoff between distortion and privacy noise. Together, the theoretical guarantees and empirical comparisons provide a foundation for designing private representations around the matching objective. The central contribution is a concrete way to turn public geometric structure into improved matching guarantees, supported by an analysis of the assignment comparisons that determine performance.

AI USE STATEMENT

In this work, we used generative AI tools to design and provide feedback on research methodology and experiments, implement methods, and interpret results. We also used these tools to develop and write proofs for the complementary constructions in Appendix D. Generative AI tools were used to write and edit all research and experimental code, draft and revise the manuscript, and search for and organize the literature discussed in the related-work section. This assistance included locating references for the K-norm mechanism and terminal dimensionality reduction. The synthetic point sets in our experiments were generated by the specified seeded random samplers, implemented with AI assistance, rather than by sampling a generative AI model. Translation and qualitative data analysis are not applicable to this work.

We have reviewed all AI-assisted work. The authors executed and verified the research code, critically assessed AI-supported experimental designs and interpretations, checked the mathematical arguments, and verified AI-identified references through manual literature searches. We take full responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

ETHICS STATEMENT

We do not identify ethical concerns specific to this work. The study is theoretical and uses synthetic point sets; it involves no human participants or personal data.

REPRODUCIBILITY STATEMENT

The assumptions and mechanisms are specified in Sections 2 and 3, with additional pseudocode in Appendix G. Complete proofs are provided in Appendices A–E. Section 5 defines the evaluation metric, and Appendix F describes the synthetic data generators, experimental configurations, and uncertainty calculations. The accompanying reproducibility materials include the source code, configurations, random seeds, software requirements, source results, and scripts for rerunning the simulations and generating every figure. Their README gives execution instructions, and the validation records document checks against the reported results.

REFERENCES

Mário S. Alvim, Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Anna Pazii. Local differential privacy on metric spaces: Optimizing the trade-off with utility. In *2018 IEEE 31st Computer Security Foundations Symposium*, pp. 262–267, 2018. doi: 10.1109/CSF.2018.00026. URL https://conferences.computer.org/sp/pdfs/csf/2018/InvitedPaperLocalDifferentialPrivacyonMetricS.pdf.

Balamurugan Anandan and Chris Clifton. Secure minimum weighted bipartite matching. In *IEEE Conference on Dependable and Secure Computing*, 2017. doi: 10.1109/DESEC.2017.8073798.

Miguel E. Andrés, Nicolás E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Geo-indistinguishability: Differential privacy for location-based systems. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, pp. 901–914, 2013. URL https://arxiv.org/abs/1212.1984.

Jordan Awan and Aleksandra Slavković. Structure and sensitivity in differential privacy: Comparing K-norm mechanisms. *Journal of the American Statistical Association*, 2021. URL https://arxiv.org/abs/1801.09236.

Arturs Backurs, Zinan Lin, Sepideh Mahabadi, Sandeep Silwal, and Jakub Tarnawski. Efficiently computing similarities to private datasets. In *International Conference on Learning Representations*, 2024. URL https://arxiv.org/abs/2403.08917.

Yuehao Bai. Optimality of matched-pair designs in randomized controlled trials. *American Economic Review*, 112(12):3911–3940, 2022. doi: 10.1257/aer.20201856. URL https://arxiv.org/abs/2206.07845.

Alexander Benvenuti, Brendan Bialy, Miriam Dennis, and Matthew Hale. Differentially private linear programming: Reduced sub-optimality and guaranteed constraint satisfaction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(24):19702–19710, 2026. doi: 10.1609/aaai.v40i24.39051.

March Boedihardjo, Thomas Strohmer, and Roman Vershynin. Metric geometry of the privacy-utility tradeoff. arXiv preprint arXiv:2405.00329, 2024. URL https://arxiv.org/abs/2405.00329.

Nicolás E. Bordenabe, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Optimal geo-indistinguishable mechanisms for location privacy. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, pp. 251–262, 2014. doi: 10.1145/2660267.2660345.

Konstantinos Chatzikokolakis, Miguel E. Andrés, Nicolás Emilio Bordenabe, and Catuscia Palamidessi. Broadening the scope of differential privacy using metrics. In *Privacy Enhancing Technologies*, pp. 82–102. Springer, 2013. doi: 10.1007/978-3-642-39077-7_5.

Joann Qiongna Chen, Tianhao Wang, Zhikun Zhang, Yang Zhang, Somesh Jha, and Zhou Li. Differentially private resource allocation. In *Annual Computer Security Applications Conference*, pp. 772–786, 2023. doi: 10.1145/3627106.3627181.

Nicolas Courty, Rémi Flamary, Devis Tuia, and Alain Rakotomamonjy. Optimal transport for domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(9):1853–1865, 2017. doi: 10.1109/TPAMI.2016.2615921. URL https://arxiv.org/abs/1507.00504.

Panayiotis Danassis, Aleksei Triastcyn, and Boi Faltings. A distributed differentially private algorithm for resource allocation in unboundedly large settings. arXiv preprint arXiv:2011.07934, 2022. URL https://arxiv.org/abs/2011.07934.

Michael Dinitz, George Z. Li, Quanquan C. Liu, and Felix Zhou. Differentially private matchings. arXiv preprint arXiv:2501.00926, 2025. URL https://arxiv.org/abs/2501.00926.

Leilei Du, Peng Cheng, Libin Zheng, Wei Xi, Xuemin Lin, Wenjie Zhang, and Jing Fang. Dynamic private task assignment under differential privacy. In *IEEE 39th International Conference on Data Engineering*, pp. 2740–2752, 2023. URL https://arxiv.org/abs/2302.09511.

Cynthia Dwork and Aaron Roth. *The Algorithmic Foundations of Differential Privacy*. Now Publishers, 2014. URL https://www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf.

Alina Ene, Huy Le Nguyen, Ta Duy Nguyen, and Adrian Vladu. Solving linear programs with differential privacy. arXiv preprint arXiv:2507.10946, 2025. URL https://arxiv.org/abs/2507.10946.

Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, pp. 705–714, 2010. URL https://arxiv.org/abs/0907.3754.

Justin Hsu, Zhiyi Huang, Aaron Roth, Tim Roughgarden, and Zhiwei Steven Wu. Private matchings and allocations. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, pp. 21–30, 2014a. doi: 10.1145/2591796.2591826. URL https://arxiv.org/abs/1311.2828.

Justin Hsu, Aaron Roth, Tim Roughgarden, and Jonathan Ullman. Privately solving linear programs. In *Automata, Languages, and Programming*, pp. 612–624. Springer, 2014b. doi: 10.1007/978-3-662-43948-7_51. URL https://arxiv.org/abs/1402.3631.

Justin Hsu, Zhiyi Huang, Aaron Roth, and Zhiwei Steven Wu. Jointly private convex programming. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 580–599, 2016. doi: 10.1137/1.9781611974331.ch43. URL https://arxiv.org/abs/1411.0998.

Zhiyi Huang and Aaron Roth. Exploiting metric structure for efficient private query release. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 523–534, 2014. doi: 10.1137/1.9781611973402.39. URL https://arxiv.org/abs/1211.7302.

Jacob Imola, Shiva Kasiviswanathan, Stephen White, Abhinav Aggarwal, and Nathanael Teissier. Balancing utility and scalability in metric differential privacy. In *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence*, volume 180 of *Proceedings of Machine Learning Research*, pp. 885–894. PMLR, 2022. URL https://proceedings.mlr.press/v180/imola22a.html.

Jacob Imola, Amrita Roy Chowdhury, and Kamalika Chaudhuri. Metric differential privacy at the user-level via the earth-mover's distance. In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*, pp. 348–362, 2024. doi: 10.1145/3658644.3690363.

Jacob Imola, Rasmus Pagh, and Lukas Retschmeier. Lower bounds for private graph optimization problems using reconstruction attacks. arXiv preprint arXiv:2609.10877, 2026. URL https://arxiv.org/abs/2609.10877.

Krishnaram Kenthapadi, Aleksandra Korolova, Ilya Mironov, and Nina Mishra. Privacy via the Johnson–Lindenstrauss transform. *Journal of Privacy and Confidentiality*, 5(1):39–71, 2013. URL https://cs.stanford.edu/people/kngk/papers/privacyViaJohnsonLindenstraussTransform-JPC2013.pdf.

Fragkiskos Koufogiannis, Shuo Han, and George J. Pappas. Optimality of the Laplace mechanism in differential privacy, 2015. URL https://arxiv.org/abs/1504.00065.

Nam LeTien, Amaury Habrard, and Marc Sebban. Differentially private optimal transport: Application to domain adaptation. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence*, pp. 2852–2858, 2019. doi: 10.24963/ijcai.2019/395. URL https://www.ijcai.org/proceedings/2019/395.

Maocheng Li, Jiachuan Wang, Libin Zheng, Han Wu, Peng Cheng, Lei Chen, and Xuemin Lin. Privacy-preserving batch-based task assignment in spatial crowdsourcing with untrusted server. In *Proceedings of the Thirtieth ACM International Conference on Information and Knowledge Management*, pp. 947–956, 2021. doi: 10.1145/3459637.3482288.

Svante Linusson and Johan Wästlund. A proof of Parisi's conjecture on the random assignment problem. *Probability Theory and Related Fields*, 128:419–440, 2004. URL https://arxiv.org/abs/math/0303214.

Chaojie Lv, Lan Zhang, and Xiang-Yang Li. Personalized differentially private online minimum bipartite matching in spatial crowdsourcing. In *8th International Conference on Big Data Computing and Communications*, pp. 134–143, 2022. doi: 10.1109/BIGCOM57025.2022.00025.

Shyam Narayanan and Jelani Nelson. Optimal terminal dimensionality reduction in Euclidean space. In *Proceedings of the Fifty-First Annual ACM SIGACT Symposium on Theory of Computing*, pp. 1064–1069, 2019. URL https://arxiv.org/abs/1810.09250.

Amanda Prorok and Vijay Kumar. Privacy-preserving vehicle assignment for mobility-on-demand systems. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1869–1876, 2017. doi: 10.1109/IROS.2017.8206003. URL https://arxiv.org/abs/1703.04738.

Adam Sealfon. Shortest paths and distances with differential privacy. In *Proceedings of the Thirty-Fifth ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, pp. 29–41, 2016. doi: 10.1145/2902251.2902291. URL https://arxiv.org/abs/1511.04631.

Qian Tao, Yongxin Tong, Zimu Zhou, Yexuan Shi, Lei Chen, and Ke Xu. Differentially private online task assignment in spatial crowdsourcing: A tree-based approach. In *IEEE 36th International Conference on Data Engineering*, pp. 517–528, 2020. doi: 10.1109/ICDE48307.2020.00051.

Hien To, Gabriel Ghinita, and Cyrus Shahabi. A framework for protecting worker location privacy in spatial crowdsourcing. *Proceedings of the VLDB Endowment*, 7(10):919–930, 2014. URL https://www.vldb.org/pvldb/vol7/p919-to.pdf.

Hengzhi Wang, En Wang, Yongjian Yang, Jie Wu, and Falko Dressler. Privacy-preserving online task assignment in spatial crowdsourcing: A graph-based approach. In *IEEE INFOCOM 2022 – IEEE Conference on Computer Communications*, pp. 570–579, 2022. doi: 10.1109/INFOCOM48880.2022.9796827.

Jiaqi Wang and Ziv Goldfeld. Differentially private Sinkhorn algorithm. In *60th Annual Allerton Conference on Communication, Control, and Computing*, 2024. doi: 10.1109/ALLERTON63246.2024.10735319. URL https://par.nsf.gov/servlets/purl/10577813.

Xinpeng Xie, Chenyang Yu, Yan Huang, Yang Cao, and Chenxi Qiu. SoK: Metric differential privacy in theory and practice. *Proceedings on Privacy Enhancing Technologies*, 2026(4):415–437, 2026. doi: 10.56553/popets-2026-0128. URL https://petsymposium.org/popets/2026/popets-2026-0128.php.

Yifan Xu, Jun Tao, Xinyu Chen, Jun Yan, and Jun Shen. Efficient privacy-preserving ridesharing: An online matching-based approach. *IEEE Transactions on Information Forensics and Security*, 21:5485–5498, 2026. doi: 10.1109/TIFS.2026.3698576.

Pengfei Zhang, Xiang Cheng, Sen Su, and Ning Wang. Task allocation under geo-indistinguishability via group-based noise addition. *IEEE Transactions on Big Data*, 9(3):860–877, 2023. doi: 10.1109/TBDATA.2022.3215467.

Zongrui Zou, Alessandro Epasto, Chenglin Fan, and Rudrajit Das. Differentially private minimum spanning tree in Euclidean graphs. In *Proceedings of the Twenty-Ninth International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 379–387. PMLR, 2026. URL https://proceedings.mlr.press/v300/zou26a.html.

A MATCHING ERROR UNDER PERTURBATION

Throughout the proofs, $\mathcal{R}_M(Q, R)$ denotes $\mathbb{E}[u(\widehat{\sigma}_M)] - \mathrm{OPT}$ for mechanism M; subscripts proj and dist refer to the projected and distance-vector mechanisms. This quantity is in total distance units.

We first establish a deterministic bound that converts errors in the cost matrix into excess matching cost.

Lemma 7 (Matching stability). *Let $C \in \mathbb{R}^{n\times n}$ be a cost matrix, let $\widehat{C} = C + E$, and let*

$$\widehat{\sigma} \in \arg\min_{\sigma\in S_n} \sum_{i=1}^{n} \widehat{C}_{i,\sigma(i)}.$$

Writing $u(\sigma) = \sum_i C_{i,\sigma(i)}$ and $\mathrm{OPT} = \min_{\sigma\in S_n} u(\sigma)$, we have

$$0 \le u(\widehat{\sigma}) - \mathrm{OPT} \le \sum_{j=1}^{n} \left(\max_i E_{ij} - \min_i E_{ij}\right). \tag{12}$$

Proof. Let σ^* minimize the true cost. Optimality of $\widehat{\sigma}$ for $\widehat{C}$ gives

$$u(\widehat{\sigma}) + \sum_{i=1}^{n} E_{i,\widehat{\sigma}(i)} \le u(\sigma^*) + \sum_{i=1}^{n} E_{i,\sigma^*(i)}.$$

Consequently,

$$u(\widehat{\sigma}) - \mathrm{OPT} \le \sum_{i=1}^{n} E_{i,\sigma^*(i)} - \sum_{i=1}^{n} E_{i,\widehat{\sigma}(i)}.$$

Each permutation selects exactly one entry from every column. Therefore,

$$\sum_{i=1}^{n} E_{i,\sigma^*(i)} \le \sum_{j=1}^{n} \max_i E_{ij}, \qquad \sum_{i=1}^{n} E_{i,\widehat{\sigma}(i)} \ge \sum_{j=1}^{n} \min_i E_{ij}.$$

Combining these inequalities proves the upper bound. Nonnegativity follows from the definition of OPT. □

The lemma shows that an error shared by every entry in a column does not affect matching quality; only variation within that column enters the bound.

Proof of Theorem 2. The identity map on $\mathbb{R}^d$ is 1-Lipschitz in the Euclidean norm. Proposition 1 therefore gives local ε-metric privacy for each release $r_j + Z_j$. Parallel composition (Proposition 8) gives εD-metric privacy for the joint releases, and post-processing (Proposition 9) preserves this guarantee when computing $\widehat{C}$ and $\widehat{\sigma}$.

For the error bound, write

$$E_{ij} = \|q_i - r_j - Z_j\|_2 - \|q_i - r_j\|_2.$$

The reverse triangle inequality gives

$$|E_{ij}| \le \|Z_j\|_2.$$

Hence, for every realization of the noise,

$$\max_i E_{ij} - \min_i E_{ij} \le 2\|Z_j\|_2.$$

Applying Lemma 7 yields

$$u(\widehat{\sigma}) - \mathrm{OPT} \le 2\sum_{j=1}^{n} \|Z_j\|_2.$$

Since

$$\|Z_j\|_2 \sim \mathrm{Gamma}(d, \text{rate } \varepsilon), \qquad \mathbb{E}\|Z_j\|_2 = \frac{d}{\varepsilon},$$

taking expectations gives

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le 2\sum_{j=1}^{n} \mathbb{E}\|Z_j\|_2 = \frac{2nd}{\varepsilon}.$$

□

Proof of Theorem 3. For each public candidate q_i, the reverse triangle inequality gives

$$|\|q_i - x\|_2 - \|q_i - x'\|_2| \le \|x - x'\|_2.$$

Consequently, for every $p \in [1, \infty]$,

$$\|\Phi_Q(x) - \Phi_Q(x')\|_p \le n^{1/p}\|x - x'\|_2.$$

Proposition 1 with noise rate $\varepsilon/n^{1/p}$ therefore gives local ε-metric privacy. Parallel composition and post-processing establish the guarantee for the joint releases and output matchings.

For utility, the cost error in column j is exactly Z_j. Lemma 7 gives

$$u(\widehat{\sigma}) - \mathrm{OPT} \le \sum_{j=1}^{n} \left(\max_i [Z_j]_i - \min_i [Z_j]_i\right) \le 2\sum_{j=1}^{n} \|Z_j\|_\infty.$$

The noise magnitude satisfies

$$\|Z_j\|_p \sim \mathrm{Gamma}\left(n, \text{rate } \frac{\varepsilon}{n^{1/p}}\right), \qquad \mathbb{E}\|Z_j\|_p = \frac{n^{1+1/p}}{\varepsilon}.$$

Using $\|Z_j\|_\infty \le \|Z_j\|_p$ and taking expectations yields

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le 2n\,\mathbb{E}\|Z_1\|_\infty \le \frac{2n^{2+1/p}}{\varepsilon}.$$

To compare the noise-based bounds across p, use

$$\|z\|_\infty \ge n^{-1/p}\|z\|_p.$$

It follows that

$$\mathbb{E}\|Z_1\|_\infty \ge n^{-1/p}\mathbb{E}\|Z_1\|_p = \frac{n}{\varepsilon}.$$

At $p = \infty$, equality holds:

$$\mathbb{E}\|Z_1\|_\infty = \frac{n}{\varepsilon}.$$

Thus $p = \infty$ minimizes $2n\,\mathbb{E}\|Z_1\|_\infty$ within the stated family, with value $2n^2/\varepsilon$. □

B Proof of the Projected-Representation Bound

Proof of Theorem 4. We first verify that f_Q preserves the matching costs and is 1-Lipschitz. Both properties are immediate when $m = d$, since f_Q is the identity. Otherwise, $q_i \in W$ gives

$$\begin{aligned}\|f_Q(x) - f_Q(q_i)\|_2^2 &= \|P_W x - q_i\|_2^2 + \|P_{W^\perp} x\|_2^2 \\ &= \|x - q_i\|_2^2.\end{aligned}$$

For any $x, x' \in \mathbb{R}^d$, the reverse triangle inequality gives

$$\begin{aligned}\|f_Q(x) - f_Q(x')\|_2^2 &= \|P_W(x - x')\|_2^2 + \left(\|P_{W^\perp} x\|_2 - \|P_{W^\perp} x'\|_2\right)^2 \\ &\le \|P_W(x - x')\|_2^2 + \|P_{W^\perp}(x - x')\|_2^2 \\ &= \|x - x'\|_2^2.\end{aligned}$$

Proposition 1 therefore gives local ε-metric privacy for each release. Parallel composition and post-processing establish the guarantee for the joint releases and output matchings.

For utility, preservation of the true distances implies

$$C_{ij} = \|f_Q(q_i) - f_Q(r_j)\|_2.$$

Consequently,

$$\begin{aligned} |E_{ij}| &= |\|f_Q(q_i) - f_Q(r_j) - Z_j\|_2 - \|f_Q(q_i) - f_Q(r_j)\|_2| \\ &\le \|Z_j\|_2. \end{aligned}$$

Applying Lemma 7 and taking expectations yields

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le 2\sum_{j=1}^{n} \mathbb{E}\|Z_j\|_2 = \frac{2nm}{\varepsilon},$$

where the last equality uses

$$\|Z_j\|_2 \sim \mathrm{Gamma}(m, \text{rate } \varepsilon).$$

Finally, $q_1 = 0$ implies $k \le n-1$, so

$$m = \min\{k+1, d\} \le \min\{n, d\}.$$

This proves the stated bound. □

C Proof of the Corrected-Projection Bound

Proof of Theorem 5. All norms in this proof are Euclidean. We first establish privacy and an error bound for arbitrary positive $\varepsilon_y, \varepsilon_s$, then choose their split using only m, n, ε.

Privacy. The representation f_Q is 1-Lipschitz by Appendix B, so the first release Y_j satisfies local ε_y-metric privacy. Conditional on any fixed first output y, the selected index $a(y)$ is fixed, and

$$\|f_Q(x) - p_{a(y)}\| = \|x - q_{a(y)}\|$$

is a 1-Lipschitz function of x. Adding scalar noise with density proportional to $e^{-\varepsilon_s|t|}$ gives local ε_s-metric privacy; clipping at zero is post-processing. Sequential composition therefore gives local $(\varepsilon_y + \varepsilon_s)$-metric privacy for the pair (Y_j, T_j). Parallel composition and post-processing give the sum-metric guarantee for the joint releases and output matching.

Reconstruction error. Let $\mathcal{V}$ contain the unit vectors $(p_i - p_\ell)/\|p_i - p_\ell\|$ for $i < \ell$ with $p_i \ne p_\ell$, and define

$$M(w) = \max\big(\{0\} \cup \{|\langle w, v\rangle| : v \in \mathcal{V}\}\big).$$

This definition also covers coincident public points, when $\mathcal{V}$ is empty and $M(w) = 0$. Fix a column and suppress j. Write $y = z + w$, $a = a(y)$, $d_i = \|z - p_i\|$, and $t = d_a$. Put

$$D_i = \|y - p_i\|^2 - \|y - p_a\|^2 \ge 0, \qquad h_i = \sqrt{t^2 + D_i}.$$

Expansion gives

$$h_i^2 - d_i^2 = 2\langle w, p_a - p_i\rangle.$$

Since $h_i \ge t$, the triangle inequality yields

$$|h_i^2 - d_i^2| \le 2M(w)\|p_a - p_i\| \le 2M(w)(t + d_i) \le 2M(w)(h_i + d_i).$$

If $h_i + d_i = 0$, both distances vanish. Otherwise division gives $|h_i - d_i| \le 2M(w)$. Thus the range of $h_i - d_i$ is at most $4M(w)$.

Now replace t with $T = \max\{0, t + \eta\}$. For every $D_i \ge 0$, the function $v \mapsto \sqrt{v^2 + D_i}$ is nondecreasing and 1-Lipschitz on $[0, \infty)$. Consequently, the changes $\widehat{C}_i - h_i$ all have the same sign and magnitude at most $|T - t| \le |\eta|$. Their range is at most $|\eta|$. Therefore, with $E_{ij} = \widehat{C}_{ij} - C_{ij}$,

$$\max_i E_{ij} - \min_i E_{ij} \le 4M(Z_j) + |\eta_j|. \tag{13}$$

Lemma 7 now gives

$$u(\widehat{\sigma}) - \mathrm{OPT} \le \sum_{j=1}^{n} \big(4M(Z_j) + |\eta_j|\big).$$

Expected noise projections. Write Euclidean Laplace noise of rate λ as $Z = H\Omega$, where $H \sim \mathrm{Gamma}(m, \text{rate } \lambda)$ and Ω is an independent uniform unit vector. For any unit vector v,

$$\mathbb{E}\langle \Omega, v\rangle^{2\ell} = \frac{(2\ell-1)!!}{m(m+2)\cdots(m+2\ell-2)} \le \frac{(2\ell-1)!!}{m^{\ell}}.$$

The equality follows by decomposing a standard Gaussian vector into its independent radius and direction and comparing moments. Odd moments vanish by symmetry, so expanding the exponential gives

$$\mathbb{E}e^{s\langle \Omega, v\rangle} \le e^{s^2/(2m)}.$$

There are at most $n(n-1)/2$ directions in $\mathcal{V}$. Applying this bound to both signs of each projection, and including the constant zero, gives for every $s > 0$

$$\mathbb{E}M(\Omega) \le \frac{\log(2n^2)}{s} + \frac{s}{2m}.$$

Optimizing over s, using $\log(2n^2) \le 2\log(2n)$, and also using $M(\Omega) \le 1$, we obtain

$$\mathbb{E}M(\Omega) \le \min\left\{1, 2\sqrt{\frac{\log(2n)}{m}}\right\}.$$

Since $\mathbb{E}H = m/\lambda$, define

$$b = \min\{m, 2\sqrt{m\log(2n)}\} \quad \text{to obtain} \quad \mathbb{E}M(Z) \le \frac{b}{\lambda}.$$

Choosing the privacy split. Combining the preceding bounds and $\mathbb{E}|\eta_j| = 1/\varepsilon_s$ yields

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le n\left(\frac{4b}{\varepsilon_y} + \frac{1}{\varepsilon_s}\right).$$

Choose

$$\varepsilon_y = \frac{2\sqrt{b}}{2\sqrt{b}+1}\varepsilon, \qquad \varepsilon_s = \frac{1}{2\sqrt{b}+1}\varepsilon.$$

Both are positive and depend only on m, n, ε. Their sum is ε, and the resulting bound is

$$\mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le \frac{n}{\varepsilon}(2\sqrt{b}+1)^2 = O\left(\frac{n\sqrt{m\log(2n)}}{\varepsilon}\right).$$

The final estimate uses $1 \le b \le 2\sqrt{m\log(2n)}$. □

The geometry-dependent split used in the experiments. For consistency with the numerical implementation, one can retain the number N of unordered pairs of distinct public points instead of bounding it by $n(n-1)/2$. When $N \ge 1$, the same argument gives

$$b_Q = \min\{m, \sqrt{2m\log(2N)}\}, \qquad \mathbb{E}[u(\widehat{\sigma})] - \mathrm{OPT} \le \frac{n}{\varepsilon}(2\sqrt{b_Q}+1)^2$$

using $\varepsilon_y : \varepsilon_s = 2\sqrt{b_Q} : 1$. Counting all such pairs, even if their directions coincide, gives a valid upper bound and matches the experimental split. If all public points coincide, every matching has the same true cost; no release is needed.

D Complementary constructions

D.1 Distance-vector bound for two populations

This argument supplies the distance-vector bound in Proposition 6. Suppose the private records have two locations, each repeated $n/2$ times. An optimal assignment partitions the public indices into the two sets assigned to those locations. Every permutation within each of these two sets has the same true cost.

Use the uniform-ball sampler for norm noise (Hardt & Talwar, 2010, Remark 4.2). Independently for each column, write $Z_{ij} = S_j(2V_{ij} - 1)$, where $S_j \sim \mathrm{Gamma}(n+1, \text{rate } \varepsilon)$ and the V_{ij} are independent uniforms on $[0,1]$, independent of S. Adding S_j to column j leaves the assignment unchanged and gives nonnegative perturbations $W_{ij} = 2S_jV_{ij}$. Couple $V_{ij} = 1-e^{-X_{ij}} \le X_{ij}$ with independent rate-one exponential variables X_{ij}. Within each of the two sets, choose an assignment minimizing X, and combine them into τ. It is independent of S and has optimal true cost. Since the private assignment minimizes $C + W$ and $W \ge 0$,

$$u(\widehat{\sigma}) - \mathrm{OPT} \le \sum_i W_{i,\tau(i)}.$$

The exponential assignment identity (Linusson & Wästlund, 2004, Theorem 1.1) therefore gives

$$\mathcal{R}_{\mathrm{dist}} \le \frac{4(n+1)}{\varepsilon} \sum_{\ell=1}^{n/2} \ell^{-2} = O(n/\varepsilon). \tag{14}$$

If private records are displaced from the two locations by v_j, every within-set assignment costs at most $2\sum_j \|v_j\|$ above the new optimum. The same comparison adds this quantity to the bound. This accounts for the private perturbations in the numerical construction; the public points may be arbitrary in this argument.

D.2 Two populations

Let $n = d \ge 4$ be even, $L = n/(2\varepsilon)$, and $\delta = L/n^3$. Set

$$q_i = \begin{cases} -Le_1 + \delta e_i & i \le n/2, \\ Le_1 + \delta e_i & i > n/2, \end{cases} \qquad r_j = \begin{cases} \frac{1}{2}Le_1 + 2Le_2 & j \le n/2, \\ \frac{3}{10}Le_1 & j > n/2. \end{cases}$$

The public row matrix is $\delta I + Lse_1^\top$, where s_i is -1 on the left and 1 on the right. Its determinant is $\delta^n(1 - L/\delta) \ne 0$. The public points are therefore affinely independent, and the projected representation uses $m = n = d$.

First remove the δe_i perturbations. The left-minus-right distance differences, divided by L, for the two private groups are

$$a = \frac{5 - \sqrt{17}}{2} \approx 0.43845, \qquad b_0 = \frac{3}{5}.$$

The optimal assignment sends the first group left. Reversing both groups incurs excess $\frac{n}{2}L(b_0 - a) = c_0 n^2/\varepsilon$, where $c_0 = (5\sqrt{17} - 19)/40 > 0$. This is also the largest excess in the ideal instance. Restoring the public perturbations changes every assignment cost by at most $n\delta$.

Write each projected representation noise vector as $Z_j = H_j\Omega_j$. Gamma and spherical concentration, followed by a union bound over the n individuals, imply

$$\max_j |H_j/L - 2| \longrightarrow 0, \qquad \max_j \max_{\ell=1,2} |\Omega_{j,\ell}| \longrightarrow 0$$

in probability. For example, use deviations $4\sqrt{\log n/n}$ in the Gamma Chernoff bound $\Pr(|\varepsilon H_j/n - 1| > t) \le 2e^{-nt^2/4}$ and the spherical bound $\Pr(|\Omega_{j,\ell}| > t) \le 2e^{-nt^2/2}$. Normalized squared distances converge uniformly to their noiseless values plus 4. The corresponding noisy left-minus-right contrasts converge to

$$\widetilde{a} = \frac{\sqrt{41} - \sqrt{33}}{2} \approx 0.32928, \qquad \widetilde{b}_0 = \frac{\sqrt{569} - \sqrt{449}}{10} \approx 0.26641.$$

Their ordering is reversed. On an event of probability $1 - o(1)$, every contrast in the first group exceeds every contrast in the second by a positive constant times L. If a noisy optimum sent any first-group point left and second-group point right, swapping these two assignments would strictly reduce its objective, even after restoring the public perturbations. Equal group sizes force complete reversal. Consequently

$$\frac{\varepsilon \mathcal{R}_{\mathrm{proj}}}{n^2} \longrightarrow c_0.$$

The upper passage to expectation is justified by the deterministic bound $u_R(\sigma) - \mathrm{OPT} \le c_0 n^2/\varepsilon + 2n\delta$ for every assignment. Equation (14) and Theorem 5 give the other rates in Proposition 6.

Gaussian perturbations. The same argument applies if the ideal public and private points receive perturbations u_i, v_j with maximum norm $o(L)$ and the public affine dimension remains $n-1$. An edge cost, including a noisy-location cost, changes by at most $\|u_i\|+\|v_j\|$. The uniform contrast gap still forces the group reversal. If $D = \sum_i \|u_i\| + \sum_j \|v_j\|$, the excess of any assignment changes by at most $2D$.

For independent $u_i, v_j \sim N(0, s^2\varepsilon^{-2} I_n)$ with fixed $s > 0$, the public affine dimension is $n-1$ almost surely. For $G \sim N(0, I_n)$, $\mathbb{E}\|G\|^4 = n(n+2)$; a union bound and Markov's inequality give, for each $t > 0$,

$$\Pr(\max\{\|u_i\|, \|v_j\|\} > tL) \le \frac{32 s^4 (n+2)}{t^4 n^2} \longrightarrow 0.$$

Also $\mathbb{E}D \le 2sn^{3/2}/\varepsilon = o(n^2/\varepsilon)$. The preceding upper and lower comparisons therefore show the same limit c_0 for joint expectation over geometry and privacy noise. Private-group centers give

$$\mathbb{E}\mathcal{R}_{\mathrm{dist}} \le \frac{2sn^{3/2}}{\varepsilon} + \frac{4(n+1)}{\varepsilon} \sum_{\ell=1}^{n/2} \ell^{-2} = o(n^2/\varepsilon).$$

The corrected projected representation has its uniform subquadratic bound. The positive-perturbation experiment uses these ideal centers before adding Gaussian noise. This is a prescribed population model with growing separation, not a statement about arbitrary Gaussian data.

D.3 Unique partners

Let $h = n/(4\varepsilon)$ and $q_i = r_i = (h/\sqrt{2})e_i$. Costs are zero on the diagonal and h elsewhere. Use the cube-mixture representation from Appendix D.1. For each of $n/2$ disjoint index pairs (i, j), consider

$$V_{ii}, V_{jj} \ge 3/4, \qquad V_{ji}, V_{ij} \le 1/4, \qquad S_i, S_j \ge n/(2\varepsilon).$$

The uniform constraints have probability $1/256$. The second-moment inequality gives $\Pr(S \ge n/(2\varepsilon)) \ge (n+1)/(4(n+2)) \ge 1/6$. Thus the event has probability at least $1/9216$. On it, swapping the two diagonal assignments changes the noisy objective by at most $2h - S_i - S_j < 0$. No noisy optimum keeps both diagonal entries. Summing over disjoint pairs, the expected number of errors is at least $n/18432$, and hence

$$\mathcal{R}_{\mathrm{dist}} \ge \frac{n^2}{73728\varepsilon}. \tag{15}$$

The trivial maximum cost nh supplies the matching upper order.

For the projected representation, $m = n$ and the cost distribution is that of ambient noise. With $v_{ij} = (q_i - q_j)/h$,

$$\|r_j + Z_j - q_i\|^2 - \|r_j + Z_j - q_j\|^2 = h^2 - 2h\langle Z_j, v_{ij}\rangle.$$

Every individual prefers its own distinct task if all projections are less than $h/2$. On $\|Z_j\| \le 2n/\varepsilon$, a failure requires $\langle \Omega_j, v_{ij}\rangle \ge 1/16$. The concentration bounds used above and a union bound yield

$$\mathcal{R}_{\mathrm{proj}} \le \frac{n^2}{4\varepsilon} \min\{1,\ n(2/e)^n + 2n(n-1)e^{-n/512}\} = o(n^2/\varepsilon). \tag{16}$$

The constants make this bound uninformative at the plotted sizes; the experiment is not presented as a validation of an exact-recovery threshold. The corrected projected representation again follows from Theorem 5.

E Public approximation

Let $Q' = (q'_1, \ldots, q'_n)$ be determined from Q alone and define $\Delta_Q = \sum_i \|q_i - q'_i\|$. Keep all n slots, including repeated proxy locations. If a locally private mechanism on (Q', R) has expected excess at most $B(Q', \varepsilon)$, using its assignment on (Q, R) gives

$$\mathcal{R}(Q, R) \le 2\Delta_Q + B(Q', \varepsilon). \tag{17}$$

Indeed, the cost of any permutation and the optimum each change by at most Δ_Q. Privacy is unaffected because the proxy is public and the private input metric is unchanged.

For public PCA of rank r, project Q onto its centered principal subspace to obtain Q'. Apply the projected mechanism for these proxies to the original private points, including the private distance to the proxy subspace. For $r > 0$ the bound is

$$2\Delta_Q + \frac{2n \min\{r+1, d\}}{\varepsilon}.$$

At $r = 0$ all proxies coincide, so the implementation releases nothing and uses a fixed assignment; its bound is $2\Delta_Q$, with no noise term. The numerical dimension sweep evaluates each specified rank separately, without selecting a rank from the private data.

For the two-population construction, replacing each q_i by its unperturbed center gives $\Delta_Q = n\delta$ and representation dimension two. Equation (17) gives $2n\delta + 4n/\varepsilon$. Thus the quadratic failure of the full projected representation is compatible with a strong approximate-location alternative.

F Experimental details and additional checks

All experiments are synthetic. The computation compares reconstructed-cost assignments with an exact optimum of the original Euclidean cost matrix. The solver uses deterministic SciPy/input-order tie-breaking. The proofs allow any fixed public tie rule. We do not measure downstream causal or predictive outcomes.

F.1 Point sets and scale

Uniform cube. Both sides are independent uniform samples from $[0, 1]^d$, using the frozen masters with root seed 20260914 and configuration indices 0, 1, 2. For the paper $n = 20, d = 50$. We retain the archived privacy grid $\{1, 6, 8, 11, 16, 21, \ldots, 201\}$ and 100 noise draws at each point. The master configurations are always reported separately. These intervals quantify conditional simulation noise, not uncertainty over all uniform point sets. The exact random-assignment total cost is

$$\mathrm{RAND} = \frac{1}{n} \sum_{i,j} C_{ij}.$$

No simulated random assignments are needed to compute it.

Two populations and unique partners. Use the constructions in Appendix D with $n = d \in \{32, 64, 128, 256\}$ and $\varepsilon = 1$. For the two populations, start at the ideal centers and add independent $N(0, 0.1^2 I_d)$ perturbations to each point on each side. Use 20 independent geometries per n. Public and private row orders are shuffled independently. The simplex is deterministic up to these label permutations; it is treated as one geometry, not as multiple independent datasets. Every randomized mechanism uses 50 noise draws. The public affine dimension is $n - 1$, and the full projected representation and original-point mechanism endpoints have identical reconstructed-cost distributions. The study explicitly reuses their samples. It also records whether the nonprivate optimum crosses the prescribed two-population labels; those crossings are not automatically errors in optimization.

Exact four-dimensional geometry. Each side consists of $n = 128$ independent $N(0, I_4/4)$ points padded with zeros to $d \in \{8, 32, 128, 512\}$; $\varepsilon = 64$. There is no sample centering or rescaling. For a given replicate, the same active coordinates are reused at every d, so all true costs are exactly the same. Mechanism noise is not paired across d. There are 20 geometries and 50 draws per geometry. The four-dimensional public affine span leads to a five-coordinate report, including the private distance to the span; privacy remains valid for inputs outside the generated subspace.

Approximate four-dimensional geometry. This experiment uses a separate generator with $n = d = 128$ and $\varepsilon \in \{4, 16, 64, 256\}$. Before centering and scaling, public coordinates are independent Gaussians with standard deviation one in the first four coordinates and $a\sqrt{4/(d-4)}$ in the remaining coordinates. Both $a = 0.02$ and $a = 0.2$ appear in the main figure. Thus a specifies a population

ratio of root-mean-square norms in the complementary and first-four-coordinate blocks before sample normalization; it is not the standard deviation of each residual coordinate. Subtract the public sample mean and divide by its root-mean-square norm to give public centered RMS radius exactly one.

Private points are independently generated from $N(0, I_d/d)$. Multiply coordinates 5 through d by 0.5 for the first half and by 4 for the second half. These are prescribed coordinate directions, not the estimated public normal space. They give heterogeneous distances from the main public subspace. There are 20 geometries and 50 noise draws per randomized arm. The plot contains only prespecified measured ranks $\{0, 1, 2, 3, 4, 5, 6, 8, 16, 32, 64, 127\}$ represented in all 20 geometries. Only this prespecified grid is used in the dimension curves.

F.2 Mechanisms, tuning, and uncertainty

All standing mechanisms use the same total per-individual privacy rate. The corrected projected representation uses the geometry-dependent split in Appendix C, counting all nonzero public pairs as an upper bound on the number of distinct directions. Its radius is measured to the public candidate selected from the noisy projected report, and its scalar is privatized. The numerical implementation tests the realization bound (13). Neither an exact radius nor an unprivatized location is used in a private arm.

For each dimension in the reported sweep, use the public principal subspace of that dimension. At rank zero there is no report; full-rank PCA reuses the full projected representation. The numerical implementation can reflect the single normal coordinate when $m = d$. Isotropy of the noise and the fact that public points lie in the reflecting hyperplane give the same reconstructed-cost law as the identity convention in the paper. Numerical SVD tolerances and residuals are recorded in the data. The Gaussian constructions plotted here avoid the tiny deterministic perturbations that could otherwise be truncated by numerical rank estimation.

Noiseless proxy matching is a nonprivate diagnostic: it uses exact private points with the public proxy to quantify distortion alone. It has one deterministic result per geometry. At dimension 127 its value is zero in every geometry, and this endpoint is retained on the linear vertical axes.

For randomized point sets, first average the 50 noise draws within each geometry and divide the resulting additional total distance by that geometry's $\mathrm{RAND} - \mathrm{OPT}$. The plotted mean is the average of these 20 ratios and the interval is $\overline{x} \pm t_{0.975,19} s_x/\sqrt{20}$. For the frozen cube and the simplex, use the analogous interval over noise draws after dividing by their fixed $\mathrm{RAND} - \mathrm{OPT}$, with 99 and 49 degrees of freedom. All intervals are pointwise; no simultaneous or paired superiority tests are asserted. Pilot and production draws are never pooled. Common points or equal sample indices are not treated as evidence of paired mechanism noise.

F.3 Additional privacy rates

F.4 Additional uniform configurations

These use the same sample sizes and privacy grid as Figure 2 (right), with two independently generated pairs of public and private point sets. Each configuration is held fixed across privacy rates and is reported separately, with 100 noise draws per mechanism and rate.

Provenance. The supplied figure manifest identifies the production fingerprints, source arms, instance identifiers, and reused endpoints. Figure CSVs contain means and unmodified interval endpoints, including exact zeros. Selected canonical source rows and experiment metadata accompany the figures. New plots were generated from existing production runs. Separate reruns from the original seeds check reproduction; their samples are not added to the paper averages.

G Composition and implementation details

The following standard composition properties support the local implementations and privacy proofs in the main text (Dwork & Roth, 2014).

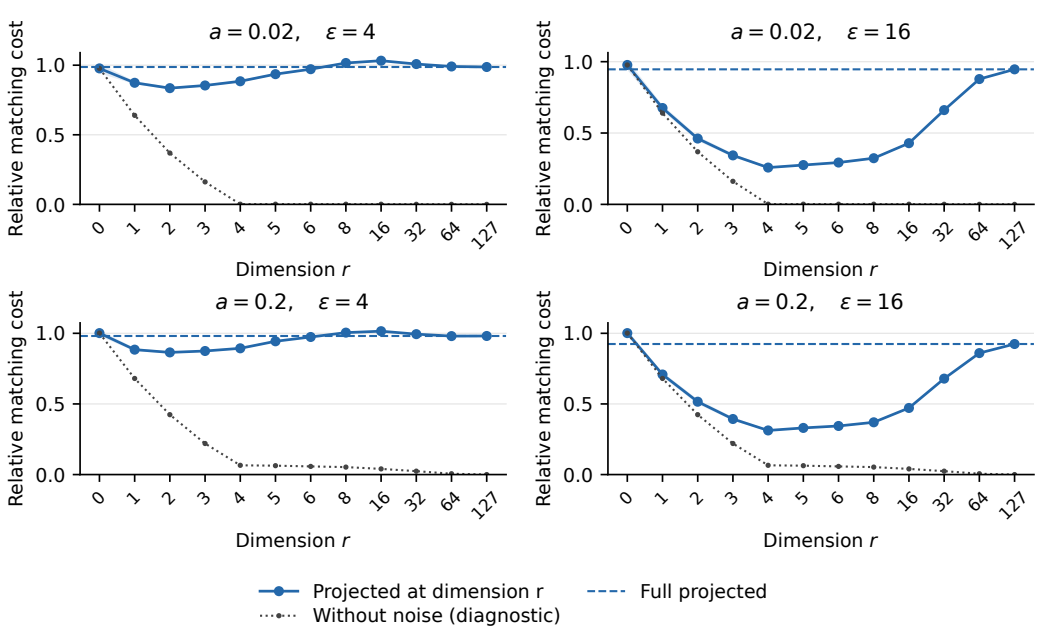

Figure 5: Relative matching cost versus retained dimension at privacy rates 4 and 16. Lower is better. Stronger privacy makes the cost of retaining many coordinates more pronounced. The dotted curve is a nonprivate diagnostic.

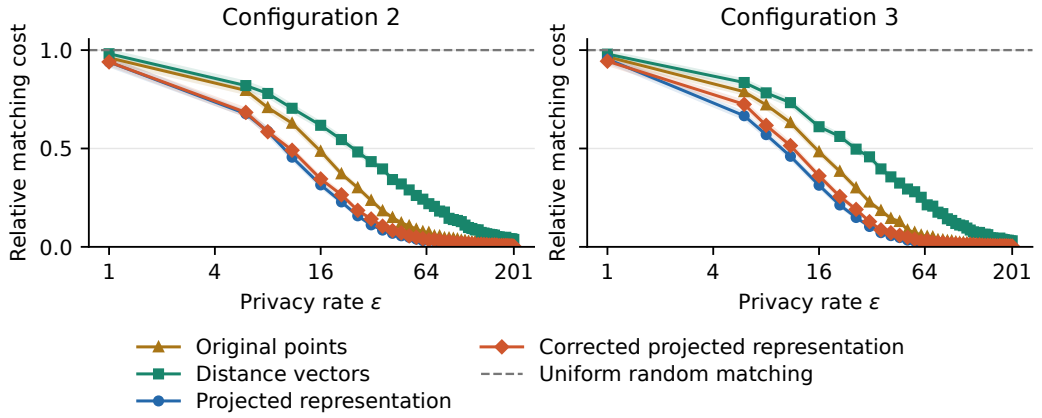

Figure 6: Two further uniform configurations. Lower relative matching cost is better. Both reproduce the mechanism ordering in Figure 2 (right); the dashed line is random matching.

Proposition 8 (Parallel composition)**.** *If each $K_{j,Q}$ satisfies the local ε-metric privacy guarantee and the releases are independent, then*

$$K_Q(R) = \big(K_{1,Q}(r_1), \ldots, K_{n,Q}(r_n)\big)$$

satisfies εD-metric privacy.

Proposition 9 (Post-processing)**.** *If M_Q satisfies εD-metric privacy, then any randomized transformation of its output and Q, without further access to R, also satisfies εD-metric privacy.*

Proposition 10 (Sequential composition)**.** *Suppose $M_{1,Q}(R)$ satisfies $\varepsilon_1 D$-metric privacy and, for every fixed y, $M_{2,Q}(R; y)$ satisfies $\varepsilon_2 D$-metric privacy. Then*

$$\big(M_{1,Q}(R),\ M_{2,Q}(R; M_{1,Q}(R))\big)$$

satisfies $(\varepsilon_1 + \varepsilon_2)D$-metric privacy.

G.1 Pseudocode for the first three mechanisms

The report formulas in the main text give the following implementations. All ties use a fixed public rule.

Algorithm 2 Matching with perturbed original points

Require: Public points Q, private points R, privacy parameter $\varepsilon > 0$

for each individual j, independently **do**
  Draw $Z_j \in \mathbb{R}^d$ with density $p(z) \propto e^{-\varepsilon\|z\|_2}$.
  Release $\widetilde{r}_j = r_j + Z_j$ to the server.
end for
Compute $\widehat{C}_{ij} = \|q_i - \widetilde{r}_j\|_2$ for all i, j.
Find $\widehat{\sigma} \in \arg\min_{\sigma \in S_n} \sum_{i=1}^n \widehat{C}_{i,\sigma(i)}$.
return $\widehat{\sigma}$.

Algorithm 3 Matching with perturbed distance vectors

Require: Public points Q, private points R, privacy parameter $\varepsilon > 0$, norm parameter $p \in [1, \infty]$

for each individual j, independently **do**
  Compute $c_j = \Phi_Q(r_j)$.
  Draw $Z_j \in \mathbb{R}^n$ with density $f_p(z) \propto \exp\big(-\frac{\varepsilon}{n^{1/p}}\|z\|_p\big)$.
  Release $\widetilde{c}_j = c_j + Z_j$ to the server.
end for
Set $\widehat{C}_{ij} = [\widetilde{c}_j]_i$ for all i, j.
Find $\widehat{\sigma} \in \arg\min_{\sigma \in S_n} \sum_{i=1}^n \widehat{C}_{i,\sigma(i)}$.
return $\widehat{\sigma}$.

Algorithm 4 Matching with a perturbed projected representation

Require: Public points Q with $q_1 = 0$, private points R, privacy parameter $\varepsilon > 0$

Construct f_Q from Q and set $p_i = f_Q(q_i)$ for all i.
for each individual j, independently **do**
  Draw $Z_j \in \mathbb{R}^m$ with density $p(z) \propto e^{-\varepsilon\|z\|_2}$.
  Release $Y_j = f_Q(r_j) + Z_j$ to the server.
end for
Compute $\widehat{C}_{ij} = \|p_i - Y_j\|_2$ for all i, j.
Find $\widehat{\sigma} \in \arg\min_{\sigma \in S_n} \sum_{i=1}^n \widehat{C}_{i,\sigma(i)}$.
return $\widehat{\sigma}$.

H Related Work

H.1 Metric differential privacy

Metric differential privacy bounds distinguishability between inputs according to their distance, protecting precise values while allowing disclosure of coarser information. Chatzikokolakis et al. (2013) propose the general framework, while Andrés et al. (2013) develop geo-indistinguishability for location-based systems. For a broader account of the development of metric privacy, its mechanism families, and its applications, we refer readers to the systematization of knowledge by Xie et al. (2026).

Subsequent work investigates how mechanisms optimize the trade-off between privacy and utility. Bordenabe et al. (2014) design location-release mechanisms subject to geo-indistinguishability, and Imola et al. (2022) address the computational challenges of constructing useful metric-private mechanisms at scale. More generally, geometric approaches relate privacy–utility trade-offs to the geometry of the released statistics and the input metric (Hardt & Talwar, 2010; Boedihardjo et al.,

2024). The choice of input metric also determines what information is protected: for example, Imola et al. (2024) use the Earth-Mover's distance to formulate user-level metric privacy for collections of records. This use of transport as an input privacy metric is distinct from minimizing a transport or matching objective. Alvim et al. (2018) study the utility of local metric privacy; Koufogiannis et al. (2015) analyze additive mechanisms for numerical reconstruction under Lipschitz privacy. The latter criterion differs from the assignment loss studied here.

H.2 Private optimization and matching

Private matching belongs to the broader study of optimization with sensitive problem data. Hsu et al. (2014b) distinguish classes of private linear programs according to whether sensitive information appears in the objective or constraints, providing algorithms and impossibility results for these different formulations. Recent work further studies private linear-program feasibility and the trade-off between constraint satisfaction and optimization accuracy (Ene et al., 2025; Benvenuti et al., 2026). These distinctions are directly relevant to assignment: its linear-program formulation has a fixed feasible region when the matching constraints are public, while private locations determine the objective coefficients. The geometric dependence between these coefficients supplies additional structure beyond a general private objective vector.

Another line of work considers optimization in which different agents receive different parts of the solution. Hsu et al. (2014a) study private matchings and allocations under joint differential privacy, and Hsu et al. (2016) extend this perspective to a broad family of convex programs with shared resource constraints. Joint privacy protects an agent's information against the outputs received by other agents, while allowing that agent's own allocation to depend on its input. Danassis et al. (2022) develop decentralized resource allocation under a piecewise local privacy formulation. Resource-allocation privacy also arises through observable contention for shared resources, as studied by Chen et al. (2023). Together, these works highlight the importance of specifying both the sensitive information and the observer's access to the solution.

Matching-specific results address several further privacy models. Dinitz et al. (2025) study matching under edge and node privacy, distinguishing explicitly released matchings from implicit representations. For minimum-weight perfect matching, Sealfon (2016, Appendix B.2) gives algorithms and lower bounds when the topology is public and edge weights are private under an ℓ_1 neighboring relation. Imola et al. (2026) strengthen reconstruction-based lower bounds in this setting. These results provide direct optimization precedents, although arbitrary changes to edge weights differ from the coordinated changes to matching costs induced by moving a private point. Anandan & Clifton (2017) combine secure bipartite-matching computation with differential privacy for the numerical matching cost, illustrating the distinction between protecting an optimum value and protecting a released assignment.

Private optimal transport is closely related because, for two equally weighted point sets of the same size, unregularized discrete transport admits an optimal permutation coupling; its objective equals the corresponding minimum matching cost divided by the number of points. LeTien et al. (2019) use private projected representations and reconstructed transport costs for domain adaptation. Wang & Goldfeld (2024) develop a private Sinkhorn procedure for entropically regularized transport, together with a further randomized transport-cost release. These methods connect private geometric representations to constrained allocation, while their privacy models and, in the latter case, regularized objective differ from pure local metric-private matching with unsquared Euclidean costs.

H.3 Private geometric representations and distance queries

Private representations can preserve information needed for a downstream task without accurately reconstructing every original coordinate. Kenthapadi et al. (2013) combine Johnson–Lindenstrauss projection with Gaussian perturbation and correct the resulting bias in squared-distance estimates. Huang & Roth (2014) exploit metric structure for private distance-query release. More recently, Backurs et al. (2024) develop private data structures for aggregate similarities and distances to a dataset. Aggregate query answers and point-specific assignment costs serve different purposes, but these works establish a common connection between geometric structure, private representations, and downstream utility. Private projected representations have also been used for Euclidean min-

imum spanning trees (Zou et al., 2026), providing an additional example of geometry informing private optimization.

Public reference points offer a further source of structure. Terminal dimensionality reduction preserves distances from designated terminals to other points. The exact projection-plus-orthogonal-norm representation used in this work appears in Narayanan & Nelson (2019, Section 1.1). Both dimensionality reduction and reconstruction of distances from private sketches therefore have established precedents. The mechanism studied here builds on this perspective by releasing a private distance to a selected public candidate and controlling the resulting assignment error. Its analysis uses the invariance of assignment to offsets shared across a cost column, focusing on errors that can change the optimizer rather than requiring uniformly accurate reconstruction of all coordinates.

The analysis of distance reports also uses classical random assignment. Linusson & Wästlund (2004) prove the exact expected optimum for a square matrix of independent exponential costs. We apply that identity inside sets of interchangeable assignments, using the uniform-ball sampler of Hardt & Talwar (2010) to handle dependence within private noise columns.

H.4 Location-private assignment and spatial crowdsourcing

Spatial crowdsourcing provides an extensive application literature on assignment with sensitive locations. Early work uses private spatial decompositions to support task dissemination while protecting worker populations (To et al., 2014). Prorok & Kumar (2017) study location-private vehicle assignment for mobility-on-demand systems. Later approaches address online task assignment through tree-based representations (Tao et al., 2020), personalized privacy through quadtree generalization (Lv et al., 2022), and graph-based online allocation after location perturbation (Wang et al., 2022). Li et al. (2021) consider batch-based task assignment with an untrusted server, while Du et al. (2023) incorporate changing privacy choices and compensation into task assignment. Zhang et al. (2023) propose CANOE, which groups workers' preferred tasks and collects distances from their true locations to perturbed group representatives, providing a closely related approach to supporting assignment through obfuscated distance information.

These formulations differ in arrival assumptions, capacities, disclosure, and utility objectives. For example, Xu et al. (2026) study online ridesharing with known-distribution arrivals and driver matching budgets, considering matching count, profit, and pickup distance. The privacy metric also matters: the mechanism of Lv et al. (2022) uses differences between numbered quadtree nodes at a selected level, rather than continuous Euclidean displacement. This literature motivates private assignment and provides application-specific mechanisms. The present formulation isolates the optimization problem of releasing a minimum-total-distance perfect matching from locally privatized points, with public assignment constraints and an additive utility guarantee that does not require bounded or distributionally specified private locations.