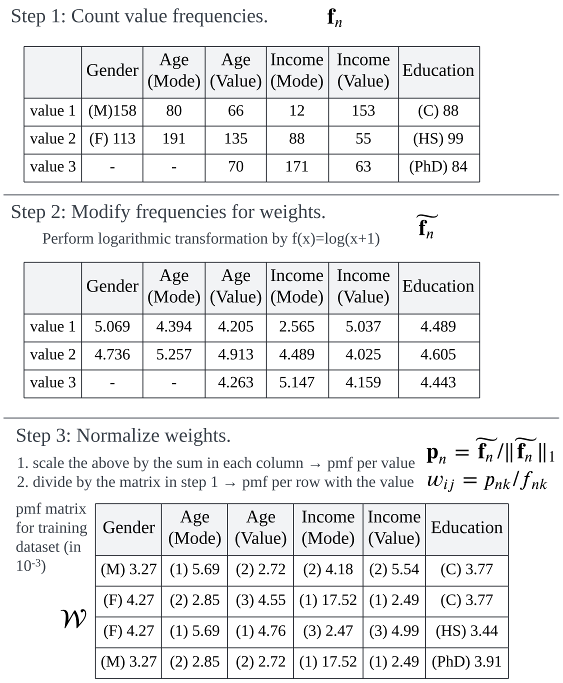

Figure 8: *Weight matrix* calculation explained with the example.

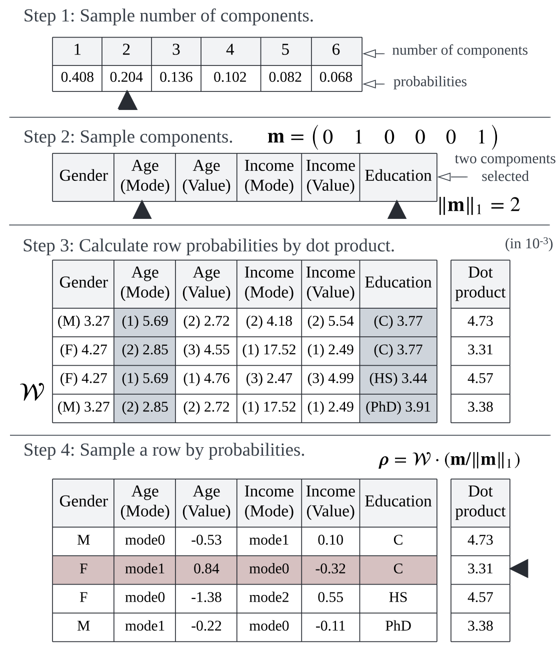

Figure 9: Multivariate data sampling using *weight matrix* explained with the example.

Appendix A. Details of TAEGAN

Algorithm 4 TAEGAN Generation Algorithm

Data: Training table $\mathcal{T}$, generator parameters $\boldsymbol{\theta}_G$

Result: Generated row $\mathbf{x}'$

```
σ ← randomly generated order (permutation of [1, 2, ..., C] with first D_d being categorical indices)
m ← 0, m_{σ_1} ← 1 ;                              /* Initial mask with one selected only */
x' ← randomly sampled γ with the current m
for i in 2, 3, ..., C do
    γ ← μ^T · x' with μ constructed based on m
    Generate x' ← G(m, γ, z) based on noise z from latent space
    m_{σ_i} ← 1 ;                                  /* Update mask with an additional component */
```

The notations related to the construction of *conditional vectors* introduced in Section 3.2 are summarized in Table 8.

Table 8: Summary of the notations related to *conditional vectors*.

component α	$\mathbf{m}_\alpha$	$\boldsymbol{\mu}_\alpha$	$\boldsymbol{\gamma}_\alpha$
maintained after mask	(1)	$\mathbf{1} \in \{0,1\}^{D_\alpha}$	$\mathbf{x}_\alpha$ ($\mathbf{c}_n$ or $\mathbf{d}_n$)
not maintained after mask	(0)	$\mathbf{0} \in \{0,1\}^{D_\alpha}$	$\mathbf{0} \in \mathbb{R}^{D_\alpha}$
empty / disallowed in condition	ϕ	ϕ	ϕ

The TAEGAN generation process described in Section 4.1 and visualized in Fig. 6 can be formally described as the pseudocode in Algorithm 4.

Fig. 8-9 provides an illustrative example with actual values of *weight matrix* calculation (Algorithm 2) and data sampling based on it (Algorithm 3) described in Appendix 4.2. This Fig. aims to aid understanding of the process of data sampling based on logarithmic frequency during training.

Appendix B. Proof of Proposition 1

In CTGAN, $\mathbb{P}(\boldsymbol{\gamma}|\mathbf{m})$ is derived from $\mathbf{p}$ of Algorithm 2 (recall Section 3.3), based on the value in the selected component by $\mathbf{m}$.

In TAEGAN,

$$\begin{aligned}
&\mathbb{P}(\boldsymbol{\gamma}|\mathbf{m})\\
=&\mathbb{P}(\boldsymbol{\gamma} = \boldsymbol{\mu}^T \cdot \mathbf{x}|\mathbf{m})\\
=&\mathbb{P}(\text{a row with } \boldsymbol{\mu}^T \cdot \mathbf{x} = \gamma \text{ is sampled}) \cdot (\text{number of rows with } \boldsymbol{\mu}^T \cdot \mathbf{x} = \gamma)\\
=&\widetilde{\mathbf{p}}^T \cdot \mathbf{f} \text{ from Algorithm 2 mapped to each row's value}\\
=&\mathbf{p} \text{ from Algorithm 2 mapped to each row's value}
\end{aligned} \tag{5}$$

Therefore, these two $\mathbb{P}(\boldsymbol{\gamma}|\mathbf{m})$ values are identical.

Appendix C. Experiment Setup

C.1. Datasets

Table 9 summarizes the 8 datasets used. All datasets are classification datasets for easiness of benchmarking. Datasets are downloaded by `sklearn.datasets.fetch_openml` with the dataset name as input.

C.2. Baselines Implementation

We use the Synthcity (Qian et al., 2023) implementation for ARF (Watson et al., 2023), CTGAN (Xu et al., 2019), GReaT (Borisov et al., 2023), TabDDPM (Kotelnikov et al., 2023), and TVAE (Xu et al., 2019). We also include CTAB-GAN+ (Zhao et al., 2024) and TabSyn (Zhang et al., 2024) using their official code on GitHub.

Table 9: Datasets used in experiments. #R, #F, #N, #D, and #C represent the number of rows, features (including target column), numeric features, discrete (i.e., categorical) features, and classes respectively. Aliases in brackets show the names used in this paper.

Name (Alias)	#R	#F	#N	#D	#C
adult	48842	15	2	13	2
bank-marketing (bank)	45211	17	7	10	2
breast-w (breast)	699	10	9	1	2
credit-g (credit)	1000	21	7	14	2
diabetes	768	9	8	1	2
iris	150	5	4	1	3
qsar-biodeg (qsar)	1055	42	41	1	2
wdbc	569	31	30	1	2

C.3. TAEGAN Configuration

Model dimensions and layers. The dimension of $\mathbf{z}$ is set to 100, among which 50 are discrete and 50 are continuous.

The encoder (of the generator), decoder (of the generator), and discriminator are all MLPs with 6 layers and a hidden dimension of 128. Each layer consists of a batch normalization (Ioffe and Szegedy, 2015) and is activated by ReLU except for the last layer. The encoded dimension of the generator (i.e., output of encoder) is 256. All MLPs are trained with a dropout rate of 0.1.

Training. In each step, the discriminator is updated three times instead of once as shown in Algorithm 1, which is aimed for higher-level illustration. Updating the discriminator more frequently than generator is a common practice for GANs trained with WGAN-GP (Gulrajani et al., 2017). Usually, the discriminator is updated five times in each step for WGAN-GP, but we have a stronger warmed up masked auto-encoder as the generator, so we reduces this value. During training, we apply a weight decay of 1×10^{-5} and gradient norm of 5.

Other hyperparameters. We set $\lambda_1 = 0.1, \lambda_2 = 1, \alpha_1 = 1$. The *outer product* is symmetric, and we use it for correlation, so it suffices to use the upper triangular only for the loss calculation. Both the mean loss and interaction loss's weight are the inverse of their dimensions (e.g., $\alpha_2 = \frac{1}{C}$).

C.4. MLE Details

Hyperparameters of downstream tasks using random forest and XGBoost are tuned by Optuna (Akiba et al., 2019) in 30 trials.

The hyperparameter space explored for random forest is

- **Number of Estimators**: Integers in $[100, 300]$ step 50;
- **Maximum Depth**: Integers in $[5, 20]$ step 5;
- **Minimum Samples Split**: Integers in $[2, 10]$ step 2;
- **Minimum Size per Leaf**: Integers $[1, 5]$;
- **Maximum Features**: Value in "sqrt", "log2", and NULL;
- **Bootstramp**: Enabled or disabled.

The hyperparameter space explored for XGBoost is:

- **Learning Rate**: Logarithmic scale float in $[0.01, 0.3]$;
- **Number of Estimators**: Integers in $[100, 300]$ step 50;
- **Maximum Depth**: Integers in $[3, 10]$;
- **Minimum Child Weight**: Logarithmic scale float in $[1.0, 10.0]$;
- **Minimum Split Loss Gamma**: Float in $[0.0, 0.5]$;
- **Subsample Ratio**: Float in $[0.5, 1.0]$;
- **Subsample Ratio of Columns by Tree**: Float in $[0.5, 1.0]$;

- **L1 Regularization**: Float in $[0.0, 10.0]$.

For all models, numeric values are standardized using standard scaling. Categorical values are one-hot encoded for logistic regression and label-encoded for random forest and XGBoost. To simplify the experiments, rows with missing values are excluded, as handling missing data is not the focus of this paper.

C.5. RSD Details

The training data for the RSD task consists of the real training set combined with a synthetic dataset of the same size as the real training set. The test data comprises the real test set combined with another synthetic dataset of the same size as the real test set. The implementation of ML tasks follow from the MLE implementation.

C.6. Fidelity Details

Fidelity metrics are calculated using SDMetrics (Dat, 2023)'s public SDK, where MD is evaluated by "Shape" and DC is evaluated by "Trend" scores.

C.7. DCR Details

Let the real training dataset be $\mathbf{X}$, real hold-out test set be $\widehat{\mathbf{X}}$, and a synthetic dataset of the same size be $\mathbf{X}'$. We compare the DCR of $(\mathbf{X}, \widehat{\mathbf{X}})$ and $(\mathbf{X}, \mathbf{X}')$. The synthetic data is deemed privacy-preserving if the former values are not smaller than the latter. This is tested using the Mann-Whitney U Test (Mann and Whitney, 1947), where the null hypothesis H_0 states that the distance between $\widehat{\mathbf{X}}$ and $\mathbf{X}$ is greater than or equal to the distance between $\mathbf{X}'$ and $\mathbf{X}$. If the p-value is below 0.05, it indicates that $\mathbf{X}'$ is closer to $\mathbf{X}$ than $\widehat{\mathbf{X}}$, suggesting a potential risk of privacy leakage.

For distance calculations, the data is preprocessed by standard scaling numeric features and applying one-hot encoding to categorical features. Cosine distance is employed to measure the distance between records. As a baseline, we calculate the cosine distances between the $\widehat{\mathbf{X}}$ and the $\mathbf{X}$, recording the minimum distance for each record in $\widehat{\mathbf{X}}$. Similarly, we compute the cosine distances between the $\mathbf{X}'$ and $\mathbf{X}$, also extracting the minimum distances. These distances are then compared to the baseline values to evaluate the similarity and privacy-preserving properties of the synthetic data.

C.8. Efficiency Details

The time records for training and generation are computed from the recorded start and end times of the operations.

C.9. Ablation Experiment Setup

For experiment efficiency, we run ablation experiments on only 3 datasets: `adult`, `credit`, and `diabetes`. The ablation experiments' configurations are shown in Table 10. More settings are provided than the main paper.

Table 10: Ablation experiments setup.

Short form	Description
w/o warmup	No warmup, warmup epochs are changed to main training epochs
gen. in 2 steps	Generating data in 2 steps instead of one component per step (C steps)
w/o log. freq.	Data sampling during training by uniform sampling per row instead of using logarithmic frequency
$\lambda_1 = \lambda_2 = 1$	Uniform weight for different mask ratio
w/o discrete noise	Discrete noise dimensions are set to continuous
w/o pac	Training without PacGAN framework, i.e., pac size is 1
w/o info. loss	Weights of all components of information loss set to 0
w/o interac. loss	Weights of interaction loss set to 0

Table 11: Raw MLE scores in each dataset. All values indicate better performance when the score is higher, except for RE (relative error), which is better when the score is lower. The best scores are highlighted in bold and underscore, and the second best scores are highlighted in bold.

Dataset	ML	real	ARF	CTAB+	CTGAN	TDDPM	GReaT	TabSyn	TVAE	TAEGAN
	LG	0.914	$0.907_{\pm 0.000}$	$0.904_{\pm 0.001}$	$0.886_{\pm 0.009}$	$0.908_{\pm 0.000}$	$\underline{\mathbf{0.911}}_{\pm 0.000}$	$\mathbf{0.910}_{\pm 0.002}$	$0.893_{\pm 0.001}$	$0.907_{\pm 0.001}$
adult	RF	0.910	$0.903_{\pm 0.001}$	$0.902_{\pm 0.002}$	$0.884_{\pm 0.003}$	$0.904_{\pm 0.001}$	$\underline{\mathbf{0.905}}_{\pm 0.000}$	$\mathbf{0.905}_{\pm 0.003}$	$0.892_{\pm 0.001}$	$0.905_{\pm 0.002}$
	XGB	0.914	$0.907_{\pm 0.001}$	$0.905_{\pm 0.001}$	$0.886_{\pm 0.002}$	$0.908_{\pm 0.000}$	$\underline{\mathbf{0.911}}_{\pm 0.000}$	$\mathbf{0.910}_{\pm 0.003}$	$0.895_{\pm 0.001}$	$0.909_{\pm 0.001}$
	LG	0.907	$0.895_{\pm 0.002}$	$\mathbf{0.901}_{\pm 0.003}$	$0.887_{\pm 0.003}$	$0.900_{\pm 0.002}$	$\underline{\mathbf{0.907}}_{\pm 0.000}$	$0.900_{\pm 0.005}$	$0.890_{\pm 0.005}$	$0.899_{\pm 0.000}$
bank	RF	0.930	$0.900_{\pm 0.001}$	$0.902_{\pm 0.002}$	$0.882_{\pm 0.002}$	$\mathbf{0.908}_{\pm 0.001}$	$0.907_{\pm 0.001}$	$0.907_{\pm 0.006}$	$0.884_{\pm 0.002}$	$\underline{\mathbf{0.914}}_{\pm 0.001}$
	XGB	0.936	$0.903_{\pm 0.005}$	$0.906_{\pm 0.000}$	$0.884_{\pm 0.002}$	$0.912_{\pm 0.001}$	$0.910_{\pm 0.001}$	$\mathbf{0.913}_{\pm 0.007}$	$0.885_{\pm 0.001}$	$\underline{\mathbf{0.920}}_{\pm 0.001}$
	LG	0.985	$0.984_{\pm 0.005}$	$0.913_{\pm 0.091}$	$0.978_{\pm 0.006}$	$0.985_{\pm 0.002}$	$0.985_{\pm 0.000}$	$\mathbf{0.987}_{\pm 0.001}$	$0.930_{\pm 0.009}$	$\underline{\mathbf{0.990}}_{\pm 0.001}$
breast	RF	0.985	$0.982_{\pm 0.003}$	$0.968_{\pm 0.012}$	$\mathbf{0.985}_{\pm 0.003}$	$0.981_{\pm 0.002}$	$0.979_{\pm 0.004}$	$0.981_{\pm 0.002}$	$0.976_{\pm 0.007}$	$\underline{\mathbf{0.985}}_{\pm 0.003}$
	XGB	0.984	$0.979_{\pm 0.001}$	$0.958_{\pm 0.025}$	$\mathbf{0.984}_{\pm 0.002}$	$0.981_{\pm 0.001}$	$0.982_{\pm 0.002}$	$\underline{\mathbf{0.987}}_{\pm 0.001}$	$0.981_{\pm 0.005}$	$0.983_{\pm 0.001}$
	LG	0.836	$0.773_{\pm 0.064}$	$0.538_{\pm 0.093}$	$\mathbf{0.814}_{\pm 0.012}$	$0.781_{\pm 0.018}$	$0.665_{\pm 0.000}$	$0.780_{\pm 0.017}$	$0.702_{\pm 0.042}$	$\underline{\mathbf{0.820}}_{\pm 0.015}$
credit	RF	0.837	$0.769_{\pm 0.030}$	$0.505_{\pm 0.097}$	$0.782_{\pm 0.033}$	$0.755_{\pm 0.013}$	$0.722_{\pm 0.014}$	$\mathbf{0.791}_{\pm 0.007}$	$0.746_{\pm 0.057}$	$\underline{\mathbf{0.819}}_{\pm 0.025}$
	XGB	0.844	$0.788_{\pm 0.019}$	$0.509_{\pm 0.088}$	$0.788_{\pm 0.023}$	$0.748_{\pm 0.026}$	$0.739_{\pm 0.006}$	$\mathbf{0.795}_{\pm 0.014}$	$0.759_{\pm 0.048}$	$\underline{\mathbf{0.833}}_{\pm 0.014}$
	LG	0.884	$0.854_{\pm 0.014}$	$0.817_{\pm 0.028}$	$0.864_{\pm 0.017}$	$0.753_{\pm 0.032}$	$0.858_{\pm 0.000}$	$\underline{\mathbf{0.883}}_{\pm 0.006}$	$0.863_{\pm 0.008}$	$\mathbf{0.874}_{\pm 0.010}$
diabetes	RF	0.869	$0.818_{\pm 0.023}$	$0.802_{\pm 0.014}$	$0.818_{\pm 0.008}$	$0.838_{\pm 0.024}$	$0.824_{\pm 0.012}$	$\mathbf{0.844}_{\pm 0.012}$	$0.817_{\pm 0.020}$	$\underline{\mathbf{0.860}}_{\pm 0.021}$
	XGB	0.867	$0.838_{\pm 0.012}$	$0.811_{\pm 0.005}$	$0.807_{\pm 0.005}$	$0.832_{\pm 0.017}$	$0.831_{\pm 0.003}$	$\mathbf{0.847}_{\pm 0.009}$	$0.820_{\pm 0.020}$	$\underline{\mathbf{0.855}}_{\pm 0.010}$
	LG	1.000	$\mathbf{0.997}_{\pm 0.005}$	$0.272_{\pm 0.068}$	$0.886_{\pm 0.025}$	$0.983_{\pm 0.015}$	$0.985_{\pm 0.000}$	$\underline{\mathbf{0.999}}_{\pm 0.002}$	$0.937_{\pm 0.030}$	$0.971_{\pm 0.015}$
iris	RF	1.000	$0.998_{\pm 0.002}$	$0.396_{\pm 0.099}$	$0.936_{\pm 0.018}$	$0.934_{\pm 0.058}$	$\underline{\mathbf{1.000}}_{\pm 0.000}$	$\underline{\mathbf{1.000}}_{\pm 0.000}$	$0.995_{\pm 0.004}$	$0.997_{\pm 0.003}$
	XGB	1.000	$\underline{\mathbf{1.000}}_{\pm 0.000}$	$0.211_{\pm 0.042}$	$0.971_{\pm 0.009}$	$0.817_{\pm 0.203}$	$\underline{\mathbf{1.000}}_{\pm 0.000}$	$\underline{\mathbf{1.000}}_{\pm 0.000}$	$0.986_{\pm 0.004}$	$0.997_{\pm 0.004}$
	LG	0.906	$\underline{\mathbf{0.888}}_{\pm 0.011}$	$0.716_{\pm 0.085}$	$\mathbf{0.869}_{\pm 0.019}$	$0.783_{\pm 0.057}$	$0.673_{\pm 0.000}$	$0.867_{\pm 0.012}$	$0.862_{\pm 0.015}$	$0.864_{\pm 0.016}$
qsar	RF	0.936	$0.877_{\pm 0.021}$	$0.648_{\pm 0.036}$	$0.874_{\pm 0.020}$	$0.714_{\pm 0.080}$	$0.616_{\pm 0.009}$	$\underline{\mathbf{0.882}}_{\pm 0.006}$	$0.850_{\pm 0.026}$	$\mathbf{0.882}_{\pm 0.008}$
	XGB	0.921	$\mathbf{0.872}_{\pm 0.011}$	$0.731_{\pm 0.020}$	$\underline{\mathbf{0.874}}_{\pm 0.010}$	$0.713_{\pm 0.056}$	$0.602_{\pm 0.020}$	$0.860_{\pm 0.016}$	$0.856_{\pm 0.022}$	$0.869_{\pm 0.010}$
	LG	0.993	$0.979_{\pm 0.006}$	$0.929_{\pm 0.053}$	$0.983_{\pm 0.007}$	$0.946_{\pm 0.028}$	$0.979_{\pm 0.000}$	$\underline{\mathbf{0.992}}_{\pm 0.001}$	$0.979_{\pm 0.013}$	$\mathbf{0.989}_{\pm 0.002}$
wdbc	RF	0.976	$0.972_{\pm 0.002}$	$0.919_{\pm 0.024}$	$0.974_{\pm 0.004}$	$0.894_{\pm 0.051}$	$0.976_{\pm 0.003}$	$\mathbf{0.984}_{\pm 0.004}$	$0.976_{\pm 0.001}$	$\underline{\mathbf{0.984}}_{\pm 0.002}$
	XGB	0.990	$0.979_{\pm 0.001}$	$0.930_{\pm 0.021}$	$0.974_{\pm 0.008}$	$0.512_{\pm 0.082}$	$0.973_{\pm 0.003}$	$\mathbf{0.986}_{\pm 0.002}$	$0.977_{\pm 0.006}$	$\underline{\mathbf{0.987}}_{\pm 0.006}$
	All	0.930	0.907	0.750	0.895	0.850	0.868	**0.913**	0.890	$\underline{\mathbf{0.917}}$
	RE (↓)	-	2.632%	19.217%	3.873%	8.604%	6.897%	**1.951%**	4.511%	$\underline{\mathbf{1.422\%}}$
Avg.	LG	0.928	0.909	0.749	0.896	0.880	0.870	$\underline{\mathbf{0.915}}$	0.882	**0.914**
	RF	0.930	0.903	0.755	0.892	0.866	0.866	**0.912**	0.892	$\underline{\mathbf{0.918}}$
	XGB	0.932	0.908	0.745	0.896	0.803	0.869	**0.912**	0.895	$\underline{\mathbf{0.919}}$

Table 12: Raw RSD scores in each dataset. All values indicate better performance when the score is lower (capped at 0.5). The best scores are highlighted in bold and underscore, and the second best scores are highlighted in bold.

Dataset	ML	ARF	CTAB+	CTGAN	TDDPM	GReaT	TabSyn	TVAE	TAEGAN
adult	LG	$0.794_{\pm 0.006}$	$0.632_{\pm 0.016}$	$0.946_{\pm 0.005}$	$\underline{\mathbf{0.585}}_{\pm 0.005}$	$0.714_{\pm 0.000}$	$\mathbf{0.614}_{\pm 0.031}$	$0.945_{\pm 0.003}$	$0.619_{\pm 0.005}$
	RF	$0.877_{\pm 0.003}$	$0.711_{\pm 0.010}$	$0.983_{\pm 0.007}$	$0.739_{\pm 0.002}$	$0.741_{\pm 0.001}$	$\mathbf{0.691}_{\pm 0.067}$	$0.978_{\pm 0.005}$	$\underline{\mathbf{0.670}}_{\pm 0.001}$
	XGB	$0.902_{\pm 0.003}$	$0.734_{\pm 0.011}$	$0.989_{\pm 0.004}$	$0.762_{\pm 0.002}$	$0.758_{\pm 0.001}$	$\mathbf{0.714}_{\pm 0.068}$	$0.985_{\pm 0.003}$	$\underline{\mathbf{0.695}}_{\pm 0.003}$
bank	LG	$0.736_{\pm 0.004}$	$0.607_{\pm 0.013}$	$0.818_{\pm 0.013}$	$\mathbf{0.581}_{\pm 0.004}$	$0.700_{\pm 0.000}$	$0.588_{\pm 0.025}$	$0.826_{\pm 0.021}$	$\underline{\mathbf{0.573}}_{\pm 0.006}$
	RF	$0.869_{\pm 0.014}$	$0.745_{\pm 0.007}$	$0.961_{\pm 0.006}$	$0.783_{\pm 0.003}$	$0.824_{\pm 0.000}$	$\mathbf{0.742}_{\pm 0.060}$	$0.968_{\pm 0.006}$	$\underline{\mathbf{0.657}}_{\pm 0.002}$
	XGB	$0.901_{\pm 0.008}$	$0.795_{\pm 0.008}$	$0.980_{\pm 0.002}$	$0.821_{\pm 0.007}$	$0.842_{\pm 0.001}$	$\mathbf{0.784}_{\pm 0.038}$	$0.983_{\pm 0.002}$	$\underline{\mathbf{0.703}}_{\pm 0.004}$
breast	LG	$\underline{\mathbf{0.605}}_{\pm 0.006}$	$0.723_{\pm 0.014}$	$0.690_{\pm 0.040}$	$\mathbf{0.619}_{\pm 0.013}$	$0.685_{\pm 0.000}$	$0.695_{\pm 0.032}$	$0.702_{\pm 0.046}$	$0.682_{\pm 0.034}$
	RF	$\underline{\mathbf{0.615}}_{\pm 0.014}$	$0.921_{\pm 0.041}$	$0.743_{\pm 0.021}$	$\mathbf{0.623}_{\pm 0.006}$	$0.668_{\pm 0.001}$	$0.670_{\pm 0.019}$	$0.823_{\pm 0.012}$	$0.661_{\pm 0.023}$
	XGB	$\underline{\mathbf{0.604}}_{\pm 0.011}$	$0.922_{\pm 0.041}$	$0.751_{\pm 0.026}$	$\mathbf{0.611}_{\pm 0.005}$	$0.652_{\pm 0.017}$	$0.700_{\pm 0.018}$	$0.850_{\pm 0.017}$	$0.674_{\pm 0.039}$
credit	LG	$\underline{\mathbf{0.505}}_{\pm 0.007}$	$0.634_{\pm 0.066}$	$0.654_{\pm 0.035}$	$0.663_{\pm 0.013}$	$0.708_{\pm 0.000}$	$0.812_{\pm 0.034}$	$0.688_{\pm 0.027}$	$\mathbf{0.564}_{\pm 0.009}$
	RF	$0.858_{\pm 0.003}$	$0.910_{\pm 0.074}$	$0.917_{\pm 0.017}$	$0.794_{\pm 0.026}$	$0.816_{\pm 0.001}$	$\mathbf{0.785}_{\pm 0.038}$	$0.921_{\pm 0.011}$	$\underline{\mathbf{0.608}}_{\pm 0.028}$
	XGB	$0.866_{\pm 0.004}$	$0.920_{\pm 0.083}$	$0.927_{\pm 0.013}$	$\mathbf{0.799}_{\pm 0.021}$	$0.835_{\pm 0.005}$	$0.856_{\pm 0.017}$	$0.942_{\pm 0.008}$	$\underline{\mathbf{0.640}}_{\pm 0.032}$
diabetes	LG	$0.547_{\pm 0.009}$	$0.629_{\pm 0.045}$	$0.767_{\pm 0.026}$	$0.714_{\pm 0.012}$	$0.727_{\pm 0.000}$	$\underline{\mathbf{0.469}}_{\pm 0.021}$	$0.573_{\pm 0.018}$	$\mathbf{0.524}_{\pm 0.029}$
	RF	$0.725_{\pm 0.021}$	$0.845_{\pm 0.016}$	$0.924_{\pm 0.038}$	$0.865_{\pm 0.017}$	$0.811_{\pm 0.001}$	$\underline{\mathbf{0.603}}_{\pm 0.022}$	$0.953_{\pm 0.009}$	$\mathbf{0.677}_{\pm 0.009}$
	XGB	$0.754_{\pm 0.018}$	$0.844_{\pm 0.014}$	$0.948_{\pm 0.019}$	$0.868_{\pm 0.011}$	$0.814_{\pm 0.003}$	$\underline{\mathbf{0.706}}_{\pm 0.012}$	$0.977_{\pm 0.009}$	$\mathbf{0.707}_{\pm 0.022}$
iris	LG	$0.518_{\pm 0.077}$	$0.796_{\pm 0.040}$	$0.867_{\pm 0.002}$	$0.806_{\pm 0.111}$	$0.534_{\pm 0.000}$	$\underline{\mathbf{0.485}}_{\pm 0.081}$	$\mathbf{0.510}_{\pm 0.055}$	$0.541_{\pm 0.096}$
	RF	$0.877_{\pm 0.023}$	$0.989_{\pm 0.008}$	$0.969_{\pm 0.007}$	$0.999_{\pm 0.002}$	$\underline{\mathbf{0.628}}_{\pm 0.029}$	$0.896_{\pm 0.067}$	$0.954_{\pm 0.014}$	$\mathbf{0.665}_{\pm 0.024}$
	XGB	$0.931_{\pm 0.037}$	$0.984_{\pm 0.010}$	$0.944_{\pm 0.003}$	$0.998_{\pm 0.003}$	$\underline{\mathbf{0.570}}_{\pm 0.009}$	$0.896_{\pm 0.052}$	$0.975_{\pm 0.006}$	$\mathbf{0.871}_{\pm 0.015}$
qsar	LG	$\underline{\mathbf{0.659}}_{\pm 0.025}$	$0.838_{\pm 0.031}$	$0.930_{\pm 0.013}$	$1.000_{\pm 0.000}$	$0.821_{\pm 0.000}$	$0.770_{\pm 0.016}$	$0.801_{\pm 0.034}$	$\mathbf{0.693}_{\pm 0.002}$
	RF	$0.972_{\pm 0.006}$	$1.000_{\pm 0.000}$	$0.999_{\pm 0.001}$	$1.000_{\pm 0.000}$	$0.979_{\pm 0.001}$	$\mathbf{0.946}_{\pm 0.018}$	$0.998_{\pm 0.001}$	$\underline{\mathbf{0.838}}_{\pm 0.019}$
	XGB	$0.984_{\pm 0.007}$	$1.000_{\pm 0.000}$	$0.999_{\pm 0.000}$	$1.000_{\pm 0.000}$	$0.989_{\pm 0.002}$	$\mathbf{0.975}_{\pm 0.003}$	$0.998_{\pm 0.001}$	$\underline{\mathbf{0.869}}_{\pm 0.009}$
wdbc	LG	$\mathbf{0.517}_{\pm 0.008}$	$0.832_{\pm 0.049}$	$0.961_{\pm 0.020}$	$0.994_{\pm 0.004}$	$0.675_{\pm 0.000}$	$0.658_{\pm 0.018}$	$0.743_{\pm 0.024}$	$\underline{\mathbf{0.510}}_{\pm 0.037}$
	RF	$0.948_{\pm 0.007}$	$1.000_{\pm 0.000}$	$0.993_{\pm 0.006}$	$1.000_{\pm 0.000}$	$0.965_{\pm 0.003}$	$\underline{\mathbf{0.616}}_{\pm 0.057}$	$0.989_{\pm 0.001}$	$\mathbf{0.633}_{\pm 0.050}$
	XGB	$0.939_{\pm 0.006}$	$0.999_{\pm 0.000}$	$0.990_{\pm 0.008}$	$1.000_{\pm 0.000}$	$0.968_{\pm 0.004}$	$\underline{\mathbf{0.677}}_{\pm 0.065}$	$0.994_{\pm 0.001}$	$\mathbf{0.698}_{\pm 0.018}$
Avg.	All	0.771	0.834	0.902	0.818	0.768	**0.723**	0.878	$\underline{\mathbf{0.666}}$
	LG	**0.610**	0.711	0.829	0.745	0.696	0.636	0.724	$\underline{\mathbf{0.588}}$
	RF	0.843	0.890	0.936	0.850	0.804	**0.743**	0.948	$\underline{\mathbf{0.676}}$
	XGB	0.860	0.900	0.941	0.857	0.803	**0.789**	0.963	$\underline{\mathbf{0.732}}$

Appendix D. Detailed Experiment Results

D.1. Baseline Comparison

Table 11 shows the raw scores of MLE comparing different baseline models with TAEGAN. TAEGAN is better than all baseline models in general. In particular, TAEGAN is significantly better than baseline GAN models.

Table 12 shows the raw scores of RSD comparing different baseline models with TAEGAN. TAEGAN demonstrates wins over all baseline models by a great margin.

Table 13 shows the raw scores of MD and CR for fidelity comparing different baseline models with TAEGAN. TAEGAN is not the best among all experimented models but generally maintains a comparable performance with the state-of-the-art models, and is the best GAN model.

Table 14 shows the raw p-values indicating the risk of privacy leakage computed based on DCR. TAEGAN and many good-performing baseline models do not suffer from outstanding privacy leakage risk.

Table 15 shows the raw computation time on each dataset. Results show that GANs are generally more efficient.

Table 13: Raw fidelity scores in each dataset. All values indicate better performance when the score is higher. The best scores are highlighted in bold and underscore, and the second best scores are highlighted in bold. The best scores among GANs are highlighted with superscript *.

Dataset	Metric	ARF	CTAB+	CTGAN	TDDPM	GReaT	TabSyn	TVAE	TAEGAN
adult	MD	$0.942_{\pm0.001}$	$0.964^{*}_{\pm0.005}$	$0.881_{\pm0.008}$	$\underline{\mathbf{0.983}}_{\pm0.001}$	$0.929_{\pm0.000}$	$\mathbf{0.977}_{\pm0.011}$	$0.877_{\pm0.004}$	$0.958_{\pm0.002}$
	CR	$0.876_{\pm0.004}$	$0.915_{\pm0.004}$	$0.760_{\pm0.005}$	$\mathbf{0.953}_{\pm0.006}$	$0.882_{\pm0.000}$	$\underline{\mathbf{0.955}}_{\pm0.018}$	$0.726_{\pm0.002}$	$0.919^{*}_{\pm0.002}$
bank	MD	$0.947_{\pm0.000}$	$0.965_{\pm0.002}$	$0.891_{\pm0.001}$	$\underline{\mathbf{0.983}}_{\pm0.001}$	$0.915_{\pm0.000}$	$\mathbf{0.983}_{\pm0.005}$	$0.894_{\pm0.003}$	$0.966^{*}_{\pm0.002}$
	CR	$0.885_{\pm0.019}$	$0.884^{*}_{\pm0.003}$	$0.870_{\pm0.009}$	$\mathbf{0.932}_{\pm0.030}$	$0.903_{\pm0.000}$	$\underline{\mathbf{0.967}}_{\pm0.008}$	$0.870_{\pm0.011}$	$0.877_{\pm0.014}$
breast	MD	$\mathbf{0.854}_{\pm0.005}$	$0.775_{\pm0.024}$	$0.832_{\pm0.019}$	$0.838_{\pm0.001}$	$0.728_{\pm0.000}$	$0.813_{\pm0.005}$	$\underline{\mathbf{0.900}}_{\pm0.013}$	$0.833^{*}_{\pm0.009}$
	CR	$\mathbf{0.767}_{\pm0.006}$	$0.616_{\pm0.063}$	$0.709_{\pm0.015}$	$0.767_{\pm0.000}$	$0.636_{\pm0.000}$	$0.709_{\pm0.006}$	$\underline{\mathbf{0.781}}_{\pm0.010}$	$0.747^{*}_{\pm0.004}$
credit	MD	$\underline{\mathbf{0.971}}_{\pm0.003}$	$0.937_{\pm0.013}$	$0.944_{\pm0.003}$	$0.926_{\pm0.007}$	$0.932_{\pm0.000}$	$0.932_{\pm0.007}$	$0.931_{\pm0.004}$	$\mathbf{0.968}^{*}_{\pm0.003}$
	CR	$\mathbf{0.925}_{\pm0.004}$	$0.862_{\pm0.030}$	$0.894_{\pm0.005}$	$0.864_{\pm0.007}$	$0.861_{\pm0.000}$	$0.862_{\pm0.008}$	$0.848_{\pm0.004}$	$\underline{\mathbf{0.934}}^{*}_{\pm0.007}$
diabetes	MD	$0.898_{\pm0.004}$	$0.923_{\pm0.006}$	$0.793_{\pm0.042}$	$0.769_{\pm0.017}$	$0.881_{\pm0.000}$	$\underline{\mathbf{0.955}}_{\pm0.007}$	$0.824_{\pm0.012}$	$\mathbf{0.929}^{*}_{\pm0.008}$
	CR	$\mathbf{0.949}_{\pm0.007}$	$0.898_{\pm0.004}$	$0.913_{\pm0.001}$	$0.875_{\pm0.003}$	$0.921_{\pm0.000}$	$\underline{\mathbf{0.964}}_{\pm0.002}$	$0.858_{\pm0.008}$	$0.946^{*}_{\pm0.001}$
iris	MD	$\mathbf{0.893}_{\pm0.011}$	$0.822_{\pm0.022}$	$0.762_{\pm0.040}$	$0.553_{\pm0.024}$	$0.870_{\pm0.000}$	$\underline{\mathbf{0.893}}_{\pm0.003}$	$0.809_{\pm0.004}$	$0.872^{*}_{\pm0.007}$
	CR	$0.897_{\pm0.007}$	$0.590_{\pm0.019}$	$0.793_{\pm0.047}$	$0.506_{\pm0.036}$	$\mathbf{0.899}_{\pm0.000}$	$\underline{\mathbf{0.917}}_{\pm0.007}$	$0.783_{\pm0.017}$	$0.874^{*}_{\pm0.010}$
qsar	MD	$\underline{\mathbf{0.931}}_{\pm0.004}$	$0.917^{*}_{\pm0.003}$	$0.891_{\pm0.001}$	$0.602_{\pm0.005}$	$0.887_{\pm0.000}$	$\mathbf{0.930}_{\pm0.005}$	$0.873_{\pm0.007}$	$0.898_{\pm0.001}$
	CR	$\mathbf{0.911}_{\pm0.009}$	$0.863_{\pm0.007}$	$0.887^{*}_{\pm0.010}$	$0.605_{\pm0.001}$	$0.855_{\pm0.000}$	$\underline{\mathbf{0.914}}_{\pm0.009}$	$0.849_{\pm0.012}$	$0.874_{\pm0.006}$
wdbc	MD	$0.923_{\pm0.003}$	$0.861_{\pm0.005}$	$0.824_{\pm0.031}$	$0.395_{\pm0.017}$	$0.873_{\pm0.000}$	$\underline{\mathbf{0.946}}_{\pm0.004}$	$0.749_{\pm0.026}$	$\mathbf{0.928}^{*}_{\pm0.005}$
	CR	$0.939_{\pm0.001}$	$0.866_{\pm0.007}$	$0.934_{\pm0.003}$	$0.772_{\pm0.010}$	$0.896_{\pm0.000}$	$\underline{\mathbf{0.976}}_{\pm0.002}$	$0.907_{\pm0.017}$	$\mathbf{0.951}^{*}_{\pm0.002}$
Avg.	MD	**0.925**	0.898	0.852	0.751	0.881	$\underline{\mathbf{0.931}}$	0.860	0.906*
	CR	**0.904**	0.825	0.857	0.783	0.867	$\underline{\mathbf{0.919}}$	0.839	0.892*

Table 14: Raw p-values of DCR values indicating risk of privacy leakage. Values smaller than 0.05 are risky, which are highlighted in red.

	ARF	CTAB+	CTGAN	TDDPM	GReaT	TabSyn	TVAE	TAEGAN
adult	$0.486_{\pm0.077}$	$0.691_{\pm0.424}$	$0.687_{\pm0.272}$	$0.440_{\pm0.125}$	$0.000_{\pm0.000}$	$0.654_{\pm0.149}$	$0.003_{\pm0.004}$	$0.938_{\pm0.036}$
bank	$0.999_{\pm0.000}$	$0.849_{\pm0.093}$	$0.996_{\pm0.003}$	$0.873_{\pm0.077}$	$0.040_{\pm0.000}$	$0.938_{\pm0.027}$	$0.709_{\pm0.377}$	$0.729_{\pm0.182}$
breast	$0.071_{\pm0.060}$	$0.333_{\pm0.468}$	$0.595_{\pm0.415}$	$0.039_{\pm0.026}$	$0.005_{\pm0.000}$	$0.357_{\pm0.256}$	$1.000_{\pm0.000}$	$0.220_{\pm0.116}$
credit	$0.197_{\pm0.068}$	$0.161_{\pm0.124}$	$0.462_{\pm0.131}$	$0.725_{\pm0.261}$	$0.990_{\pm0.000}$	$0.689_{\pm0.246}$	$0.581_{\pm0.194}$	$0.442_{\pm0.134}$
diabetes	$0.698_{\pm0.230}$	$0.890_{\pm0.036}$	$0.835_{\pm0.187}$	$0.990_{\pm0.005}$	$0.971_{\pm0.000}$	$0.825_{\pm0.093}$	$0.909_{\pm0.049}$	$0.840_{\pm0.176}$
iris	$0.796_{\pm0.127}$	$0.970_{\pm0.039}$	$0.932_{\pm0.032}$	$0.972_{\pm0.038}$	$0.187_{\pm0.000}$	$0.765_{\pm0.149}$	$0.770_{\pm0.118}$	$0.677_{\pm0.229}$
qsar	$0.668_{\pm0.231}$	$0.546_{\pm0.193}$	$0.669_{\pm0.186}$	$0.485_{\pm0.306}$	$0.063_{\pm0.000}$	$0.664_{\pm0.238}$	$0.351_{\pm0.160}$	$0.865_{\pm0.100}$
wdbc	$0.821_{\pm0.029}$	$0.647_{\pm0.283}$	$0.909_{\pm0.068}$	$1.000_{\pm0.000}$	$0.591_{\pm0.000}$	$0.853_{\pm0.096}$	$0.307_{\pm0.170}$	$0.935_{\pm0.036}$
# vio.	0	0	0	1	3	0	1	0

D.2. Ablation Study

Table 16 shows the raw results of the ablation study on MLE. Original TAEGAN setting has the best score compared to all the other ablation settings, which are TAEGAN with one design component removed. The advantage is more obvious with more complicated downstream task (XGB).

Table 17 shows the raw results of the ablation study on fidelity metrics. Note that when uniform sampling instead of data sampling following logarithmic frequency is used, all fidelity metrics in all datasets are improved.

Table 15: Time used for training and generation respectively in seconds of different models on each dataset.

Dataset	Action	ARF	CTAB+	CTGAN	TDDPM	GReaT	TabSyn	TVAE	TAEGAN
adult	Train	$258.758_{\pm 55.030}$	$787.379_{\pm 3.334}$	$753.259_{\pm 104.080}$	$285.153_{\pm 0.410}$	$10397.040_{\pm 9.203}$	$1199.823_{\pm 146.744}$	$1111.927_{\pm 2.790}$	$758.236_{\pm 1.879}$
	Generate	$41.029_{\pm 0.363}$	$0.429_{\pm 0.001}$	$0.112_{\pm 0.001}$	$63.388_{\pm 0.083}$	$138.023_{\pm 0.205}$	$2.409_{\pm 0.010}$	$0.357_{\pm 0.005}$	$2.204_{\pm 0.079}$
bank	Train	$121.480_{\pm 39.164}$	$737.314_{\pm 0.138}$	$544.578_{\pm 62.697}$	$252.016_{\pm 0.266}$	$10295.769_{\pm 21.720}$	$1865.473_{\pm 418.665}$	$878.170_{\pm 122.998}$	$832.752_{\pm 0.931}$
	Generate	$42.208_{\pm 5.408}$	$1.126_{\pm 0.006}$	$0.733_{\pm 0.001}$	$55.038_{\pm 0.077}$	$149.872_{\pm 0.104}$	$3.009_{\pm 0.063}$	$0.975_{\pm 0.003}$	$2.782_{\pm 0.006}$
breast	Train	$0.585_{\pm 0.015}$	$8.479_{\pm 0.026}$	$13.571_{\pm 2.035}$	$6.136_{\pm 0.099}$	$153.352_{\pm 0.134}$	$1037.126_{\pm 281.166}$	$6.340_{\pm 0.925}$	$37.559_{\pm 0.480}$
	Generate	$0.876_{\pm 0.013}$	$0.056_{\pm 0.008}$	$0.051_{\pm 0.000}$	$3.123_{\pm 0.007}$	$3.434_{\pm 0.016}$	$0.249_{\pm 0.003}$	$0.055_{\pm 0.000}$	$0.090_{\pm 0.000}$
credit	Train	$2.262_{\pm 0.089}$	$8.400_{\pm 0.035}$	$32.463_{\pm 8.548}$	$8.102_{\pm 0.215}$	$277.555_{\pm 0.388}$	$726.006_{\pm 67.662}$	$13.246_{\pm 1.809}$	$68.799_{\pm 1.766}$
	Generate	$2.198_{\pm 0.088}$	$0.043_{\pm 0.000}$	$0.035_{\pm 0.000}$	$4.526_{\pm 0.007}$	$5.845_{\pm 0.008}$	$0.167_{\pm 0.003}$	$0.041_{\pm 0.000}$	$0.133_{\pm 0.001}$
diabetes	Train	$1.100_{\pm 0.094}$	$8.524_{\pm 0.407}$	$20.414_{\pm 2.831}$	$5.799_{\pm 1.191}$	$130.793_{\pm 0.077}$	$758.140_{\pm 41.333}$	$7.944_{\pm 1.272}$	$54.576_{\pm 0.720}$
	Generate	$0.943_{\pm 0.047}$	$0.047_{\pm 0.001}$	$0.036_{\pm 0.000}$	$2.329_{\pm 0.822}$	$1.496_{\pm 0.008}$	$0.139_{\pm 0.000}$	$0.039_{\pm 0.000}$	$0.096_{\pm 0.001}$
iris	Train	$0.153_{\pm 0.002}$	$7.267_{\pm 0.021}$	$6.538_{\pm 0.504}$	$3.485_{\pm 0.210}$	$24.975_{\pm 0.018}$	$620.864_{\pm 35.603}$	$2.811_{\pm 0.518}$	$13.702_{\pm 0.040}$
	Generate	$0.690_{\pm 0.010}$	$0.033_{\pm 0.006}$	$0.016_{\pm 0.000}$	$6.208_{\pm 2.124}$	$0.360_{\pm 0.015}$	$0.098_{\pm 0.002}$	$0.016_{\pm 0.001}$	$0.039_{\pm 0.000}$
qsar	Train	$18.992_{\pm 1.985}$	$18.389_{\pm 0.217}$	$64.331_{\pm 11.582}$	$7.369_{\pm 0.156}$	$566.658_{\pm 23.451}$	$694.890_{\pm 76.489}$	$23.013_{\pm 2.183}$	$209.272_{\pm 3.010}$
	Generate	$2.842_{\pm 0.012}$	$0.206_{\pm 0.006}$	$0.192_{\pm 0.000}$	$2.825_{\pm 0.826}$	$46.712_{\pm 3.062}$	$0.336_{\pm 0.005}$	$0.203_{\pm 0.001}$	$0.506_{\pm 0.070}$
wdbc	Train	$5.344_{\pm 0.860}$	$10.795_{\pm 0.089}$	$24.175_{\pm 7.507}$	$4.249_{\pm 0.056}$	$265.340_{\pm 2.411}$	$631.546_{\pm 68.070}$	$9.356_{\pm 1.236}$	$109.288_{\pm 0.061}$
	Generate	$1.410_{\pm 0.007}$	$0.159_{\pm 0.001}$	$0.154_{\pm 0.002}$	$1.934_{\pm 0.010}$	$65.821_{\pm 1.106}$	$0.276_{\pm 0.005}$	$0.159_{\pm 0.000}$	$0.302_{\pm 0.002}$
Avg.	Train	51.933	187.517	178.590	67.330	2660.705	1023.474	226.858	254.823
	Generate	10.147	0.302	0.208	15.776	47.157	0.952	0.270	0.798

Table 16: Results of ablation study on MLE.

	adult			credit			diabetes			Avg.			
ML	LG	RF	XGB	LG	RF	XGB	LG	RF	XGB	LG	RF	XGB	All
TAEGAN	$0.907_{\pm 0.001}$	$0.905_{\pm 0.002}$	$0.909_{\pm 0.001}$	$0.820_{\pm 0.015}$	$0.819_{\pm 0.025}$	$0.833_{\pm 0.014}$	$0.874_{\pm 0.010}$	$0.860_{\pm 0.021}$	$0.855_{\pm 0.010}$	0.867	0.861	0.866	0.865
w/o warmup	$0.907_{\pm 0.001}$	$0.905_{\pm 0.002}$	$0.910_{\pm 0.001}$	$0.814_{\pm 0.007}$	$0.801_{\pm 0.016}$	$0.821_{\pm 0.011}$	$0.863_{\pm 0.016}$	$0.837_{\pm 0.037}$	$0.838_{\pm 0.029}$	0.861	0.848	0.857	0.855
gen. in 2 steps	$0.907_{\pm 0.001}$	$0.905_{\pm 0.001}$	$0.910_{\pm 0.000}$	$0.820_{\pm 0.015}$	$0.818_{\pm 0.024}$	$0.830_{\pm 0.015}$	$0.874_{\pm 0.010}$	$0.855_{\pm 0.022}$	$0.854_{\pm 0.022}$	0.867	0.859	0.864	0.864
w/o log. freq.	$0.906_{\pm 0.001}$	$0.904_{\pm 0.002}$	$0.910_{\pm 0.001}$	$0.822_{\pm 0.023}$	$0.817_{\pm 0.014}$	$0.817_{\pm 0.025}$	$0.876_{\pm 0.009}$	$0.856_{\pm 0.012}$	$0.851_{\pm 0.004}$	0.868	0.859	0.859	0.862
$\lambda_1 = \lambda_2 = 1$	$0.905_{\pm 0.001}$	$0.903_{\pm 0.001}$	$0.909_{\pm 0.001}$	$0.804_{\pm 0.008}$	$0.809_{\pm 0.032}$	$0.798_{\pm 0.042}$	$0.870_{\pm 0.010}$	$0.863_{\pm 0.014}$	$0.850_{\pm 0.012}$	0.859	0.858	0.852	0.857
w/o discrete noise	$0.907_{\pm 0.001}$	$0.905_{\pm 0.002}$	$0.910_{\pm 0.001}$	$0.807_{\pm 0.014}$	$0.816_{\pm 0.027}$	$0.830_{\pm 0.025}$	$0.876_{\pm 0.008}$	$0.856_{\pm 0.006}$	$0.853_{\pm 0.011}$	0.864	0.859	0.864	0.862
w/o pac	$0.906_{\pm 0.001}$	$0.904_{\pm 0.001}$	$0.909_{\pm 0.001}$	$0.816_{\pm 0.019}$	$0.803_{\pm 0.022}$	$0.828_{\pm 0.003}$	$0.869_{\pm 0.014}$	$0.854_{\pm 0.003}$	$0.851_{\pm 0.003}$	0.864	0.854	0.862	0.860
w/o info. loss	$0.907_{\pm 0.001}$	$0.905_{\pm 0.001}$	$0.910_{\pm 0.000}$	$0.819_{\pm 0.014}$	$0.821_{\pm 0.020}$	$0.819_{\pm 0.029}$	$0.874_{\pm 0.009}$	$0.857_{\pm 0.016}$	$0.855_{\pm 0.012}$	0.867	0.861	0.861	0.863
w/o interac. loss	$0.907_{\pm 0.001}$	$0.904_{\pm 0.003}$	$0.909_{\pm 0.001}$	$0.819_{\pm 0.016}$	$0.819_{\pm 0.016}$	$0.833_{\pm 0.017}$	$0.873_{\pm 0.009}$	$0.860_{\pm 0.019}$	$0.852_{\pm 0.014}$	0.866	0.861	0.865	0.864

Table 17: Results of ablation study on fidelity.

	adult		credit		diabetes		Avg.	
Metric	MD	CR	MD	CR	MD	CR	MD	CR
TAEGAN	$0.958_{\pm 0.002}$	$0.919_{\pm 0.002}$	$0.968_{\pm 0.003}$	$0.934_{\pm 0.007}$	$0.929_{\pm 0.008}$	$0.946_{\pm 0.001}$	0.952	0.933
w/o warmup	$0.959_{\pm 0.003}$	$0.919_{\pm 0.005}$	$0.969_{\pm 0.002}$	$0.918_{\pm 0.006}$	$0.933_{\pm 0.004}$	$0.948_{\pm 0.003}$	0.954	0.928
gen. in 2 steps	$0.958_{\pm 0.002}$	$0.919_{\pm 0.002}$	$0.968_{\pm 0.003}$	$0.934_{\pm 0.007}$	$0.929_{\pm 0.008}$	$0.946_{\pm 0.001}$	0.952	0.933
w/o log. freq.	$0.968_{\pm 0.000}$	$0.932_{\pm 0.000}$	$0.970_{\pm 0.004}$	$0.935_{\pm 0.007}$	$0.932_{\pm 0.007}$	$0.948_{\pm 0.001}$	0.957	0.938
$\lambda_1 = \lambda_2 = 1$	$0.963_{\pm 0.002}$	$0.922_{\pm 0.003}$	$0.967_{\pm 0.004}$	$0.931_{\pm 0.007}$	$0.931_{\pm 0.011}$	$0.945_{\pm 0.002}$	0.953	0.933
w/o discrete noise	$0.960_{\pm 0.001}$	$0.920_{\pm 0.001}$	$0.966_{\pm 0.004}$	$0.922_{\pm 0.015}$	$0.929_{\pm 0.008}$	$0.945_{\pm 0.004}$	0.952	0.929
w/o pac	$0.959_{\pm 0.001}$	$0.919_{\pm 0.003}$	$0.965_{\pm 0.002}$	$0.913_{\pm 0.005}$	$0.929_{\pm 0.001}$	$0.946_{\pm 0.002}$	0.951	0.926
w/o info. loss	$0.958_{\pm 0.001}$	$0.918_{\pm 0.002}$	$0.968_{\pm 0.004}$	$0.934_{\pm 0.008}$	$0.930_{\pm 0.008}$	$0.946_{\pm 0.002}$	0.952	0.933
w/o interac. loss	$0.958_{\pm 0.002}$	$0.919_{\pm 0.002}$	$0.968_{\pm 0.003}$	$0.934_{\pm 0.007}$	$0.929_{\pm 0.009}$	$0.946_{\pm 0.002}$	0.952	0.933