Data Collaboration Analysis with Orthonormal Basis Selection and Alignment

Keiyu Nosaka, Yuichi Takano, Akiko Yoshise

***Abstract*—Data Collaboration (DC) enables multiple parties to jointly train a model without exposing their private datasets. Each party privately transforms its data using a secret linear basis and shares only the resulting intermediate representations. Existing theory asserts that any target basis spanning the same subspace as the secret bases should suffice; however, empirical evidence reveals that the particular choice of target basis significantly influences model accuracy and stability. In this paper, we introduce Orthonormal Data Collaboration (ODC), a novel DC framework that explicitly enforces orthonormality constraints on both the secret and target bases. Under these constraints, the basis alignment step reduces precisely to the classical *Orthogonal Procrustes Problem*, admitting a closed-form solution. We rigorously establish that the resulting orthonormal change-of-basis matrices achieve *orthogonal concordance*, aligning all parties' intermediate representations up to a common orthogonal transformation. Consequently, downstream model performance becomes invariant to the specific choice of orthonormal target basis. Computationally, ODC substantially reduces alignment complexity from $O\big(\min\{a\,(cl)^2,\ a^2\,cl\}\big)$ to $O(a\,cl^2)$, where a denotes anchor data size, l the latent dimension, and c the number of collaborating parties. Extensive empirical evaluations confirm the theoretical advantages of ODC, demonstrating alignment speed-ups of up to two orders of magnitude compared to state-of-the-art DC methods, alongside comparable or superior accuracy across multiple benchmark datasets. ODC maintains robust privacy under the semi-honest threat model and requires only a single round of communication. These results establish ODC as a practically advantageous and computationally efficient enhancement to existing DC pipelines, particularly when orthonormal secret bases are naturally feasible.**

***Index Terms*—Data Collaboration Analysis, Orthogonal Procrustes Problem, Privacy-Preserving Machine Learning, Data Privacy**

I. Introduction

The effectiveness of machine learning (ML) algorithms strongly depends on the quality, diversity, and comprehensiveness of training datasets. High-quality datasets enhance predictive performance and improve the generalization capabilities of ML models across diverse real-world applications. To overcome biases and limitations inherent to datasets obtained from single sources, data aggregation from multiple origins has become common practice. Nonetheless, this practice introduces significant ethical and privacy concerns, particularly regarding unauthorized access and disclosure of sensitive user information. Recent literature highlights a growing awareness and reports an increasing number of privacy breaches linked to large-scale personal data collection and analysis [1].

To address these privacy concerns, Privacy-Preserving Machine Learning (PPML) has emerged as a crucial approach, facilitating the secure and effective utilization of sensitive data—such as medical records, financial transactions, and geolocation traces—without compromising individual privacy. Among various PPML methodologies, *Federated Learning* (FL) [2] has gained prominence for enabling collaborative model training among decentralized data sources, safeguarding data privacy by limiting direct data exchange.

While FL offers promising solutions, it encounters notable challenges in *cross-silo* scenarios, primarily stemming from its dependence on iterative communication among participating entities during training. This communication overhead constitutes a critical obstacle, especially in privacy-sensitive sectors such as healthcare and finance, where institutions typically operate under stringent regulatory constraints and within isolated network infrastructures. Furthermore, FL inherently lacks formal privacy guarantees, necessitating supplementary privacy-enhancing mechanisms to achieve explicit privacy assurances.

A widely adopted strategy to mitigate these shortcomings is employing *Differential Privacy* (DP) [3], [4], typically through additive noise perturbation mechanisms. Nevertheless, approaches based on DP frequently involve significant trade-offs between privacy preservation and model utility [5], thereby constraining their practicality and effectiveness in real-world scenarios.

To overcome limitations associated with conventional PPML methods, *Data Collaboration* (DC) has emerged as a promising alternative [6], [7]. Unlike FL, DC leverages the central aggregation of *secure intermediate representations* computed locally from raw data. Specifically, each participating entity independently transforms its dataset using a privately selected basis matrix. Subsequently, a central aggregator aligns these transformed datasets within a common representation space by constructing a shared target basis—*without knowledge of the secret bases*. This approach facilitates collaborative model training while ensuring robust privacy preservation, thus eliminating the requirement for iterative communication among semi-honest parties [8].

Recent advancements in DC have expanded its applicability to scenarios involving stronger adversarial threats. Notably, methods have been proposed to counter re-identification

K. Nosaka is with the Graduate School of Science and Technology, University of Tsukuba, Ibaraki, Japan (email:s2430118@u.tsukuba.ac.jp).

Y. Takano and A. Yoshise are with the Institute of Systems and Information Engineering, University of Tsukuba, Ibaraki, Japan, and the Center for Artificial Intelligence Research, Tsukuba Institute for Advanced Research (TIAR), University of Tsukuba, Tsukuba, Japan. (email: {ytakano, yoshise}@sk.tsukuba.ac.jp).

This work was partially supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant Numbers JP22K18866, JP23K26327, and JP25KJ0701, as well as the Japan Science and Technology Agency (JST) under Grant Number JPMJSP2124.

Earlier versions of this article have been circulated under the titles "Data Collaboration Analysis Over Matrix Manifolds" and "Data Collaboration Analysis with Orthogonal Basis Alignment."

attacks targeting intermediate representations [9], ensuring that transformed datasets remain unlinkable to their original sources. Additionally, integrating differential privacy mechanisms has been explored to further mitigate the risks associated with malicious collusion among participants [10]. Moreover, hybrid approaches such as FedDCL [11] demonstrate how DC's advantages can be effectively combined with FL paradigms, resulting in scalable, privacy-preserving collaborative ML solutions.

Although DC has empirically demonstrated significant potential in balancing privacy and utility without iterative communication, its theoretical foundations remain underdeveloped. Existing theoretical analyses typically assume that *any* target basis spanning the same subspace as the secret bases is sufficient. However, recent empirical studies indicate that the selection of the target basis substantially influences the performance of downstream models [12], [13]. Specifically, choosing target bases that disproportionately emphasize certain directions in the feature space can degrade model accuracy and utility. These findings highlight a clear discrepancy between current theoretical guarantees and observed empirical behavior, emphasizing the need for improved basis selection and alignment strategies within the DC framework.

A. Our Contributions

We propose a novel framework termed *Orthonormal Data Collaboration (ODC)* to bridge the gap between the theoretical foundations and empirical performance of Data Collaboration (DC). The central innovation of ODC is the explicit enforcement of *orthonormality* constraints on both the secret and target bases during basis selection and alignment. This design leverages common practices in DC, as conventional dimensionality-reduction methods such as Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) naturally produce orthonormal bases, thus imposing minimal additional overhead.

The orthonormality constraint leads to two significant theoretical advantages:

1) **Alignment Efficiency:** The basis alignment simplifies precisely to the classical *Orthogonal Procrustes Problem* [14], for which a closed-form analytical solution is available. This simplification significantly reduces the computational complexity relative to existing DC approaches.
2) **Orthogonal Concordance:** All orthonormal target bases spanning the same subspace as the secret bases result in identical downstream model performance. Thus, the specific choice of an orthonormal target basis becomes inconsequential, effectively resolving previous instabilities in DC.

Empirical evaluations confirm these theoretical results by: (a) demonstrating that ODC achieves substantially faster alignment compared to state-of-the-art DC methods, validating its theoretical complexity advantages; (b) highlighting the practical benefits of orthogonal concordance in stabilizing and enhancing model accuracy; and (c) assessing the robustness of ODC under realistic conditions where theoretical assumptions may not strictly hold.

Notably, the ODC framework extends traditional DC by adding only a single practical assumption—the orthonormality of secret bases. Empirical results demonstrate that relaxing this assumption notably degrades performance, underscoring its necessity. Because orthonormal bases already emerge naturally in existing DC workflows (e.g., [9], DP integration [10], and FL integration [11]), ODC can seamlessly integrate into current pipelines. Moreover, our empirical comparisons position ODC within the broader context of PPML, highlighting its advantages over mainstream techniques such as DP-based perturbations and federated learning.

Fig. 1 provides a conceptual illustration of ODC, highlighting the principle of orthogonal concordance achieved via basis alignment.

B. Organization

The remainder of this paper is organized as follows. § II provides a comprehensive overview of the state-of-the-art DC framework, emphasizing its algorithmic procedures, privacy guarantees, and communication overhead. § III presents an in-depth review of related literature focused on the basis alignment phase of DC and formulates the central research question addressed in this study. § IV introduces the primary theoretical contributions of this paper. Specifically, it first defines the notion of *orthogonal concordance* as a clear objective aimed at resolving key inconsistencies observed in contemporary DC. It then proposes the ODC framework, explicitly designed to achieve orthogonal concordance. § V reports empirical evaluations comparing ODC with existing DC methods, analyzing computational efficiency, the practical superiority of orthogonal concordance, and the influence of theoretical assumptions under realistic settings. Additionally, we contextualize ODC within the broader landscape of PPML by comparing it against mainstream PPML techniques (DP and FL), acknowledging that assumptions and use cases may differ. Finally, § VI summarizes the key findings and outlines future research directions.

II. Preliminaries

In this section, we present the necessary preliminaries on DC analysis. Specifically, we begin with an overview of the DC algorithm, followed by an examination of its privacy-preserving mechanisms and communication overhead..

A. The Data Collaboration Algorithm

We consider a general DC framework for supervised machine learning [6], [10]. Let $\boldsymbol{X} \in \mathbb{R}^{n\times m}$ represent a dataset containing n training samples, each characterized by m features, and let $\boldsymbol{L} \in \mathbb{R}^{n\times l}$ denote the corresponding label set with l labels. For privacy-preserving analysis across multiple entities, we assume the dataset is horizontally partitioned among c distinct entities, expressed as:

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_c \end{bmatrix}, \quad \boldsymbol{L} = \begin{bmatrix} \boldsymbol{L}_1 \\ \vdots \\ \boldsymbol{L}_c \end{bmatrix},$$

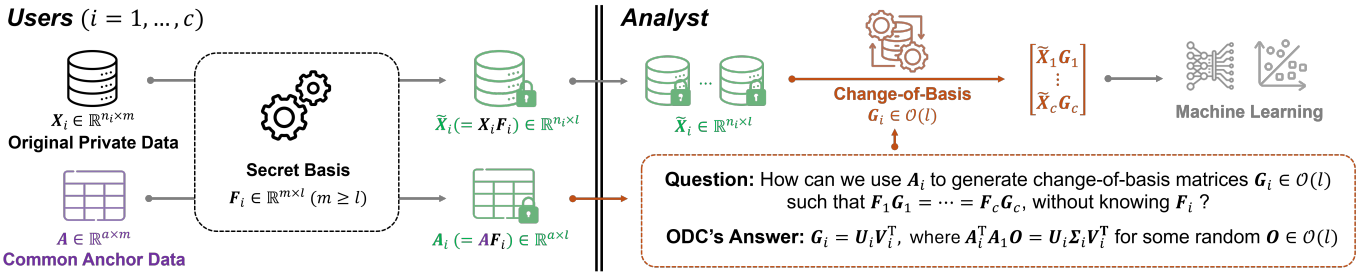


Fig. 1: Conceptual illustration of the Orthonormal Data Collaboration (ODC) framework. Each participating user independently projects their private dataset $\boldsymbol{X}_i \in \mathbb{R}^{n_i \times m}$ and a common anchor dataset $\boldsymbol{A} \in \mathbb{R}^{a \times m}$ into intermediate representations $\tilde{\boldsymbol{X}}_i = \boldsymbol{X}_i \boldsymbol{F}_i$ and $\boldsymbol{A}_i = \boldsymbol{A}\boldsymbol{F}_i$, respectively, using a privately selected orthonormal secret basis $\boldsymbol{F}_i \in \mathbb{R}^{m \times l}$. To collaboratively train machine learning models without revealing their private raw data, an analyst constructs orthogonal change-of-basis matrices $\boldsymbol{G}_i \in \mathcal{O}(l) := \{\boldsymbol{O} \in \mathbb{R}^{l \times l} : \boldsymbol{O}^\top \boldsymbol{O} = \boldsymbol{O}\boldsymbol{O}^\top = \boldsymbol{I}\}$ to align these representations onto a shared orthonormal target basis, without directly accessing the private secret bases. The ODC framework ensures these matrices achieve *orthogonal concordance*, aligning all user representations up to a common orthogonal transformation. Consequently, the analyst can safely aggregate and analyze the aligned representations $\tilde{\boldsymbol{X}}_i \boldsymbol{G}_i$ to perform downstream machine learning tasks. ODC explicitly addresses a key practical question: "How can we use only the intermediate anchor representations $\boldsymbol{A}_i$ to generate alignment matrices $\boldsymbol{G}_i$ without explicitly knowing the secret bases $\boldsymbol{F}_i$?" The analytical solution, illustrated in the figure, is $\boldsymbol{G}_i = \boldsymbol{U}_i \boldsymbol{V}_i^\top$, computed via the singular value decomposition $\boldsymbol{A}_i^\top \boldsymbol{A}_1 \boldsymbol{O} = \boldsymbol{U}_i \boldsymbol{\Sigma}_i \boldsymbol{V}_i^\top$, where $\boldsymbol{O} \in \mathcal{O}(l)$ is an arbitrarily selected orthogonal matrix.

where each entity i possesses a subset of the data $\boldsymbol{X}_i \in \mathbb{R}^{n_i \times m}$ and corresponding labels $\boldsymbol{L}_i \in \mathbb{R}^{n_i \times l}$. The total number of samples satisfies $n = \sum_{i \in [c]} n_i$, where $[c] := \{1, 2, \ldots, c\}$. Additionally, each entity holds a test dataset $\boldsymbol{Y}_i \in \mathbb{R}^{s_i \times m}$, for which the goal is to predict the corresponding labels $\boldsymbol{L}_{\boldsymbol{Y}_i} \in \mathbb{R}^{s_i \times l}$. The DC framework can also be extended to handle more complex scenarios, such as partially shared features [15] or data partitioned both horizontally and vertically [7].

The framework defines two primary roles: the *user* and the *analyst*. Users possess their private datasets $\boldsymbol{X}_i$ and corresponding labels $\boldsymbol{L}_i$, and their objective is to enhance local model performance by leveraging insights derived from other users' data without revealing their own. The analyst's role is to facilitate this collaborative process by providing the necessary resources and infrastructure for the machine learning workflow.

At the outset, each user generates a shared anchor dataset, denoted as $\boldsymbol{A} \in \mathbb{R}^{a \times m}$ $(a > m)$. This anchor dataset may consist of publicly available data or synthetically generated dummy data [6], [7], [16]. Notably, the anchor dataset remains hidden from the analyst to preserve privacy.

Each user independently selects an m-dimensional basis with size l, denoted by $\boldsymbol{F}_i \in \mathbb{R}^{m \times l}$ $(m \geq l)$, to linearly transform their private dataset $\boldsymbol{X}_i$ and the anchor dataset $\boldsymbol{A}$ into secure intermediate representations. A common basis selection method employs truncated SVD with a random orthogonal mapping [11], [17]:

$$\boldsymbol{F}_i = \boldsymbol{V}_i \boldsymbol{E}_i, \tag{1}$$

where, $\boldsymbol{E}_i \in \mathcal{O}(l) := \{\boldsymbol{O} \in \mathbb{R}^{l \times l} : \boldsymbol{O}^\top \boldsymbol{O} = \boldsymbol{O}\boldsymbol{O}^\top = \boldsymbol{I}\}$ and $\boldsymbol{V}_i \in \mathbb{R}^{m \times l}$ denotes the top l right singular vectors of $\boldsymbol{X}_i$. Notably, this typical method inherently produces orthonormal bases, i.e., $\boldsymbol{F}_i^\top \boldsymbol{F}_i = \boldsymbol{I}$.

Once the secret bases $\boldsymbol{F}_i$ are chosen for each user, the secure intermediate representations of the private dataset $\boldsymbol{X}_i$ and the anchor dataset $\boldsymbol{A}$ are computed as follows.

$$\tilde{\boldsymbol{X}}_i = \boldsymbol{X}_i \boldsymbol{F}_i, \quad \boldsymbol{A}_i = \boldsymbol{A}\boldsymbol{F}_i. \tag{2}$$

Each user shares $\tilde{\boldsymbol{X}}_i$, $\boldsymbol{L}_i$, and $\boldsymbol{A}_i$ with the analyst, whose task is to construct a collaborative ML model based on all $\tilde{\boldsymbol{X}}_i$ and $\boldsymbol{L}_i$. However, directly concatenating $\tilde{\boldsymbol{X}}_i$ and building a model from it is futile, as the bases were selected privately and are generally different. Within the DC framework, the analyst aims to align the secret bases using change-of-basis matrices $\boldsymbol{G}_i \in \mathbb{R}^{l \times l}$, and constructs $\hat{\boldsymbol{X}}$ as follows:

$$\hat{\boldsymbol{X}} = \begin{bmatrix} \tilde{\boldsymbol{X}}_1 \boldsymbol{G}_1 \\ \vdots \\ \tilde{\boldsymbol{X}}_c \boldsymbol{G}_c \end{bmatrix}.$$

After successfully creating the change-of-basis matrices from the aggregated $\boldsymbol{A}_i$, as detailed in § III and IV, the analyst utilizes $\hat{\boldsymbol{X}}$ and $\boldsymbol{L}$ to construct a supervised classification model h:

$$\boldsymbol{L} \approx h(\hat{\boldsymbol{X}}).$$

This model h can simply be distributed to the users along with $\boldsymbol{G}_i$ to predict the labels $\boldsymbol{L}_{\boldsymbol{Y}_i}$ of the test dataset $\boldsymbol{Y}_i$:

$$\boldsymbol{L}_{\boldsymbol{Y}_i} = h(\boldsymbol{Y}_i \boldsymbol{F}_i \boldsymbol{G}_i),$$

or employed in other DC-based applications [9], [18], [19], [20], [21], [22], [23] for enhanced privacy or utility. An overview of the DC algorithm is presented in **Algorithm** 1. Notably, cross-entity communication within this framework is strictly limited to Steps 2, 5, and 15 of **Algorithm** 1. The primary challenge of the DC algorithm lies in Step 10, which poses the key question: *How can we generate the change-of-basis matrices $\boldsymbol{G}_i$ without access to the secret bases $\boldsymbol{F}_i$?* This question is addressed in detail in § III and § IV through the development of the proposed ODC framework.

Algorithm 1: Overview of the DC algorithm *(Adapted from Algorithm 1 in [9])*

Input : $\boldsymbol{X}_i \in \mathbb{R}^{n_i \times m}$, $\boldsymbol{L}_i \in \mathbb{R}^{n_i \times l}$, $\boldsymbol{Y}_i \in \mathbb{R}^{s_i \times m}$ for each user $i \in [c]$

Output: $\boldsymbol{L}_{\boldsymbol{Y}_i} \in \mathbb{R}^{s_i \times l}$ for each user $i \in [c]$

User-side ($i \in [c]$): **begin**
  Generate $\boldsymbol{A} \in \mathbb{R}^{a \times m}$ and share it with all users
  Select a secret basis $\boldsymbol{F}_i \in \mathbb{R}^{m \times l}$
  Compute $\tilde{\boldsymbol{X}}_i = \boldsymbol{X}_i \boldsymbol{F}_i$ and $\boldsymbol{A}_i = \boldsymbol{A}\boldsymbol{F}_i$
  Share $\tilde{\boldsymbol{X}}_i$, $\boldsymbol{A}_i$, and $\boldsymbol{L}_i$ with the analyst
end

Analyst-side: **begin**
  for *each user* $i \in [c]$ **do**
    Obtain $\tilde{\boldsymbol{X}}_i$, $\boldsymbol{A}_i$, and $\boldsymbol{L}_i$
    Generate a change-of-basis matrix $\boldsymbol{G}_i \in \mathbb{R}^{l \times l}$
    Compute $\hat{\boldsymbol{X}}_i = \tilde{\boldsymbol{X}}_i \boldsymbol{G}_i$
  end
  Set $\hat{\boldsymbol{X}}$ and $\boldsymbol{L}$ by aggregating all $\hat{\boldsymbol{X}}_i$ and $\boldsymbol{L}_i$
  Analyze $\hat{\boldsymbol{X}}$ to obtain h such that $\boldsymbol{L} \approx h(\hat{\boldsymbol{X}})$
  Return $\boldsymbol{G}_i$ and h to each user $i \in [c]$
end

User-side ($i \in [c]$): **begin**
  Obtain $\boldsymbol{G}_i$ and h
  Predict $\boldsymbol{L}_{\boldsymbol{Y}_i} = h(\boldsymbol{Y}_i \boldsymbol{F}_i \boldsymbol{G}_i)$
end

B. Privacy Analysis

This section briefly examines the privacy considerations and limitations inherent in the DC framework [8]. For clarity and analytical convenience, we assume the threat model of *semi-honest* participants, where all involved entities faithfully follow prescribed procedures but might attempt to extract sensitive information by exploiting any vulnerabilities. We consider a threat model in which either the semi-honest user i or the semi-honest analyst acts as an adversary attempting to infer the private dataset $\boldsymbol{X}_j$ of some other user $j(\neq i)$.

Theorem II.1. ***Privacy Against Semi-Honest Users*** (Adapted from Theorem 1 in [8]) *Any semi-honest user i in the DC framework cannot infer the private dataset $\boldsymbol{X}_j$ of any other user $j(\neq i)$.*

Proof. A semi-honest user i has access only to their own private dataset, the shared anchor dataset $\boldsymbol{A}$, and the collaboratively trained ML model h. Since the model h is constructed using the transformed datasets $\hat{\boldsymbol{X}}$, the most that user i can observe regarding another user $j(\neq i)$ is the transformed representation $\hat{\boldsymbol{X}}_j$.

The anchor dataset $\boldsymbol{A}$ is either publicly available or synthetically generated and does not contain any private information from $\boldsymbol{X}_j$. Therefore, $\boldsymbol{A}$ cannot be used to infer $\boldsymbol{X}_j$.

Furthermore, the transformed data $\hat{\boldsymbol{X}}_j$ is computed as $\hat{\boldsymbol{X}}_j = \boldsymbol{X}_j \boldsymbol{F}_j \boldsymbol{G}_j$, where both transformation matrices $\boldsymbol{F}_j$ and $\boldsymbol{G}_j$ are inaccessible to user i. Without knowledge of $\boldsymbol{F}_j$ and $\boldsymbol{G}_j$, it is infeasible for user i to recover any meaningful information about $\boldsymbol{X}_j$ from $\hat{\boldsymbol{X}}_j$.

Hence, a semi-honest user cannot infer the private dataset $\boldsymbol{X}_j$ of any other user $j(\neq i)$. □

Theorem II.2. ***Privacy Against a Semi-Honest Analyst*** (Adapted from Theorem 2 in [8]) *A semi-honest analyst in the DC framework cannot infer the private dataset $\boldsymbol{X}_j$ of any user j.*

Proof. The semi-honest analyst has access only to the outputs of the linear transformations, specifically $\tilde{\boldsymbol{X}}_j = \boldsymbol{X}_j \boldsymbol{F}_j$ and $\boldsymbol{A}_j = \boldsymbol{A}\boldsymbol{F}_j$, for each user j. However, both the transformation matrix $\boldsymbol{F}_j$ and the anchor dataset $\boldsymbol{A}$ are unknown to the analyst.

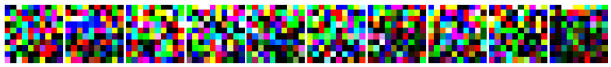
Since $\boldsymbol{F}_j$ and $\boldsymbol{A}$ are inaccessible to the analyst, the pair $(\tilde{\boldsymbol{X}}_j, \boldsymbol{A}_j)$ reveals no meaningful information about the private dataset $\boldsymbol{X}_j$. The analyst lacks sufficient information to invert the transformation or isolate $\boldsymbol{X}_j$.

Therefore, a semi-honest analyst cannot infer any user's private dataset $\boldsymbol{X}_j$. □

Contemporary DC techniques typically use orthonormal bases to transform the private data. Intuitively, imposing orthonormality on the secret bases introduces additional structure that preserves the geometric properties of the data (such as distances and angles) but does not facilitate reconstruction or inference without explicit knowledge of the bases. Therefore, orthonormal bases inherently preserve existing privacy guarantees under the semi-honest adversary model.

To visually reinforce this intuition, we provide a concise illustrative example (Fig. 2). We demonstrate that orthonormal projections significantly degrade visual recognizability, highlighting that orthonormality does not negatively impact privacy relative to general projections. Detailed accuracy and utility analyses involving visual obfuscation are provided later in § V-C.

(a) Original images



(b) Orthogonally projected images ($l = 300$)

(c) Randomly projected images ($l = 300$)

Fig. 2: Visual privacy verification using CelebA [24]. Original images (panel a) compared to images after orthonormal (panel b) and random non-orthogonal projections (panel c). Both transformations clearly obfuscate the visual content, demonstrating that the orthonormality assumption does not compromise visual privacy compared to general projections.

A natural extension of DC's threat model considers the possibility of *malicious collusion* between participants. Specifically, when a user i colludes with the analyst, they collectively

gain access to another user j's transformed data, namely $\boldsymbol{A}_j$ and $\tilde{\boldsymbol{X}}_j$. In this setting, both the input $\boldsymbol{A}$ and the output $\boldsymbol{A}_j$ of user j's secret transformation matrix $\boldsymbol{F}_j$ become observable to the colluding parties.

Given that the anchor matrix $\boldsymbol{A}$ is selected to have full column rank, the primary privacy threat arises from the potential reconstruction of $\boldsymbol{F}_j$ using the Moore-Penrose pseudoinverse: $\boldsymbol{F}_j = \boldsymbol{A}^\dagger \boldsymbol{A}_j$. According to [8], since $\boldsymbol{F}_j \in \mathbb{R}^{m\times l}$ with $m > l$, the matrix $\boldsymbol{F}_j$ acts as a projection. Thus, selecting a smaller value of l makes it substantially harder to reconstruct the original dataset $\boldsymbol{X}_j$ from the projected data $\tilde{\boldsymbol{X}}_j = \boldsymbol{X}_j \boldsymbol{F}_j$, consistent with the privacy guarantees of ε-DR privacy [25].

To further reduce this risk and ensure compliance with stronger privacy standards, differential privacy techniques have been incorporated into the DC framework, as proposed in [10], offering enhanced protection under this collusion-based threat model.

Notably, the privacy guarantees of the DC framework under assumptions of arbitrary participant behavior, where users or analysts may deviate from the prescribed protocol, remain an open research question. Additional privacy threats, such as the potential identifiability of individuals within the data, have been explored in prior work [9], to which we refer readers for further insight.

Importantly, this study neither introduces additional privacy compromises nor seeks to strengthen the existing privacy guarantees of the DC framework. Rather, the focus is on addressing challenges related to the *concordance* of DC, as detailed in §IV.

Our methodology is designed to be compatible with enhanced variants of DC that offer stronger privacy protections, such as those proposed in [9], [10]. Nonetheless, a formal analysis and empirical evaluation of these extended applications are left for future work.

C. Communication Overhead

In this subsection, we analyze the communication overhead incurred within the DC framework. As illustrated in Algorithm 1, DC requires exactly three communication steps:

1) An initial broadcast of a common anchor dataset among all users (Step 2).
2) A single uplink from each user to the analyst, containing transformed representations (Step 5).
3) A final downlink from the analyst back to each user, comprising the trained model and the orthogonal change-of-basis matrix (Step 15).

To quantify the associated overhead precisely, we adopt the following notation: let q represent the number of bits per scalar element, and let N denote the parameter count of the downstream machine learning model h.

For each user $i \in [c]$, the *uplink* data comprises the transformed representations of both its private dataset and the anchor dataset:

$$B_i^{\mathrm{DC}\uparrow} = \frac{(n_i + a) l\, q}{8} \quad \text{[bytes]}. \tag{3}$$

The corresponding *downlink* data transmitted from the analyst to each user includes the orthogonal alignment matrix $G_i \in O(l)$ and the trained model h:

$$B_i^{\mathrm{DC}\downarrow} = \frac{(l^2 + N)\, q}{8} \quad \text{[bytes]}. \tag{4}$$

Additionally, all parties receive the anchor dataset $\boldsymbol{A} \in \mathbb{R}^{a\times m}$ exactly once:

$$B^{\mathrm{anchor}} = \frac{a\, m\, q}{8} \quad \text{[bytes]}. \tag{5}$$

Summing the contributions in (3), (4), and (5), the aggregate communication overhead for DC becomes:

$$\begin{aligned} B^{\mathrm{DC}} &= \sum_{i=1}^{c} \left(B_i^{\mathrm{DC}\uparrow} + B_i^{\mathrm{DC}\downarrow} \right) + B^{\mathrm{anchor}} \\ &= \frac{q}{8} \left[(\bar{n} + a) lc + (l^2 + N)c + am \right], \end{aligned} \tag{6}$$

where $\bar{n} = \frac{1}{c}\sum_{i=1}^{c} n_i$ denotes the average number of samples per user.

To contextualize the efficiency of DC within the broader landscape of PPML, we compare it against FL. Let R denote the number of FL communication rounds, and p the fraction of total users c selected at each round (assumed constant across rounds). Each selected FL participant uploads and downloads the full model once per round, thus incurring twice the model size per round. Consequently, the cumulative communication cost for FL is:

$$B^{\mathrm{FL}} = 2\, R\, p\, c\, \frac{N\, q}{8} \quad \text{[bytes]}. \tag{7}$$

We find the conditions under which DC achieves lower communication overhead by setting $B^{\mathrm{DC}} \le B^{\mathrm{FL}}$:

$$(\bar{n} + a) lc + (l^2 + N)c + am \le 2RpcN.$$

Solving explicitly for R yields:

$$R^\star = \frac{(\bar{n} + a) l}{2pN} + \frac{l^2 + N}{2pN} + \frac{am}{2pNc}, \tag{8}$$

which represents the minimum number of FL rounds required to match DC's cumulative communication overhead.

We identify two practical regimes of particular interest:

- For scenarios where $N \gg l^2$ and $N \gg (\bar{n} + a) l$—typical of deep convolutional neural networks (CNNs) or transformer models—(8) simplifies significantly to:

  $$R^\star \approx \frac{1}{2p} \left[1 + \frac{am}{Nc} \right].$$

  Notably, even with full participation ($p = 1$), FL surpasses DC's communication load within a single communication round unless the number of users c is extremely small.
- Conversely, when the model is small such that $N \ll l^2$—for example, tiny-ML linear models—and user count c is large, we obtain:

  $$R^\star \approx \frac{(\bar{n} + a) l}{2pN}.$$

Here, even a modest latent dimension l ensures $R^\star \gg 10$, implying DC remains advantageous even after several FL rounds.

To illustrate numerically, consider the following realistic setup: $c = 100$ hospitals, each with $n_i = 10^3$ samples; anchor size $a = 10^3$; input dimension $m = 784$ (e.g., flattened 28×28 images); latent dimension $l = 100$; ResNet-50 model with $N \simeq 2.5 \times 10^7$ parameters; 32-bit quantization ($q = 32$). Substituting into (6) yields: $B^{\mathrm{DC}} \approx 10.1$ GB, while one FL round with full participation incurs: $B^{\mathrm{FL}} \approx 20$ GB.

Thus, DC achieves nearly 50% communication savings even at the first FL round ($R = 1$). Under partial participation ($p = 0.1$), DC retains an advantage until approximately $R^\star \approx 5$ FL rounds, according to (8). This numerical example highlights DC's practical efficiency in realistic scenarios.

III. Research Question and Related Works

Focusing on Step 10 of **Algorithm** 1, we address the central research question:

How can the change-of-basis matrices $\boldsymbol{G}_i$ be constructed without access to the secret bases $\boldsymbol{F}_i$?

Effective change-of-basis matrices $\boldsymbol{G}_i$ must satisfy two key properties:

- **Efficiency** — their construction must be computationally scalable to large datasets and many participants.
- **Concordance** — the matrices should perfectly align the local secret bases so as to ensure both stable and high-quality downstream model performance.

This section examines the basis alignment methodologies employed in existing DC frameworks, especially emphasizing their inherent limitations in achieving concordance.

A. Imakura's Basis Alignment (Imakura-DC)

Here, we review the basis alignment method proposed by Imakura *et al.* [8]. The central task for the analyst is to construct suitable change-of-basis matrices $\boldsymbol{G}_i$ that effectively *align* the secret bases $\boldsymbol{F}_i$, even though these secret bases are unknown to the analyst.

Formally, the goal is to identify invertible matrices $\boldsymbol{G}_i \in \mathbb{R}^{l\times l}$ for each $i \in [c]$ that satisfy:

$$\boldsymbol{F}_1\boldsymbol{G}_1 \;=\; \boldsymbol{F}_2\boldsymbol{G}_2 \;=\; \ldots \;=\; \boldsymbol{F}_c\boldsymbol{G}_c, \tag{9}$$

where each basis matrix $\boldsymbol{F}_i \in \mathbb{R}^{m\times l}$ remains inaccessible to the analyst.

Since the analyst also has access to the intermediate representations of a common anchor dataset, defined as $\boldsymbol{A}_i = \boldsymbol{A}\boldsymbol{F}_i \in \mathbb{R}^{a\times l}$, where $\boldsymbol{A}$ remains consistent across all users, it follows that any set of matrices $\boldsymbol{G}_i$ satisfying (9) necessarily satisfy the following condition:

$$\boldsymbol{A}_1\boldsymbol{G}_1 = \boldsymbol{A}_2\boldsymbol{G}_2 = \cdots = \boldsymbol{A}_c\boldsymbol{G}_c. \tag{10}$$

Consequently, the matrices $\boldsymbol{G}_i$ that satisfy equation (10) are necessarily those that minimize the following optimization problem:

$$\begin{aligned} \min_{\boldsymbol{Z}\in\mathbb{R}^{a\times l},\, \boldsymbol{G}_i\in\mathbb{R}^{l\times l}} \quad & \sum_{i=1}^{c} \|\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{Z}\|_{\mathrm{F}}^2 \\ \text{s.t.} \quad & \mathrm{Rank}(\boldsymbol{G}_i) = l. \end{aligned} \tag{11}$$

In Problem (11), the optimization variable $\boldsymbol{Z}$ represents a target matrix. A crucial observation is that the feasible solution space of Problem (11) is inherently non-compact. This non-compactness poses significant difficulties: specifically, one can construct a sequence of invertible matrices whose singular values diminish progressively, causing the sequence to converge toward the zero matrix. Such convergence undermines the stability of the optimization and complicates the identification of meaningful solutions. This ill-posedness can, however, be alleviated by a judicious, *a priori* selection of the target matrix $\boldsymbol{Z}$. Specifically, define $\boldsymbol{Z} = \boldsymbol{U}\boldsymbol{R}$, where $\boldsymbol{U}$ comprises the top l left singular vectors of the concatenated matrix $\begin{bmatrix}\boldsymbol{A}_1 & \cdots & \boldsymbol{A}_c\end{bmatrix}$, and $\boldsymbol{R} \in \mathbb{R}^{l\times l}$ is an arbitrary invertible matrix. Problem (11) reduces to:

$$\min_{\boldsymbol{G}_i\in\mathbb{R}^{l\times l}} \|\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{Z}\|_{\mathrm{F}}^2. \tag{12}$$

The optimization problem in (12) admits closed-form analytical solutions given by $\boldsymbol{G}_i^* = \boldsymbol{A}_i^\dagger\boldsymbol{Z}$ for each $i \in [c]$.

The primary computational cost arises from performing the SVD of the dense concatenated matrix $\tilde{\boldsymbol{A}} \in \mathbb{R}^{a\times lc}$. Consequently, the overall computational complexity, expressed in Big-O notation, is

$$O\left(\min\{a(cl)^2,\, a^2cl\}\right).$$

A natural theoretical approach for evaluating downstream model performance related to the change-of-basis matrices $\boldsymbol{G}_i^*$ involves analyzing their sufficiency concerning equation (9). To this end, we introduce a formal definition for *weak concordance*:

Definition III.1. ***(Weak Concordance)*** *The change-of-basis matrices $\boldsymbol{G}_i \in \mathbb{R}^{l\times l}$, for $i \in [c]$, satisfy weak concordance if*

$$\boldsymbol{F}_1\boldsymbol{G}_1 = \boldsymbol{F}_2\boldsymbol{G}_2 = \cdots = \boldsymbol{F}_c\boldsymbol{G}_c.$$

To rigorously establish weak concordance, we begin by introducing **Assumption** III.2.

Assumption III.2. *We impose the following assumptions on the anchor dataset $\boldsymbol{A} \in \mathbb{R}^{a\times m}$ and the secret bases $\boldsymbol{F}_i \in \mathbb{R}^{m\times l}$ $(m \geq l)$, for all $i \in [c]$:*

1) $\mathrm{Rank}(\boldsymbol{A}) = m$*;*

2) There exists invertible transformation matrices $\boldsymbol{E}_i \in \mathbb{R}^{l\times l}$ such that $\boldsymbol{F}_i = \boldsymbol{F}_1\boldsymbol{E}_i$.

Condition 1) is a common requirement in the DC literature. For instance, selecting $\boldsymbol{A}$ to be a uniformly random matrix (with appropriate dimensions) ensures $\mathrm{Rank}(\boldsymbol{A}) = m$, which is standard practice.

Condition 2) requires all intermediate representations to span an identical l-dimensional subspace. This condition is

both necessary and sufficient for the existence of weakly concordant change-of-basis matrices $\boldsymbol{G}_i \in \mathbb{R}^{l\times l}$. Hence, it plays a crucial role in the theoretical analysis of weak concordance.

The following theorem, established in [8], [11], guarantees weak concordance for the optimal solutions obtained from solving Problem (12):

Theorem III.3. *(Adapted from [8], [11])*
*Suppose that we observe matrices $\boldsymbol{A}_i = \boldsymbol{A}\boldsymbol{F}_i$, $i \in [c]$, with $\boldsymbol{A} \in \mathbb{R}^{a\times m}$ and $\boldsymbol{F}_i \in \mathbb{R}^{m\times l}$. Under **Assumption** III.2, let $\boldsymbol{U}$ denote the matrix formed by the top l left singular vectors of the concatenation $\begin{bmatrix}\boldsymbol{A}_1 & \cdots & \boldsymbol{A}_c\end{bmatrix}$.*

Then, for each $i \in [c]$, the solution $\boldsymbol{G}_i^* = \boldsymbol{A}_i^\dagger \boldsymbol{Z}$ to the optimization problem:

$$\min_{\boldsymbol{G}_i\in\mathbb{R}^{l\times l}} \|\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{Z}\|_{\mathrm{F}}^2,$$

where $\boldsymbol{Z} = \boldsymbol{U}\boldsymbol{R}$ for an arbitrary invertible matrix $\boldsymbol{R} \in \mathbb{R}^{l\times l}$, is weakly concordant.

Proof. See [11]. □

Theorem III.3 establishes that the invertible right factor $\boldsymbol{R} \in \mathbb{R}^{l\times l}$ of the target matrix $\boldsymbol{Z} = \boldsymbol{U}\boldsymbol{R}$ can be chosen arbitrarily while preserving weak concordance. This flexibility naturally prompts the question of whether such arbitrary choices could negatively impact downstream model performance. Unfortunately, the answer is affirmative. Empirical evidence examining this issue is presented in §V-B. Intuitively, selecting a target matrix that disproportionately emphasizes certain directions within the feature space may adversely affect model accuracy and utility. Although recent studies empirically indicate improved performance when choosing $\boldsymbol{R} = \boldsymbol{I}$ [6], this particular choice remains heuristic and lacks rigorous theoretical justification, suggesting potential suboptimality.

Consequently, the practical utility of **Definition** III.1 and **Theorem** III.3 is inherently limited. Indeed, Imakura's basis alignment method exhibits a notable discrepancy between its theoretical guarantees and empirical performance. As a result, concordance is not achieved.

B. Kawakami's Basis Alignment (Kawakami-DC)

To mitigate the arbitrariness associated with the choice of the target matrix $\boldsymbol{Z}$ and its potential negative impact on model accuracy and utility, Kawakami *et al.* [12] proposed an alternative formulation. Their approach involves decomposing each change-of-basis matrix $\boldsymbol{G}_i$ into its constituent column vectors and introducing norm constraints, thereby removing the necessity for an explicit, a priori selection of the target matrix. Specifically, the optimization is formulated as follows:

$$\begin{aligned}\min_{g_{i,k}\in\mathbb{R}^l} \quad & \sum_{i=1}^{c}\sum_{j=1}^{c} \|\boldsymbol{A}_i g_{i,k} - \boldsymbol{A}_j g_{j,k}\|_2^2, \\ \text{s.t.} \quad & \sum_{i=1}^{c} \|\boldsymbol{A}_i g_{i,k}\|_2^2 = 1,\end{aligned} \tag{13}$$

where $g_{i,k}$ denotes the k-th column vector of the matrix $\boldsymbol{G}_i$, defined as:

$$\boldsymbol{G}_i = [g_{i,1} \;\; g_{i,2} \;\; \cdots \;\; g_{i,l}], \quad \forall i \in [c].$$

Define the symmetric block matrix $\boldsymbol{C} \in \mathbb{R}^{cl\times cl}$ and the diagonal block matrix $\boldsymbol{D} \in \mathbb{R}^{cl\times cl}$ as:

$$\boldsymbol{C} := \begin{bmatrix} \boldsymbol{A}_{1,1} & \cdots & \boldsymbol{A}_{1,c} \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_{c,1} & \cdots & \boldsymbol{A}_{c,c} \end{bmatrix},$$

$$\boldsymbol{D} := \begin{bmatrix} \boldsymbol{A}_{1,1} & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & \boldsymbol{A}_{c,c} \end{bmatrix},$$

where $\boldsymbol{A}_{i,j} = \boldsymbol{A}_i^\top \boldsymbol{A}_j$. Additionally, let $\boldsymbol{S} := 2c\boldsymbol{D} - 2\boldsymbol{C}$ and $v_k \in \begin{bmatrix} g_{1k}^\top \cdots g_{ck}^\top \end{bmatrix}^\top \in \mathbb{R}^{cl}$. [12] shows by solving

$$\boldsymbol{S}v_k = \lambda_k \boldsymbol{D}v_k \qquad (v_k^\top \boldsymbol{D}v_k = 1),$$

and computing v_k associated with the smallest l generalized eigenvalues, the solutions of Problem (13) can be efficiently obtained. Furthermore, [12] proposed enhancements to this alignment process. Specifically, they also introduced transformations leveraging QR decomposition and SVD to achieve computational efficiency comparable to that of [8], i.e., $O\left(\min\{a(cl)^2,\, a^2cl\}\right)$. We reference their approach explicitly here, as subsequent experimental validations leverage this improved formulation.

It is important to emphasize that the method proposed by Kawakami *et al.* [12] does not adequately justify the introduction of column-norm constraints. These constraints primarily serve to avoid trivial zero-matrix solutions rather than explicitly guaranteeing the invertibility of the resulting change-of-basis matrices $\boldsymbol{G}_i$. Furthermore, Kawakami *et al.* [12] do not provide any type of analysis regarding the sufficient conditions necessary to ensure any type of concordance.

IV. Orthonormal Basis Selection and Alignment

In our analysis of existing basis alignment methodologies, we identified a significant shortcoming in contemporary approaches concerning the achievement of concordance, highlighting a notable gap between their theoretical foundations and empirical outcomes. Specifically, we demonstrated that the practical relevance of the notion of *weak concordance* (**Definition** III.1) is limited. To address this deficiency, we formally introduce the concept of *orthogonal concordance*, thereby establishing a rigorous mathematical framework that closely aligns with empirical performance.

Definition IV.1. ***(Orthogonal Concordance)***
The orthogonal change-of-basis matrices $\boldsymbol{G}_i \in \mathcal{O}(l)$, for $i \in [c]$, satisify orthogonal concordance if

$$\boldsymbol{F}_1\boldsymbol{G}_1 = \boldsymbol{F}_2\boldsymbol{G}_2 = \cdots = \boldsymbol{F}_c\boldsymbol{G}_c.$$

In contrast to weakly concordant change-of-basis matrices, where the common invertible transformation can be chosen arbitrarily—preserving weak concordance but adversely affecting downstream model performance—orthogonally concordant

change-of-basis matrices constrain the transformation to orthogonal matrices. Since orthogonal transformations inherently preserve distances and angles, any selection of a common orthogonal transformation should theoretically maintain invariance in the performance of distance-based machine learning models. Furthermore, it is well-established that orthogonal transformations exert negligible influence even on models that are not explicitly distance-based. We empirically investigate these theoretical assertions and provide comprehensive insights in §V-B.

Our proposed basis alignment method aims to achieve orthogonal concordance for the resulting change-of-basis matrices. Importantly, orthogonal concordance holds if and only if there exists a common invertible matrix $\boldsymbol{F} \in \mathbb{R}^{m\times l}$ and orthogonal matrices $\boldsymbol{O}_i \in \mathcal{O}(l)$ satisfying:

$$\boldsymbol{F}_i = \boldsymbol{F}\boldsymbol{O}_i. \tag{14}$$

To satisfy this requirement, we first introduce **Assumption** IV.2:

Assumption IV.2. *We impose the following conditions on the anchor dataset $\boldsymbol{A} \in \mathbb{R}^{a\times m}$ and the secret bases $\boldsymbol{F}_i \in \mathbb{R}^{m\times l}$ $(m \geq l)$, for all $i \in [c]$:*

1) $\mathrm{Rank}(\boldsymbol{A}) = m$;
2) *There exist invertible transformation matrices $\boldsymbol{E}_i \in \mathbb{R}^{l\times l}$ such that $\boldsymbol{F}_i = \boldsymbol{F}_1\boldsymbol{E}_i$;*
3) $\boldsymbol{F}_i^\top \boldsymbol{F}_i = \boldsymbol{I}$.

Conditions 1) and 2) coincide with those previously established in **Assumption** III.2 by [8]. We additionally impose condition 3), requiring orthonormality constraints on the secret bases within the ODC framework. It is noteworthy that standard methods such as PCA and SVD naturally yield orthonormal bases. Furthermore, to our knowledge, all existing DC applications can readily accommodate this additional orthonormality constraint.

From conditions 2) and 3), it immediately follows that the transformation matrices $\boldsymbol{E}_i$ are orthogonal, because:

$$\boldsymbol{F}_i^\top \boldsymbol{F}_i = (\boldsymbol{F}_1\boldsymbol{E}_i)^\top(\boldsymbol{F}_1\boldsymbol{E}_i) = \boldsymbol{E}_i^\top \boldsymbol{F}_1^\top \boldsymbol{F}_1\boldsymbol{E}_i = \boldsymbol{E}_i^\top \boldsymbol{E}_i = \boldsymbol{I}_l.$$

This demonstrates that the if-and-only-if requirement (14) for orthogonal concordance is satisfied.

Since our objective is to derive change-of-basis matrices that satisfy orthogonal concordance, we naturally arrive at the following optimization problem:

$$\min_{\boldsymbol{Z}\in\mathbb{R}^{a\times l},\, \boldsymbol{G}_i\in\mathcal{O}(l)} \quad \sum_{i=1}^{c} \|\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{Z}\|_{\mathrm{F}}^2. \tag{15}$$

Importantly, Problem (15) involves a matrix norm minimization objective, which inherently exhibits invariance under arbitrary common orthogonal transformations from the right. Specifically, let $\boldsymbol{O} \in \mathcal{O}(l)$ be an arbitrary orthogonal matrix. Then, choosing $\boldsymbol{Z}^* = \boldsymbol{A}_1\boldsymbol{O}$ and $\boldsymbol{G}_i^* = \boldsymbol{E}_i^\top\boldsymbol{O}$ yields global minimizers for Problem (15), because:

$$\begin{aligned}\sum_{i=1}^{c} \|\boldsymbol{A}_i\boldsymbol{E}_i^\top\boldsymbol{O} - \boldsymbol{A}_1\boldsymbol{O}\|_{\mathrm{F}}^2 &= \sum_{i=1}^{c} \|\boldsymbol{A}\boldsymbol{F}_i\boldsymbol{E}_i^\top\boldsymbol{O} - \boldsymbol{A}\boldsymbol{F}_1\boldsymbol{O}\|_{\mathrm{F}}^2\\ &= \sum_{i=1}^{c} \|\boldsymbol{A}\boldsymbol{F}_1\boldsymbol{O} - \boldsymbol{A}\boldsymbol{F}_1\boldsymbol{O}\|_{\mathrm{F}}^2\\ &= 0.\end{aligned}$$

Given that the analyst has access to $\boldsymbol{A}_i$ for all $i \in [c]$, we may fix $\boldsymbol{Z} = \boldsymbol{A}_1\boldsymbol{O}$ in Problem (15), leading directly to the classical *Orthogonal Procrustes Problem (OPP)*:

$$\min_{\boldsymbol{G}_i\in\mathcal{O}(l)} \quad \sum_{i=1}^{c} \|\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{A}_1\boldsymbol{O}\|_{\mathrm{F}}^2. \tag{OPP}$$

The analyst's objective is twofold: (**A**) compute the analytical solutions $\boldsymbol{G}_i^*$ to (OPP) [14], and (**B**) establish that these solutions satisfy orthogonal concordance without requiring explicit knowledge of $\boldsymbol{F}_i$. We formally state these results in the following theorem.

Theorem IV.3. *Suppose that we observe matrices $\boldsymbol{A}_i = \boldsymbol{A}\boldsymbol{F}_i$, $i \in [c]$, with $\boldsymbol{A} \in \mathbb{R}^{a\times m}$ and $\boldsymbol{F}_i \in \mathbb{R}^{m\times l}$. Under **Assumption IV.2**, the following assertions hold for every orthogonal matrix $\boldsymbol{O} \in \mathcal{O}(l)$:*

(A) *The solution matrices $\boldsymbol{G}_i^*$ to the Orthogonal Procrustes Problem* (OPP) *have analytical forms given by:*

$$\boldsymbol{G}_i^* = \boldsymbol{U}_i\boldsymbol{V}_i^\top, \tag{16}$$

where:

$$\boldsymbol{A}_i^\top\boldsymbol{A}_1\boldsymbol{O} = \boldsymbol{U}_i\boldsymbol{\Sigma}_i\boldsymbol{V}_i^\top,$$

as established in [14].

(B) *These solutions satisfy:*

$$\boldsymbol{G}_i^* = \boldsymbol{E}_i^\top\boldsymbol{O},$$

thereby guaranteeing orthogonal concordance.

Proof. We prove for all $i \in [c]$. Given (OPP), we can write:

$$\begin{aligned}&\|\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{A}_1\boldsymbol{O}\|_{\mathrm{F}}^2\\ &\quad= \mathrm{tr}\left((\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{A}_1\boldsymbol{O})^\top(\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{A}_1\boldsymbol{O})\right)\\ &\quad= \|\boldsymbol{A}_i\|_{\mathrm{F}}^2 + \|\boldsymbol{A}_1\|_{\mathrm{F}}^2 - 2\,\mathrm{tr}(\boldsymbol{G}_i^\top\boldsymbol{A}_i^\top\boldsymbol{A}_1\boldsymbol{O}),\end{aligned}$$

where $\mathrm{tr}(\cdot)$ denotes the matrix trace. Minimizing $\|\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{A}_1\boldsymbol{O}\|_{\mathrm{F}}^2$ for each $\boldsymbol{G}_i$ individually is equivalent to minimizing $\sum_{i=1}^{N}\|\boldsymbol{A}_i\boldsymbol{G}_i - \boldsymbol{A}_1\boldsymbol{O}\|_{\mathrm{F}}^2$ for all $\boldsymbol{G}_i$. Therefore, solving (OPP) is equivalent to solving:

$$\max_{\boldsymbol{G}_i\in\mathcal{O}(l)} \quad \mathrm{tr}(\boldsymbol{G}_i^\top\boldsymbol{A}_i^\top\boldsymbol{A}_1\boldsymbol{O}), \tag{17}$$

for each $i \in [c]$. Consider the SVD $\boldsymbol{A}_i^\top\boldsymbol{A}_1\boldsymbol{O} = \boldsymbol{U}_i\boldsymbol{\Sigma}_i\boldsymbol{V}_i^\top$. Then

$$\begin{aligned}\mathrm{tr}(\boldsymbol{G}_i^\top\boldsymbol{A}_i^\top\boldsymbol{A}_1\boldsymbol{O}) &= \mathrm{tr}(\boldsymbol{G}_i^\top\boldsymbol{U}_i\boldsymbol{\Sigma}_i\boldsymbol{V}_i^\top)\\ &= \mathrm{tr}(\boldsymbol{V}_i^\top\boldsymbol{G}_i^\top\boldsymbol{U}_i\boldsymbol{\Sigma}_i)\\ &= \mathrm{tr}(\boldsymbol{W}_i\boldsymbol{\Sigma}_i)\\ &= \sum_{s=1}^{l} w_{i,(s,s)}\sigma_{i,(s,s)}\end{aligned} \tag{18}$$

where $\boldsymbol{W}_i = \boldsymbol{V}_i^\top \boldsymbol{G}_i^\top \boldsymbol{U}_i$, and $w_{i,(s,t)}, \sigma_{i,(s,t)}$ denote the (s,t)-th elements of matrices $\boldsymbol{W}_i$ and $\boldsymbol{\Sigma}_i$, respectively. Since $\boldsymbol{W}_i \in \mathcal{O}(l)$, $w_{i,(s,t)} \leq 1$ for all s,t. Thus, the sum in (18) is maximized when $\boldsymbol{W}_i = \boldsymbol{I}$, yielding the solution $\boldsymbol{G}_i^* = \boldsymbol{U}_i \boldsymbol{V}_i^\top$, which proves (A).

From condition 3) of Assumption IV.2, we have:

$$\boldsymbol{A}_1 \boldsymbol{O} = \boldsymbol{A} \boldsymbol{F}_1 \boldsymbol{O} = \boldsymbol{A} \boldsymbol{F}_i \boldsymbol{E}_i^\top \boldsymbol{O} = \boldsymbol{A}_i \boldsymbol{E}_i^\top \boldsymbol{O}.$$

Substitute this into Problem (17), and let $\boldsymbol{A}_i^\top \boldsymbol{A}_i = \boldsymbol{Q}_i \boldsymbol{\Lambda}_i \boldsymbol{Q}_i^\top$ be the eigenvalue decomposition. We have:

$$\begin{aligned}
\operatorname{tr}(\boldsymbol{G}_i^\top \boldsymbol{A}_i^\top \boldsymbol{A}_1 \boldsymbol{O}) &= \operatorname{tr}(\boldsymbol{G}_i^\top \boldsymbol{A}_i^\top \boldsymbol{A}_i \boldsymbol{E}_i^\top \boldsymbol{O}) \\
&= \operatorname{tr}(\boldsymbol{G}_i^\top \boldsymbol{Q}_i \boldsymbol{\Lambda}_i \boldsymbol{Q}_i^\top \boldsymbol{E}_i^\top \boldsymbol{O}) \\
&= \operatorname{tr}(\boldsymbol{Q}_i^\top \boldsymbol{E}_i^\top \boldsymbol{O} \boldsymbol{G}_i^\top \boldsymbol{Q}_i \boldsymbol{\Lambda}_i) \\
&= \operatorname{tr}(\boldsymbol{W}_i' \boldsymbol{\Lambda}_i) \\
&= \sum_{s=1}^{l} w'_{i,(s,s)} \lambda_{i,(s,s)}, \qquad (19)
\end{aligned}$$

where $\boldsymbol{W}_i' = \boldsymbol{Q}_i^\top \boldsymbol{E}_i^\top \boldsymbol{O} \boldsymbol{G}_i^\top \boldsymbol{Q}_i$, and $w'_{i,(s,t)}, \lambda_{i,(s,t)}$ denote the (s,t)-th elements of matrices W_i' and $\boldsymbol{\Lambda}_i$, respectively. Since $W_i' \in \mathcal{O}(l)$, $w'_{i,(s,t)} \leq 1$ for all s,t. Thus, the sum in (19) is maximized when $\boldsymbol{W}_i' = \boldsymbol{I}$, which gives:

$$\begin{aligned}
\boldsymbol{G}_i^* &= \boldsymbol{Q}_i \boldsymbol{Q}_i^\top \boldsymbol{E}_i^\top \boldsymbol{O} \\
\boldsymbol{G}_i^* &= \boldsymbol{E}_i^\top \boldsymbol{O},
\end{aligned}$$

and therefore, we have

$$\boldsymbol{F}_1 \boldsymbol{G}_1^* = \cdots = \boldsymbol{F}_c \boldsymbol{G}_c^*,$$

which proves (B). □

ODC's basis alignment procedure with orthogonal concordance can be summarized as follows:

1) Randomly generate $\boldsymbol{O} \in \mathcal{O}(l)$ and compute the SVD

$$\boldsymbol{A}_i^\top \boldsymbol{A}_1 \boldsymbol{O} = \boldsymbol{U}_i \boldsymbol{\Sigma}_i \boldsymbol{V}_i^\top. \qquad (20)$$

2) Set $\boldsymbol{G}_i^* = \boldsymbol{U}_i \boldsymbol{V}_i^\top$.

The primary computational expense in the ODC framework arises from the matrix multiplication of $\boldsymbol{A}_i^\top \boldsymbol{A}_1 \boldsymbol{O}$ for all $i \in [c]$. Consequently, the overall computational complexity of ODC, expressed in Big-O notation, is:

$$O(acl^2).$$

V. Empirical Evaluations

In this section, we empirically evaluate the proposed ODC framework relative to contemporary DC methods. Table I summarizes the theoretical attributes distinguishing ODC from existing DC approaches. Our empirical analysis systematically addresses each attribute highlighted in the table to substantiate the practical advantages conferred by ODC. Specifically, we begin by quantitatively assessing the computational efficiency of ODC compared to established DC techniques. Subsequently, we analyze the practical implications of concordance, focusing particularly on the significance of orthogonal concordance relative to weak concordance. Finally, we investigate the robustness of ODC by evaluating its performance under realistic conditions wherein certain theoretical assumptions may not strictly hold.

A. Empirical Evaluation on Computational Efficiency

Table I demonstrates that, under the natural dimensional ordering $l \leq m < a$, the computational time complexity of ODC satisfies the following inequality relative to existing DC methods:

$$acl^2 \;<\; \min\big\{\, a(cl)^2,\, a^2 cl \big\}.$$

This indicates that ODC inherently incurs lower computational cost compared to Imakura-DC and Kawakami-DC. To empirically corroborate this theoretical advantage, we measured actual wall-clock execution times under controlled conditions.

All three methods admit closed-form solutions for the change-of-basis matrices $\boldsymbol{G}_i$; therefore, their running times depend exclusively on the dimensions of the intermediate anchor representations $\boldsymbol{A}_i \in \mathbb{R}^{a \times l}$ $(i \in [c])$. Since the original anchor matrix $\boldsymbol{A} \in \mathbb{R}^{a \times m}$ is drawn from a standard i.i.d. uniform distribution independently from the private datasets $\boldsymbol{X}_i \in \mathbb{R}^{n_i \times m}$, the computational cost of alignment is isolated from user-specific data characteristics.

For the empirical evaluation, we generated uniformly random matrices $\boldsymbol{A}_i \in \mathbb{R}^{a \times l}$, $i \in [c]$, and varied each of the three primary parameters (anchor size a, latent dimension l, and the number of users c), while keeping the other two fixed. The specific experimental settings are summarized in Table II. As prior literature [12] suggests, randomized SVD enhances computational efficiency for contemporary DC methods; thus, we employed randomized SVD for both Imakura-DC and Kawakami-DC. Conversely, ODC remained unchanged, as it inherently requires only a single full SVD of an $l \times l$ matrix (see Eq. (20)). Each experiment was repeated 100 times, and we report median runtimes to mitigate transient system effects.

Figure 3 plots the measured median running times. Across all parameter variations, ODC consistently outperforms contemporary DC approaches by at least an order of magnitude, validating the theoretical complexity hierarchy detailed in Table I.

1) Scaling with Latent Dimension l: Fig. 3(a) indicates an approximate power-law scaling of computation time with respect to the dimension l. We applied ordinary least squares (OLS) regression to the linear model:

$$\log_{10}(\text{time}) = \kappa + \alpha \, \log_{10}(l),$$

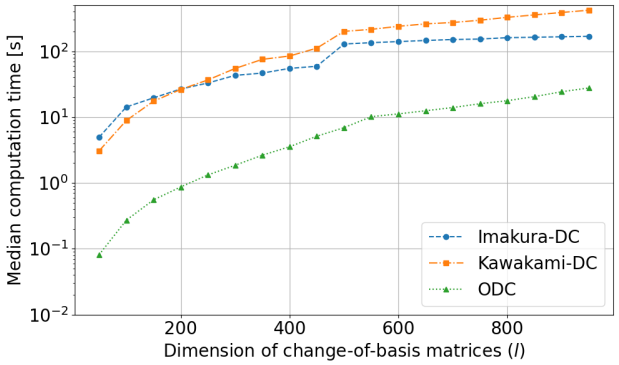
with the constant κ to empirically estimate the exponent α. The resulting empirical estimates $\widehat{\alpha}_l$ are:

$$\widehat{\alpha}_l = \begin{cases} 1.26, & \text{Imakura–DC} \\ 1.75, & \text{Kawakami–DC} \\ 2.05, & \text{ODC.} \end{cases}$$

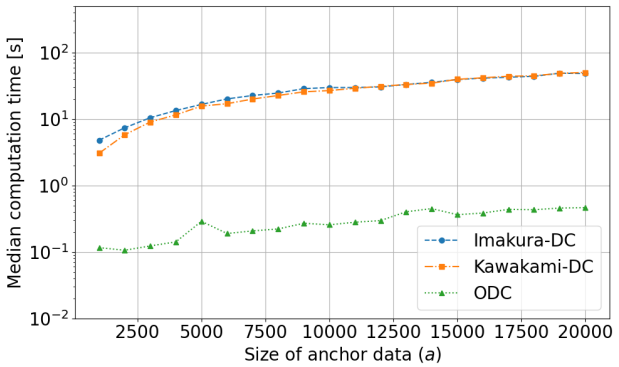
The observed exponent $\alpha \approx 2$ for ODC aligns closely with its theoretically derived complexity $O(acl^2)$. Conversely, contemporary DC methods exhibit somewhat smaller empirical exponents, reflecting complexity consistent with their theoretical predictions of $O\big(\min\{a(cl)^2, a^2 cl\}\big)$. Although ODC has a slightly higher exponent with respect to l, its absolute computational cost remains significantly lower—approximately 6–15 times faster—across the tested range of dimensions. This practical efficiency arises primarily from

TABLE I: Theoretical comparison between contemporary DC methods and the proposed ODC framework.

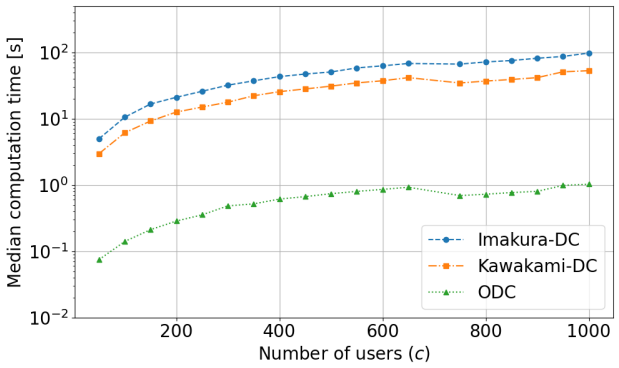
	Imakura-DC	**Kawakami-DC**	**ODC**
Efficiency	$O\left(\min\{a(cl)^2,\, a^2cl\}\right)$	$O\left(\min\{a(cl)^2,\, a^2cl\}\right)$	$O(acl^2)$
Concordance	Weak Concordance	N/A	Orthogonal Concordance
Assumptions	**Assumption** III.2	N/A	**Assumption** IV.2



(a) Wall-clock time versus latent dimension $l\,(a, c) = (1000, 50)$.



(b) Wall-clock time versus anchor size $a\,(l, c) = (50, 50)$.



(c) Wall-clock time versus number of users $c\,(a, l) = (1000, 50)$.

Fig. 3: Wall-clock time with varying parameters (l, a, c). Across all parameter variations, ODC consistently outperforms contemporary DC approaches by at least an order of magnitude, validating the theoretical complexity hierarchy detailed in Table I.

ODC avoiding explicit construction and manipulation of the large, dense $a \times cl$ concatenated matrix, thereby circumventing costly large-scale singular value decompositions. Moreover,

TABLE II: Design of the efficiency experiment. Two parameters are held fixed while the third is swept over the indicated range.

Figure	**Free parameter (fixed pair)**	**Range**
Fig. 3(a)	dimension l $(a, c) = (1000, 50)$	50:50:950
Fig. 3(b)	anchor size a $(l, c) = (50, 50)$	1000:1000:20000
Fig. 3(c)	users c $(a, l) = (1000, 50)$	50:50:1000

typical practical scenarios involve $a \gg l$; thus, the substantial scaling advantage in a (discussed further in § V-A2) renders the modestly larger exponent in l negligible in realistic deployment.

2) Scaling with Anchor Size a: In Fig. 3(b), the relationship between computational wall-clock time and anchor size a again exhibits an approximate power-law scaling. Applying OLS regression to the log-log relationship $\log_{10}(\text{time}) = \kappa + \alpha \log_{10}(a)$ yields the following empirical estimates $\widehat{\alpha}_a$:

$$\widehat{\alpha}_a = \begin{cases} 0.79 & \text{Imakura–DC}, \\ 0.82 & \text{Kawakami–DC}, \\ 0.56 & \text{ODC}. \end{cases}$$

Notably, all empirically observed slopes are lower than their corresponding theoretical predictions ($\alpha = 1$ for ODC and $\alpha \in \{1, 2\}$ for baseline DC methods). The significantly lower slope of 0.56 observed for ODC is particularly noteworthy. This empirical observation can be attributed to ODC's dominant computational workload comprising matrix multiplications—specifically, performing c independent multiplications of relatively small $l \times a$ and $a \times l$ matrices—which are highly optimized and efficiently executed in modern computational environments, thus performing substantially better in practice than theoretically anticipated.

At the maximum tested anchor size $a = 20\,000$, median wall-clock runtimes are $171\,\text{s}$ for Imakura–DC, $165\,\text{s}$ for Kawakami–DC, and $1.6\,\text{s}$ for ODC, corresponding to an empirical speed-up factor exceeding two orders of magnitude ($\sim 100\times$) in favor of the proposed ODC method.

3) Scaling with the Number of Users c: Fig. 3(c) examines computational scaling with respect to the number of users c, holding the dimensions $(a, l) = (1000, 50)$ fixed. Applying OLS regression to the log-log relationship $\log_{10}(\text{time}) = \kappa + \alpha \log_{10}(c)$, we obtain the following empirical estimates $\widehat{\alpha}_c$:

$$\widehat{\alpha}_c = \begin{cases} 0.95 & \text{Imakura–DC}, \\ 0.92 & \text{Kawakami–DC}, \\ 0.84 & \text{ODC}. \end{cases}$$

For ODC, the theoretical complexity is $O(acl^2)$, predicting a slope $\alpha \approx 1$. The slightly lower empirical slope of 0.84 is

primarily due to efficient parallelization and vectorization in modern CPU architectures.

Conversely, the theoretical complexity for Imakura–DC and Kawakami–DC methods is $O(\min\{a(cl)^2,\, a^2cl\})$. Under the tested dimensions, the a^2cl term dominates for $c > 20$, predicting linear scaling in c. The empirical slopes of 0.95 and 0.92 closely align with these theoretical predictions.

Although asymptotic scaling behaviors are similar, ODC exhibits significantly improved absolute performance. Median runtimes at $c = 1000$ report that ODC completes computations in only **1.0 second**, compared to 52 or 96 seconds required by the baseline methods. Thus, ODC achieves empirical speed-ups of at least $50\times$.

This substantial performance gap arises because ODC employs optimized batched matrix multiplications and small-scale ($l \times l$) singular value decompositions, while baseline methods incrementally extend memory-intensive randomized SVD computations with each additional user.

In practice, ODC maintains nearly constant incremental computational overhead per user. This efficiency renders the basis-alignment phase computationally negligible in realistic deployments.

B. Empirical Evaluation on Orthogonal Concordance

Theoretically, orthogonally concordant change-of-basis matrices preserve downstream performance in distance-based models. However, in practice, the presence of finite data, non-distance-based architectures, and numerical precision issues can introduce performance variability. This section empirically evaluates whether orthogonal concordance (ODC) offers a practical advantage over weak concordance (Imakura-DC).

The methods evaluated are:

- **DC-random**: Weakly concordant change-of-basis matrices constructed via **Theorem** III.3, using a uniformly random matrix $\boldsymbol{R}$.
- **DC-identity**: Weakly concordant matrices from **Theorem** III.3, with $\boldsymbol{R} = \boldsymbol{I}$.
- **ODC-random**: Orthogonally concordant matrices from **Theorem** IV.3, using a uniformly random orthogonal matrix $\boldsymbol{O}$ (Haar distribution).
- **ODC-identity**: Orthogonally concordant matrices **Theorem** IV.3 with $\boldsymbol{O} = \boldsymbol{I}$.

We follow the canonical DC protocol with $c = 100$ users, each providing $n_i = 100$ samples. The anchor matrix $\boldsymbol{A} \in \mathbb{R}^{1000\times 784}$ is generated as a uniformly random matrix, and the secret bases $\boldsymbol{F}_i \in \mathbb{R}^{784\times 100}$ satisfy Assumption IV.2. Experiments are conducted on the MNIST [26] and Fashion-MNIST [27] datasets. For each dataset, we train (i) a Support Vector Machine (SVM), representative of distance-based classifiers, and (ii) a single-hidden-layer MLP with 256 ReLU neurons.

Each configuration is repeated with 100 independent random seeds. For every method, we report the mean test accuracy along with the 95% confidence interval. To assess statistical significance, we perform paired, one-sided t-tests at the 1% level, testing the null hypothesis that the *random* variant performs *no worse* than the identity variant. Results are presented in Table III

Three key observations follow:

(a) **DC-random** consistently underperforms **DC-identity** by 4.0 pp on MNIST (SVM), 4.6 pp on Fashion-MNIST (SVM), and approximately 1.3 pp on both MLP tasks. Strongly negative t-statistics and vanishing p-values ($< 10^{-26}$) confirm that arbitrary $\boldsymbol{R}$ significantly harms model accuracy.

(b) **ODC-random** performs comparably to **ODC-identity**, with differences never exceeding 0.38 pp and well within the 95% confidence intervals. In three out of four settings, $t > 0$ and $p \approx 1$, indicating that random $\boldsymbol{O}$ may even slightly outperform the identity. In the remaining case (Fashion-MNIST (MLP)), the difference is statistically insignificant ($p = 0.085$).

(c) Since orthogonal transformations are Euclidean isometries, the performance of distance-based models (e.g., SVMs) should remain invariant. The MLPs, although not explicitly distance-based, exhibit negligible sensitivity. These empirical findings strongly support the practical significance of Orthogonal Concordance.

In summary, the experiments demonstrate that orthogonal concordance effectively eliminates the performance variability observed with weak concordance. Random orthogonal matrices are as safe to use for ODC, whereas arbitrary invertible matrices can introduce substantial degradation in performance for Imakura-DC.

C. Empirical Evaluation Under Relaxed Assumptions

Our theoretical analysis in Section IV relies on two key assumptions regarding the secret bases $\boldsymbol{F}_i$:

1) **Identical Span**: all $\boldsymbol{F}_i$ share an identical column space (Assumption IV.2-2);
2) **Orthonormality**: each $\boldsymbol{F}_i$ has orthonormal columns (Assumption IV.2-3).

In practical scenarios, these ideal conditions may not strictly hold, particularly when bases are generated through computationally inexpensive random projections or derived from heterogeneous local datasets. We thus empirically quantify how deviations from these assumptions impact performance.

1) Experimental Design:

Secret-base conditions: We evaluate four controlled scenarios with varying adherence to these assumptions (see Table IV):

Datasets and metrics: We group tasks into five domains:

- **Image Classification (MNIST, Fashion-MNIST):** Images are flattened into 28×28 vectors and normalized to the interval $[0, 1]$. Performance metrics reported are mean classification accuracies obtained using SVM and MLP models over 100 independent runs.
- **Biomedical Compound Classification (TDC [28]):** Molecular data are represented using 2048-bit Morgan fingerprints (radius = 2). Data are partitioned among four users, with each user exclusively holding samples from a single class (illustrated in Fig. 4). Performance is evaluated via ROC-AUC and PR-AUC scores (MLP), averaged over 100 runs.

TABLE III: Evaluation of the practical significance of orthogonal concordance compared to weak concordance on image classification tasks (MNIST and Fashion-MNIST datasets) using SVM and MLP classifiers. The table reports mean accuracies (%) ± margin of error (95% confidence interval, 100 runs each) for four concordance conditions: **DC-Identity** (weak concordance with identity matrix $\boldsymbol{R} = \boldsymbol{I}$), **DC-Random** (weak concordance via uniformly random matrix $\boldsymbol{R}$), **ODC-Identity** (orthogonal concordance with identity matrix $\boldsymbol{O} = \boldsymbol{I}$), and **ODC-Random** (orthogonal concordance via uniformly random orthogonal matrix $\boldsymbol{O}$, Haar-distributed). Columns labeled **DC-Δ [%]** and **ODC-Δ [%]** show the difference in mean accuracy between the random and identity matrix conditions (random minus identity), where negative values indicate worse performance for random matrices. Statistical significance is indicated by * for $p < 0.01$, based on one-tailed paired t-tests ($n = 100$), testing the null hypothesis that "choosing a uniformly random matrix performs greater than or equal to choosing the identity matrix." $^-$ denotes differences that are not statistically significant ($p \geq 0.01$). The results demonstrate that DC performance significantly deteriorates with random matrices, whereas ODC consistently maintains robust performance irrespective of random orthogonal matrix choice, highlighting the practical effectiveness and stability of orthogonal concordance.

Dataset	Classifier	DC-Identity	DC-Random	DC-Δ [%]	ODC-Identity	ODC-Random	ODC-Δ [%]
MNIST	SVM	94.50 ± 0.00	90.50 ± 0.10	$-4.00^\star$	94.50 ± 0.00	94.52 ± 0.01	$+0.02^-$
MNIST	MLP	94.40 ± 0.00	93.05 ± 0.13	$-1.35^\star$	94.40 ± 0.00	94.78 ± 0.07	$+0.38^-$
Fashion-MNIST	SVM	85.20 ± 0.00	80.64 ± 0.11	$-4.56^\star$	84.40 ± 0.00	84.42 ± 0.01	$+0.02^-$
Fashion-MNIST	MLP	84.40 ± 0.00	83.22 ± 0.16	$-1.18^\star$	84.50 ± 0.00	84.41 ± 0.13	-0.09^-

TABLE IV: Summary of secret-base conditions evaluated.

Condition	Identical Span	Orthonormality
SameSpan-Orth	✓	✓
SameSpan	✓	×
DiffSpan-Orth	×	✓
DiffSpan	×	×

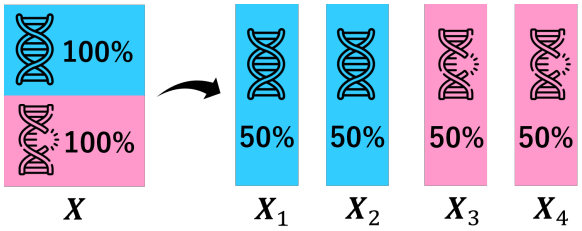


Fig. 4: Illustration of extremely heterogeneous splitting applied to TDC datasets. Binary-labeled data are partitioned across four users, each exclusively holding samples of a single label.

- **Income Classification (Adult [29]):** A widely-used dataset for income prediction tasks, which we specifically partitioned horizontally and *vertically* across 200 users. Half of the users possess only one subset of features, while the remaining users possess the complementary feature subset. We report the mean classification accuracy obtained using an MLP classifier, averaged over 100 runs.
- **Facial Attribute Classification (CelebA [24]):** RGB images (128×128 pixels) are flattened and normalized to the interval $[0, 1]$. Classification performance (binary gender prediction) is measured via the accuracy of an MLP classifier averaged over 100 runs. Results are further benchmarked against (ε, δ)-DP baselines.
- **Clinical Regression (eICU-CRD [30]):** Data from the 50 largest hospitals are used to predict patient length-of-stay (days) based on 25 routinely collected clinical features. Performance is measured using RMSE (lower values indicate better predictive performance), averaged across hospitals over 100 runs with an MLP model. Federated averaging (FedAvg), executed over 40 rounds with full participation, provides a baseline for federated learning.

Compared methods: We compare a centralized oracle (**Central**); per-user local models (**Local**); two contemporary DC baselines (**Imakura-DC** [8], **Kawakami-DC** [12]); and the proposed **ODC**. For CelebA, we additionally include Gaussian **DP**, and for eICU, we include **FedAvg**, to place ODC in the wider PPML context.

2) Results and Discussion: Tables V–VIII comprehensively summarize model performance under four distinct secret-basis scenarios, while Figure 5 provides complementary visual evidence of privacy preservation. The following discussion is organized around three comparative perspectives:

I. ODC *vs.* Existing DC Methods Under Various Secret-Basis Conditions (Tables V and VI)

- **SameSpan-Orth (identical span and orthonormality satisfied):** Under these ideal conditions, ODC achieves performance equal to or marginally better than the centralized oracle across multiple datasets (e.g., MNIST MLP accuracy $95.9\,\%$, matching the oracle; see Table V). Similar performance from Imakura-DC and Kawakami-DC confirms effective utilization of this ideal scenario by all methods.
- **SameSpan (orthonormality violated):** Relaxing orthonormality significantly impacts ODC's performance—MNIST SVM accuracy drops from $95.6\,\%$ to $88.8\,\%$ and AMES ROC-AUC from $88.6\,\%$ to $59.9\,\%$. Imakura-DC and Kawakami-DC remain largely unaffected, underscoring ODC's theoretical dependence on orthonormal secret bases.
- **DiffSpan-Orth (orthonormality maintained, identical span violated):** When orthonormality is enforced, ODC demonstrates robustness against subspace misalignment, maintaining accuracy in the range $94.9\,\%$–$95.3\,\%$ on MNIST, and often surpassing DC baselines in biomedical tasks (e.g., HIV

TABLE V: Performance comparison of ODC and baseline DC methods across multiple datasets (MNIST, Fashion-MNIST, and biomedical TDC datasets) using various classifiers and performance metrics. The table summarizes the mean scores ± margin of error at 95% confidence level over 100 runs. Four distinct secret-base conditions are evaluated: **SameSpan-Orth**: Secret bases share identical subspaces and maintain orthonormality (column spaces transformed individually by uniformly random orthogonal matrices). **SameSpan**: Secret bases share identical subspaces but do not enforce orthonormality (column spaces transformed by uniformly random invertible, non-orthogonal matrices). **DiffSpan-Orth**: Secret bases independently generated with orthonormal columns, without enforcing identical subspaces. **DiffSpan**: Secret bases independently generated without constraints on subspace similarity or orthonormality. Evaluated methods include: a fully centralized (**Central**) oracle, a purely local (**Local**) model (for image datasets only), two existing DC methods (**Imakura-DC** [8], **Kawakami-DC** [12]), and the proposed **ODC** method. Key observations: When assumptions are strictly satisfied (**SameSpan-Orth**), ODC achieves performance comparable or superior to the centralized oracle, validating its theoretical advantages. Relaxing the orthonormality constraint alone (**SameSpan**) notably deteriorates ODC's performance, highlighting orthonormality as crucial. Existing DC baselines are relatively insensitive to orthonormality violations, underscoring the unique dependency of ODC on orthonormal bases. ODC remains competitive when subspace alignment assumptions are relaxed (**DiffSpan-Orth**), outperforming baselines in several tasks (e.g., HIV, CYP3A4), indicating practical robustness driven primarily by orthonormality.

Dataset	Secret Bases	SVM Accuracy [%]					MLP Accuracy [%]				
		Central	**Local**	**Imakura-DC**	**Kawakami-DC**	**ODC**	**Central**	**Local**	**Imakura-DC**	**Kawakami-DC**	**ODC**
MNIST	**SameSpan-Orth**	96.0±0.1	66.0±0.9	96.1±0.1	96.1±0.1	95.6±0.1	95.1±0.1	59.7±1.4	96.0±0.1	95.9±0.1	95.9±0.1
	SameSpan			96.1±0.1	96.1±0.1	88.8±0.2			95.9±0.1	95.8±0.1	92.6±0.2
	DiffSpan-Orth			95.1±0.1	94.7±0.1	94.9±0.1			94.4±0.2	93.2±0.2	95.3±0.1
	DiffSpan			94.9±0.1	94.7±0.1	86.5±0.3			94.9±0.1	93.2±0.2	91.2±0.2
Fashion-MNIST	**SameSpan-Orth**	86.1±0.2	62.9±0.8	85.4±0.2	85.4±0.2	83.7±0.3	86.1±0.2	61.1±1.0	86.1±0.2	86.0±0.2	86.0±0.2
	SameSpan			85.4±0.2	85.4±0.2	77.7±0.3			86.1±0.2	85.9±0.2	80.5±0.3
	DiffSpan-Orth			83.4±0.2	83.0±0.3	82.5±0.2			81.8±0.3	81.5±0.3	84.3±0.2
	DiffSpan			82.9±0.2	83.0±0.3	76.3±0.3			82.5±0.3	81.5±0.3	78.8±0.4

Dataset	Secret Bases	ROC-AUC [%]				PR-AUC [%]			
		Central	**Imakura-DC**	**Kawakami-DC**	**ODC**	**Central**	**Imakura-DC**	**Kawakami-DC**	**ODC**
AMES [31]	**SameSpan-Orth**	87.1±0.0	86.2±0.4	82.4±0.3	88.6±0.1	88.7±0.0	87.9±0.4	83.8±0.4	89.9±0.1
	SameSpan		86.6±0.3	82.2±0.3	59.9±0.4		88.2±0.3	83.7±0.3	62.5±0.4
	DiffSpan-Orth		63.1±0.4	65.4±0.4	67.9±0.3		66.9±0.4	69.1±0.4	71.5±0.3
	DiffSpan		61.9±0.3	65.2±0.4	58.1±0.4		65.6±0.3	69.0±0.4	62.0±0.4
Tox21_SR-ARE [32]	**SameSpan-Orth**	75.0±0.0	69.4±1.0	64.5±1.3	75.4±0.3	42.2±0.0	30.8±1.1	25.2±1.3	38.4±0.6
	SameSpan		69.0±1.0	64.5±1.3	54.0±0.3		30.4±1.1	25.2±1.2	16.7±0.2
	DiffSpan-Orth		57.4±0.3	58.3±0.3	61.0±0.3		19.4±0.2	20.2±0.2	22.6±0.3
	DiffSpan		56.6±0.3	58.5±0.3	53.9±0.3		18.9±0.2	20.3±0.2	17.3±0.2
HIV [33]	**SameSpan-Orth**	78.8±0.0	79.6±0.4	78.1±0.5	80.8±0.2	42.0±0.0	41.6±0.5	39.1±0.7	41.8±0.4
	SameSpan		80.1±0.3	78.0±0.5	58.8±0.4		42.2±0.4	38.9±0.8	5.3±0.1
	DiffSpan-Orth		61.1±0.4	60.9±0.4	67.4±0.3		7.8±0.2	8.7±0.2	13.4±0.3
	DiffSpan		58.1±0.4	60.8±0.4	56.3±0.4		6.6±0.2	8.5±0.3	4.6±0.1
CYP3A4 [34]	**SameSpan-Orth**	87.2±0.0	84.8±0.2	82.9±0.2	86.2±0.1	83.5±0.0	80.6±0.3	78.1±0.3	82.1±0.1
	SameSpan		84.7±0.2	82.9±0.2	59.2±0.4		80.5±0.3	78.1±0.3	49.8±0.5
	DiffSpan-Orth		57.8±0.3	59.4±0.4	65.5±0.3		49.8±0.3	51.5±0.5	57.2±0.3
	DiffSpan		58.3±0.3	59.1±0.4	58.2±0.3		50.4±0.3	51.1±0.5	48.9±0.4
CYP2D6 [34]	**SameSpan-Orth**	83.5±0.0	81.4±0.3	80.4±0.2	83.3±0.1	63.5±0.0	60.2±0.4	58.5±0.3	62.3±0.2
	SameSpan		81.8±0.2	80.3±0.2	59.5±0.4		60.6±0.4	58.3±0.3	24.2±0.3
	DiffSpan-Orth		62.4±0.4	64.6±0.4	65.8±0.2		27.5±0.5	30.3±0.5	31.2±0.3
	DiffSpan		62.6±0.3	64.7±0.4	57.1±0.4		27.9±0.4	30.4±0.5	23.0±0.3
CYP1A2 [34]	**SameSpan-Orth**	91.2±0.0	89.7±0.1	89.5±0.2	90.4±0.2	90.1±0.0	88.8±0.1	88.4±0.3	89.8±0.2
	SameSpan		90.3±0.1	89.2±0.2	65.1±0.5		89.4±0.1	88.1±0.3	59.4±0.4
	DiffSpan-Orth		67.3±0.4	68.6±0.4	67.9±0.3		62.8±0.5	64.9±0.5	63.9±0.3
	DiffSpan		67.8±0.4	68.7±0.4	60.2±0.5		63.3±0.4	64.9±0.5	55.9±0.5

ROC-AUC $67.4\,\%$ compared to Imakura-DC $61.1\,\%$; Table V). Thus, exact span alignment is advantageous but not critical, provided orthonormality holds.

- **DiffSpan (both conditions violated):** Violating both conditions negatively affects all methods, with ODC showing the most severe degradation (e.g., MNIST SVM accuracy reduces to $86.5\,\%$). This observation highlights the practical necessity of orthonormal bases, particularly in heterogeneous, high-variance settings.

Regarding Table VI, we observe performance trends mirror those observed under purely horizontal splits, despite the additional vertical data splitting.

II. ODC *vs.* DP-based Perturbation (Table VII)
Table VII shows that under realistic conditions (DiffSpan–Orth), ODC attains $83.6\,\%$ accuracy, surpassing DP at lower privacy budgets by substantial margins ($+14.5$ pp for $\varepsilon = 0.5$ and $+4.1$ pp for $\varepsilon = 2$), while nearly matching DP at $\varepsilon = 8$ (only 0.2 pp below). Figure 5 further demonstrates that DP at high privacy budgets preserves visually identifiable

TABLE VI: Performance comparison of ODC and baseline DC methods on the Adult dataset using an MLP classifier. The results represent mean accuracy ($\pm$ margin of error at 95% confidence, computed over 100 independent trials) under the four distinct secret-basis conditions described in Table V. Evaluated methods match those in Table V. Data are partitioned horizontally and *vertically* across 200 users, with each user possessing only half of the total feature set. Key observations: Under strict orthogonality assumptions (**SameSpan-Orth**, **DiffSpan-Orth**), ODC achieves performance comparable to the centralized oracle. Relaxing orthonormality (**SameSpan**, **DiffSpan**) significantly impairs ODC, highlighting its critical reliance on orthonormal bases. Existing DC methods are relatively insensitive to orthonormality violations, emphasizing ODC's unique dependence. Despite vertical data splitting, performance trends mirror those observed under purely horizontal splits (Table V).

Secret Bases	Central	Local	Imakura-DC	Kawakami-DC	ODC
SameSpan-Orth	84.9 ± 0.2	74.2 ± 1.0	84.9 ± 0.3	84.8 ± 0.3	85.0 ± 0.2
SameSpan			84.9 ± 0.3	84.8 ± 0.2	83.0 ± 0.3
DiffSpan-Orth			84.5 ± 0.3	84.4 ± 0.3	84.5 ± 0.3
DiffSpan			84.3 ± 0.3	84.4 ± 0.3	82.4 ± 0.3

TABLE VII: Comparison of classification accuracy on the CelebA dataset using an MLP classifier across different PPML approaches. The reported results represent the mean accuracy ($\pm$ margin of error at 95% confidence, computed over 100 independent trials) under the four distinct secret-base conditions described in Table V. Compared methods include: a fully centralized (**Central**) oracle, a purely local (**Local**) model, models trained using **DP**-based additive Gaussian noise at three privacy budget levels ($\varepsilon = 0.5, 2, 8$), two existing DC methods (**Imakura-DC** [8], **Kawakami-DC** [12]), and the proposed **ODC** method. Key observations: Under the orthogonal assumption (**DiffSpan-Orth**), ODC significantly surpasses the lower-privacy DP models ($\varepsilon = 0.5, 2$) and closely approaches the DP model with $\varepsilon = 8$, highlighting its competitive accuracy while providing robust visual privacy. Relaxing orthonormality moderately affects ODC performance, demonstrating sensitivity to theoretical assumptions but still maintaining higher accuracy than the more privacy-restrictive DP conditions.

Secret Bases	Central	Local	DP ($\varepsilon = 0.5$)	DP ($\varepsilon = 2$)	DP ($\varepsilon = 8$)	Imakura-DC	Kawakami-DC	ODC
SameSpan-Orth	88.4±0.2	75.7±0.6	69.1±0.5	79.5±0.3	83.8±0.3	86.5±0.2	86.2±0.2	86.2±0.2
SameSpan						86.3±0.2	85.8±0.3	79.9±0.3
DiffSpan-Orth						82.7±0.2	82.8±0.3	83.6±0.3
DiffSpan						82.6±0.2	82.8±0.3	78.6±0.3

TABLE VIII: Comparison of RMSE on the eICU-CRD dataset using an MLP classifier across different PPML approaches. The reported results represent the mean RMSE ($\pm$ margin of error at 95% confidence, computed over 100 independent trials) under the four distinct secret-base conditions described in Table V. Compared methods include: a fully centralized (**Central**) oracle, a purely local (**Local**) model, a **FedAvg** model with full participation, two existing DC methods (**Imakura-DC** [8], **Kawakami-DC** [12]), and the proposed **ODC** method. Key observations: **FedAvg** performs best, followed by **ODC** with the **DiffSpan-Orth** condition. Interestingly, the identical span assumption does not positively affect the DC-based method's performance. This observation implies that in realistic heterogenous settings, the reference subspace to meet the identical span condition must be carefully chosen to increase performance.

Secret Bases	Central	Local	FedAvg	Imakura-DC	Kawakami-DC	ODC
SameSpan-Orth	3.996±0.046	4.174±0.048	4.064±0.049	4.089±0.049	4.096±0.049	4.083±0.048
SameSpan				4.088±0.049	4.096±0.049	4.154±0.049
DiffSpan-Orth				4.085±0.049	4.090±0.049	4.070±0.048
DiffSpan				4.085±0.049	4.090±0.049	4.150±0.049

features, whereas ODC fully obfuscates visual identity. Thus, ODC provides superior privacy–utility trade-offs without explicitly calibrating privacy budgets.

However, the formal privacy mechanisms underpinning ODC and contemporary DC currently rely on a semi-honest assumption; consequently, these mechanisms remain underdeveloped for scenarios involving stronger or malicious adversaries. Suppose the semi-honest privacy assumption cannot be guaranteed or tolerated. In that case, DP-based additive perturbation may still be preferred, despite the advantageous privacy-performance trade-offs offered by DC methods. This preference arises from the rigorous privacy guarantees inherent in DP.

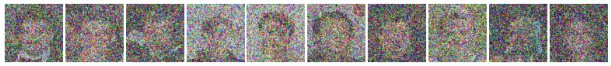
III. ODC *vs.* Federated Learning (Table VIII)

Table VIII shows FedAvg achieving the lowest RMSE (4.060). ODC, under DiffSpan–Orth, is statistically indistinguishable (4.065, 0.005 difference), clearly outperforming Imakura-DC and Kawakami-DC ($\sim$4.086). Given FedAvg's requirement of iterative communication, ODC presents a strong alternative, offering near-FedAvg performance with significantly reduced communication overhead. Interestingly, enforcing identical span conditions negatively affects performance for eICU-CRD, indicating that arbitrary subspace selection is detrimental for realistic heterogeneous tasks.

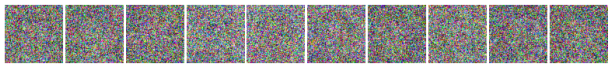
Summary of Empirical Insights.: ODC achieves *optimal* performance under ideal conditions (**SameSpan-Orth**), shows

(a) Original images

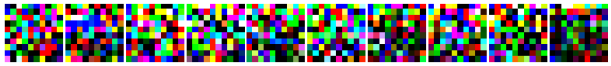
(b) Noise-added images with (ε, δ)-DP, $(\varepsilon, \delta) = (8, 10^{-3})$



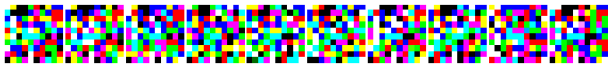
(c) Noise-added images with (ε, δ)-DP, $(\varepsilon, \delta) = (2, 10^{-3})$



(d) Noise-added images with (ε, δ)-DP, $(\varepsilon, \delta) = (0.5, 10^{-3})$



(e) Orthogonally projected images ($l = 300$)



(f) Randomly projected images ($l = 300$)

Fig. 5: Visual privacy analysis using CelebA. Panel (a) depicts original images with dimensions $128 \times 128 \times 3$. Panels (b)–(d) illustrate that, even under additive perturbation with DP, image contents remain visually discernible at higher privacy budgets (e.g., $\varepsilon = 8$). As privacy strength intensifies (i.e., decreasing ε), the visual recognizability diminishes; however, this improvement in visual privacy is accompanied by a substantial deterioration in classification accuracy (Table VII). In sharp contrast, panels (e)–(f), depicting DC projection methods employing orthogonal and random bases, consistently achieve strong visual obfuscation.

sensitivity to orthonormality violations, yet remains *robust* against subspace misalignment provided orthonormality is maintained. Across diverse tasks, ODC consistently matches or exceeds DC baselines and provides privacy–utility trade-offs competitive with DP and FL approaches. Table IX provides a qualitative summary comparing these representative PPML frameworks. The observations discussed above strongly advocate enforcing orthonormal secret bases (via PCA, SVD, or QR) while tolerating moderate span misalignment to effectively handle realistic data heterogeneity.

VI. Conclusion

In this paper, we revisited the Data Collaboration (DC) paradigm and identified a critical theoretical and practical limitation: the downstream model performance can significantly vary with the specific choice of the target basis. To address this fundamental issue, we proposed the **Orthonormal Data Collaboration (ODC)** framework, a principled refinement of DC explicitly enforcing orthonormality constraints on both secret and target bases. Under this orthonormality assumption, we showed that the basis alignment step simplifies elegantly to the classical Orthogonal Procrustes Problem [14], which admits an efficient closed-form analytical solution. Consequently, ODC significantly reduces the computational complexity of the alignment phase from

$$O\big(\min\{a(cl)^2,\, a^2cl\}\big) \quad \rightarrow \quad O(a\,c\,l^2),$$

where a denotes the size of the anchor dataset, l the latent dimension, and c the number of participating entities.

We provided theoretical proof that the resulting orthonormal change-of-basis matrices satisfy *orthogonal concordance*. Specifically, the intermediate representations from all participating parties become aligned up to a common orthogonal transform, thereby preserving distances and inner products. A direct consequence of this property is that downstream analytical tasks become invariant to the specific orthonormal basis selection. Our theoretical findings hold rigorously under the conventional semi-honest threat model and are fully compatible with established privacy-enhancing techniques and extensions of the DC paradigm.

Extensive empirical evaluations strongly corroborate our theoretical insights. On synthetic benchmarks isolating computational cost, ODC achieved alignment speed-ups of one to two orders of magnitude relative to state-of-the-art DC methods. Moreover, practical evaluations conducted across diverse application domains—including standard image classification benchmarks (MNIST and Fashion-MNIST) and various heterogeneous biomedical datasets (e.g., AMES, HIV, CYP450 enzyme prediction)—demonstrated that ODC consistently matches or surpasses the accuracy of contemporary DC methods while preserving its marked computational efficiency advantage.

Importantly, ODC maintains a minimal communication footprint, requiring only a single transfer of intermediate representations from users to the analyst and one broadcast of the (potentially synthetic) anchor dataset. This streamlined design sharply contrasts with the extensive iterative communication overhead characteristic of federated learning approaches. Additionally, given that orthonormal bases naturally arise from widely adopted dimensionality reduction techniques such as PCA, integration of ODC into existing DC pipelines requires negligible additional effort.

Future research should address several promising directions. First, while our work is explicitly grounded in the assumption of semi-honest participants, an in-depth exploration and reinforcement of this assumption's privacy guarantees remains an important open challenge. Additionally, extending the orthonormality-based alignment framework to accommodate non-linear or partially overlapping feature spaces, stronger adversarial threat models, and privacy-preserving integration with differential privacy mechanisms or federated learning constitutes meaningful directions for subsequent theoretical and empirical investigation.

References

[1] P. Rosati, P. Deeney, M. Cummins, L. van der Werff, and T. Lynn, "Social media and stock price reaction to data breach announcements:

TABLE IX: Qualitative comparison of representative PPML frameworks.

Framework	Communication rounds for training	Typical accuracy*	Privacy guarantee
DP (additive perturbation)	One user → analyst	Low	Rigorous (ε, δ)–DP
Federated Learning	Iterative bidirectional rounds until convergence	High	No intrinsic formalism†
ODC (this work)	One user → analyst+ one user-wide anchor broadcast	Medium$	Rigorous under the semi-honest model

* Relative rankings derived empirically; specific accuracies vary by task.
† FL can incorporate DP externally, often reducing accuracy.
$ Performance approaches FedAvg under strict orthonormality.

Evidence from us listed companies," *Research in International Business and Finance*, vol. 47, pp. 458–469, 2019.

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial Intelligence and Statistics*. PMLR, 2017, pp. 1273–1282.

[3] C. Dwork, "Differential privacy: A survey of results," in *International Conference on Theory and Applications of Models of Computation*. Springer, 2008, pp. 1–19.

[4] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. Quek, and H. V. Poor, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE transactions on information forensics and security*, vol. 15, pp. 3454–3469, 2020.

[5] R. Xu, N. Baracaldo, and J. Joshi, "Privacy-preserving machine learning: Methods, challenges and directions," arXiv preprint arXiv:2108.04417, 2021.

[6] A. Imakura and T. Sakurai, "Data collaboration analysis framework using centralization of individual intermediate representations for distributed data sets," *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, vol. 6, no. 2, p. 04020018, 2020.

[7] A. Imakura, X. Ye, and T. Sakurai, "Collaborative data analysis: Non-model sharing-type machine learning for distributed data," in *Knowledge Management and Acquisition for Intelligent Systems: 17th Pacific Rim Knowledge Acquisition Workshop, PKAW 2020, Yokohama, Japan, January 7–8, 2021, Proceedings 17*. Springer, 2021, pp. 14–29.

[8] A. Imakura, A. Bogdanova, T. Yamazoe, K. Omote, and T. Sakurai, "Accuracy and privacy evaluations of collaborative data analysis," *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

[9] A. Imakura, T. Sakurai, Y. Okada, T. Fujii, T. Sakamoto, and H. Abe, "Non-readily identifiable data collaboration analysis for multiple datasets including personal information," *Information Fusion*, vol. 98, p. 101826, 2023.

[10] H. Yamashiro, K. Omote, A. Imakura, and T. Sakurai, "Toward the application of differential privacy to data collaboration," *IEEE Access*, vol. PP, pp. 1–1, 2024.

[11] A. Imakura and T. Sakurai, "Feddcl: a federated data collaboration learning as a hybrid-type privacy-preserving framework based on federated learning and data collaboration," *arXiv preprint arXiv:2409.18356*, 2024.

[12] Y. Kawakami, Y. Takano, and A. Imakura, "New solutions based on the generalized eigenvalue problem for the data collaboration analysis," arXiv preprint arXiv:2404.14164, 2024.

[13] K. Nosaka and A. Yoshise, "Creating collaborative data representations using matrix manifold optimal computation and automated hyperparameter tuning," in *2023 IEEE 3rd International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB)*. IEEE, 2023, pp. 180–185.

[14] P. H. Schönemann, "A generalized solution of the orthogonal procrustes problem," *Psychometrika*, vol. 31, no. 1, pp. 1–10, 1966.

[15] A. Mizoguchi, A. Imakura, and T. Sakurai, "Application of data collaboration analysis to distributed data with misaligned features," *Informatics in Medicine Unlocked*, vol. 32, p. 101013, 2022.

[16] A. Imakura, X. Ye, and T. Sakurai, "Collaborative novelty detection for distributed data by a probabilistic method," in *Proceedings of The 13th Asian Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 157. PMLR, Nov 17–19 2021, pp. 932–947.

[17] A. Mizoguchi, A. Bogdanova, A. Imakura, and T. Sakurai, "Data collaboration analysis applied to compound datasets and the introduction of projection data to non-iid settings," 2023.

[18] T. Nakayama, Y. Kawamata, A. Toyoda, A. Imakura, R. Kagawa, M. Sanuki, R. Tsunoda, K. Yamagata, T. Sakurai, and Y. Okada, "Data collaboration for causal inference from limited medical testing and medication data," 2025. [Online]. Available: https://arxiv.org/abs/2501.06511

[19] Y. Kawamata, R. Motai, Y. Okada, A. Imakura, and T. Sakurai, "Collaborative causal inference on distributed data," *Expert Systems with Applications*, vol. 244, p. 123024, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417423035261

[20] A. Bogdanova, A. Imakura, and T. Sakurai, "Dc-shap method for consistent explainability in privacy-preserving distributed machine learning," *Human-Centric Intelligent Systems*, vol. 3, no. 3, pp. 197–210, 2023.

[21] A. Imakura, R. Tsunoda, R. Kagawa, K. Yamagata, and T. Sakurai, "Dc-cox: Data collaboration cox proportional hazards model for privacy-preserving survival analysis on multiple parties," *Journal of Biomedical Informatics*, vol. 137, p. 104264, 2023.

[22] A. Imakura, H. Inaba, Y. Okada, and T. Sakurai, "Interpretable collaborative data analysis on distributed data," *Expert Systems with Applications*, vol. 177, p. 114891, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417421003328

[23] T. Yanagi, S. Ikeda, N. Sukegawa, and Y. Takano, "Privacy-preserving recommender system using the data collaboration analysis for distributed datasets," *arXiv preprint arXiv:2406.01603*, 2024.

[24] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proceedings of International Conference on Computer Vision (ICCV)*, December 2015.

[25] H. Nguyen, D. Zhuang, P.-Y. Wu, and M. Chang, "Autogan-based dimension reduction for privacy preservation," *Neurocomputing*, vol. 384, pp. 94–103, 2020.

[26] L. Deng, "The mnist database of handwritten digit images for machine learning research [best of the web]," *IEEE signal processing magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[27] H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms," *arXiv preprint arXiv:1708.07747*, 2017.

[28] K. Huang, T. Fu, W. Gao, Y. Zhao, Y. Roohani, J. Leskovec, C. W. Coley, C. Xiao, J. Sun, and M. Zitnik, "Therapeutics data commons: Machine learning datasets and tasks for drug discovery and development," *arXiv preprint arXiv:2102.09548*, 2021.

[29] B. Becker and R. Kohavi, "Adult," UCI Machine Learning Repository, 1996, DOI: https://doi.org/10.24432/C5XW20.

[30] T. J. Pollard, A. E. Johnson, J. D. Raffa, L. A. Celi, R. G. Mark, and O. Badawi, "The eicu collaborative research database, a freely available multi-center database for critical care research," *Scientific data*, vol. 5, no. 1, pp. 1–13, 2018.

[31] C. Xu, F. Cheng, L. Chen, Z. Du, W. Li, G. Liu, P. W. Lee, and Y. Tang, "In silico prediction of chemical ames mutagenicity," *Journal of chemical information and modeling*, vol. 52, no. 11, pp. 2840–2847, 2012.

[32] A. Mayr, G. Klambauer, T. Unterthiner, and S. Hochreiter, "Deeptox: toxicity prediction using deep learning," *Frontiers in Environmental Science*, vol. 3, p. 80, 2016.

[33] Z. Wu, B. Ramsundar, E. N. Feinberg, J. Gomes, C. Geniesse, A. S. Pappu, K. Leswing, and V. Pande, "Moleculenet: a benchmark for molecular machine learning," *Chemical science*, vol. 9, no. 2, pp. 513–530, 2018.

[34] H. Veith, N. Southall, R. Huang, T. James, D. Fayne, N. Artemenko, M. Shen, J. Inglese, C. P. Austin, D. G. Lloyd *et al.*, "Comprehensive characterization of cytochrome p450 isozyme selectivity across chemical libraries," *Nature biotechnology*, vol. 27, no. 11, pp. 1050–1055, 2009.