Proof

Notations. For each database X_i of size n, we represent it by $1/c$ blocks (each has cn rows). Let $X_i = [x_{i1}, x_{i2}, \ldots, x_{i\frac{1}{c}}]^\top$, where x_{ij} denote the $j$the block in X_i. Let $X_{i\downarrow} = [x_{i2}, \ldots, x_{i\frac{1}{c}}]^\top$ and $X_{i\uparrow} = [x_{i1}, \ldots, x_{i(\frac{1}{c}-1)}]^\top$. Let $\Delta = \{D^n\}$ denote all distributions over the database, with each row chosen i.i.d. from some D. $\Delta_\downarrow$ contains Δ and also the auxiliary information $X_\downarrow$. $\Delta_\uparrow$ contains Δ and also the auxiliary information $X_\uparrow$.

Result 1. Assume (1) Mechanism F is $(\epsilon_\downarrow, \delta_\downarrow, \Delta_\downarrow, \Gamma)$-CW private with simulator $sim_{F\downarrow}$; (2) Mechanism G is $(\epsilon_\uparrow, \delta_\uparrow, \Delta_\uparrow, \Gamma)$-CW private with simulator $sim_{G\uparrow}$. Then, Mechanism $H(X) = (F(X_1), G(X_2))$ is $(\epsilon_\downarrow + \epsilon_\uparrow, \delta_\downarrow + \delta_\uparrow, \Delta_{2\uparrow}, \Gamma)$-CW private with simulator $(sim_{F\downarrow}, sim_{G\uparrow})$.

Proof. For any set $S = (S_1, S_2)$, we have

$$Pr[H(X) \in S \,|\, X_{2\uparrow} = z, \mathbf{priv}(X) = v]$$

(omit $\mathbf{priv}(X) = v$ from now on)

$$= Pr[F(X_1, G(X_2) \in (S_1, S_2) \,|\, X_{2\uparrow} = z]$$

$$= Pr[F(x_{11}, z) \in S_1, G(z, x_{2\frac{1}{c}}) \in S_2 \,|\, X_{2\uparrow} = z]$$

Since (1) $F(x_{11}, z)/G(z, x_{2\frac{1}{c}})$ is a (measurable) function on $x_{11}/x_{2\frac{1}{c}}$, where $x_{11}/x_{2\frac{1}{c}}$ is a block of cn rows; (2) x_{11} and $x_{2\frac{1}{c}}$ are i.i.d. from some D, we know $F(x_{11}, z)$ and $G(z, x_{2\frac{1}{c}})$ are independent with each other. Therefore, we have

$$= Pr[F(x_{11}, z) \in S_1 \,|\, X_{2\uparrow} = z] \cdot Pr[G(z, x_{2\frac{1}{c}}) \in S_2 \,|\, X_{2\uparrow} = z]$$

$$= Pr[F(x_{11}, z) \in S_1 \,|\, X_{1\downarrow} = z] \cdot Pr[G(z, x_{2\frac{1}{c}}) \in S_2 \,|\, X_{2\uparrow} = z]$$

$$= Pr[F(X_1) \in S_1 \,|\, X_{1\downarrow} = z] \cdot Pr[G(X_2) \in S_2 \,|\, X_{2\uparrow} = z]$$

By the assumptions on F and G, we have

$$\leq (e^{\epsilon_\downarrow} Pr[sim_{F\downarrow}(\mathbf{alt}(X_1)) \in S_1 \mid X_{1\downarrow} = z] + \delta_\downarrow) \cdot (e^{\epsilon_\uparrow} Pr[sim_{G\uparrow}(\mathbf{alt}(X_2)) \in S_2 \mid X_{2\uparrow} = z] + \delta_\uparrow)$$

$$\leq e^{\epsilon_\downarrow + \epsilon_\uparrow}(Pr[sim_{F\downarrow}(\mathbf{alt}(X_1)) \in S_1 \mid X_{1\downarrow} = z] \cdot Pr[sim_{G\uparrow}(\mathbf{alt}(X_2)) \in S_2 \mid X_{2\uparrow} = z]) + (\delta_\downarrow + \delta_\uparrow)$$

$$= e^{\epsilon_\downarrow + \epsilon_\uparrow}(Pr[sim_{F\downarrow}(\mathbf{alt}(x_{11}), \mathbf{alt}(z)) \in S_1 \mid X_{1\downarrow} = z] \cdot Pr[sim_{G\uparrow}(\mathbf{alt}(z), \mathbf{alt}(x_{2\frac{1}{c}})) \in S_2 \mid X_{2\uparrow} = z])$$
$$+ (\delta_\downarrow + \delta_\uparrow)$$

Notice $sim_{F\downarrow}(\mathbf{alt}(X)$ can be considered as a composed function $sim_{F\downarrow} \circ \mathbf{alt}$ on X. In our case when $X_{1\downarrow} = X_{2\uparrow} = z$, $sim_{F\downarrow}(\mathbf{alt}(X_1)$ is a function on x_{11} and $sim_{G\uparrow}(\mathbf{alt}(X_2)$ is a function on $x_{2\frac{1}{c}}$. Since x_{11} and $x_{2\frac{1}{c}}$ are i.i.d. from some D, we have the independence.

$$= e^{\epsilon_\downarrow + \epsilon_\uparrow}(Pr[sim_{F\downarrow}(\mathbf{alt}(x_{11}), \mathbf{alt}(z)) \in S_1, sim_{G\uparrow}(\mathbf{alt}(z), \mathbf{alt}(x_{2\frac{1}{c}})) \in S_2 \mid X_{2\uparrow} = z] + (\delta_\downarrow + \delta_\uparrow)$$

$$= e^{\epsilon_\downarrow + \epsilon_\uparrow}(Pr[(sim_{F\downarrow}(\mathbf{alt}(X_1)), sim_{G\uparrow}(\mathbf{alt}(X_2))) \in (S_1, S_2) \mid X_{2\uparrow} = z] + (\delta_\downarrow + \delta_\uparrow) \quad \square$$

Result 2 Let $G(X) = (F(X_1), F(X_2), \ldots, F(X_t))$. Let $F : \mathcal{U}^n \to \mathbb{R}^d$, such that $F(X) = \Sigma_{i=1}^{\frac{1}{c}} F(x_i)$ for all database $X = [x_1, \ldots, x_{\frac{1}{c}}]^\top \in \mathcal{U}^n$, where x_i is a block of cn rows. Assume (1) Mechanism F is $(\epsilon_\downarrow, \delta_\downarrow, \Delta_\downarrow, \Gamma)$-CW private with simulator $sim_{F\downarrow}$; (2) Mechanism F is $(\epsilon_\uparrow, \delta_\uparrow, \Delta_\uparrow, \Gamma)$-CW private with simulator $sim_{F\uparrow}$. Then,

1. $G(X)$ is $(t\epsilon_\downarrow, t\delta_\downarrow +, \Delta_{t\downarrow}, \Gamma)$-CW private with simulator $(sim_{F\downarrow})^t$;
2. $G(X)$ is $((t-1)\epsilon_\downarrow + \epsilon_\uparrow, (t-1)\delta_\downarrow + \delta_\uparrow, \Delta_{t\uparrow}, \Gamma)$-CW private with simulator $((sim_{F\downarrow})^{t-1}, sim_{F\uparrow})$;

Additional Notations. Let $X_{i\downarrow k} = [x_{i(k+1)}, \ldots, x_{i\frac{1}{c}}]^\top$. We still use $X_{i\downarrow}$ to denote $X_{i\downarrow 1}$. Let $S = (S_1, \ldots, S_{t-1}, S_t)$ and $S_{-t} = (S_1, \ldots, S_{t-1})$. Let $S_{-t}(v)$ denote the set of all $\mathbf{v_{-t}}$ such that $(\mathbf{v_{-t}}, v) \in S$.

Proof (1). For any set $S = (S_1, \ldots, S_{t-1}, S_t)$, we have

$$Pr[G(X) \in S \mid X_{t\downarrow} = z]$$

$$= Pr[(F(X_1, \ldots, F(X_{t-1})) \in (S_1, \ldots, S_{t-1}), F(X_t) \in S_t \mid X_{t\downarrow} = z]$$
$$= \Pi_{j=1}^t Pr[F(X_j) \in S_j \mid F(X_{j+1}) \in S_{j=1}, \ldots, F(X_t) \in S_t, X_{t\downarrow} = z]$$

For ease of analysis, assume $t < n$. That is, X_1 and X_t still have at least one overlapped row. For each j, we have

$$= Pr[F(x_{j1}, \ldots, x_{j(t+1-j)}, X_{t\downarrow(t+1-j)}) \in S_j \mid F(X_{j+1}) \in S_{j=1}, \ldots, F(X_t) \in S_t, X_{t\downarrow} = z]$$

$$= \Sigma_{v_{j+1} \in S_{j+1}} Pr[F(x_{j1}, \ldots, x_{j(t+1-j)}, X_{t\downarrow(t+1-j)}) \in S_j \mid F(X_{j+1}) = v_{j+1}, F(X_{j+2}) \in S_{j+2},$$
$$\ldots, F(X_t) \in S_t, X_{t\downarrow} = z] \cdot Pr[F(X_{j+1}) = v_{j+1}]$$

Notice that $F(x_{j1}, \ldots, x_{j(t+1-j)}, X_{t\downarrow(t+1-j)}) = F(x_{j1}) + F(X_{j+1}) - F(x_{(j+1)\frac{1}{c}})$, of which $x_{(j+1)\frac{1}{c}} \in X_{t\downarrow}$. Let $F(x_{(j+1)\frac{1}{c}}) = v^*_{j+1}$, and we have

$$= \Sigma_{v_{j+1} \in S_{j+1}} Pr[(F(x_{j1}) + F(X_{j\downarrow})) \in S_j \mid F(X_{j\downarrow}) = v_{j+1} - v^*_{j+1}, F(X_{j+1}) = v_{j+1}, F(X_{j+2}) \in S_{j+2},$$
$$\ldots, F(X_t) \in S_t, X_{t\downarrow} = z] \cdot Pr[F(X_{j+1}) = v_{j+1}]$$

Notice block x_{j1} is independent w.r.t. all the conditions, and $X_j \cap X_{t\downarrow} = X_{t\downarrow(t+1-j)})$. Let $X_{t\downarrow(t+1-j)}) = z_j$, we have

$$= \Sigma_{v_{j+1} \in S_{j+1}} Pr[(F(X_j) \in S_j \mid F(X_{j\downarrow}) = (v_{j+1} - v^*_{j+1})$$
$$, X_{t\downarrow(t+1-j)} = z_j] \cdot Pr[F(X_{j+1}) = v_{j+1}]$$

Not finished. Not correct below

Since F is $(\epsilon_\downarrow, \delta_\downarrow, \Delta_\downarrow, \Gamma)$-CW private for all database X, F is $(\epsilon_\downarrow, \delta_\downarrow, (\Delta_\downarrow \cup F(X_\downarrow)), \Gamma)$-CW private, because $F(X_\downarrow)$ does not provide any unknown auxiliary information. By the definition of CW-privacy, F should be $(\epsilon_\downarrow, \delta_\downarrow, \Delta^{'}, \Gamma)$-CW private, for any $\Delta^{'}$ with less auxiliary information than $X_\downarrow \cup F(X_\downarrow)$. Then, F is $(\epsilon_\downarrow, \delta_\downarrow, (\Delta_{\downarrow k} \cup F(X_\downarrow)), \Gamma)$-CW private for any $k \in [1/c]$, where $\Delta_{\downarrow k}$ contains auxiliary information $X_{\downarrow k}$. Let F be $(\epsilon_\downarrow, \delta_\downarrow, (\Delta_{\downarrow k} \cup F(X_\downarrow)), \Gamma)$-CW private with simulator $sim_{F\downarrow k}$.

$$\leq \Sigma_{v_{j+1} \in S_{j+1}} (e^{\epsilon_\downarrow} Pr[sim_{F\downarrow j}(\mathbf{alt}(X_j)) \in S_j \mid F(X_{j\downarrow}) = (v_{j+1} - v^*_{j+1})$$
$$, X_{t\downarrow(t+1-j)} = z_j] + \delta_\downarrow) \cdot Pr[F(X_{j+1}) = v_{j+1}]$$
$$= \Sigma_{v_{j+1} \in S_{j+1}} (e^{\epsilon_\downarrow} Pr[sim_{F\downarrow j}(\mathbf{alt}(X_j)) \in S_j \mid F(X_{j+1}) = v_{j+1}, F(X_{j+2}) \in S_{j+2},$$
$$\ldots, F(X_t) \in S_t, X_{t\downarrow} = z] + \delta_\downarrow) \cdot Pr[F(X_{j+1}) = v_{j+1}]$$
$$= e^{\epsilon_\downarrow} Pr[sim_{F\downarrow j}(\mathbf{alt}(X_j)) \in S_j \mid F(X_{j+1}) \in S_{j+1}, F(X_{j+2}) \in S_{j+2},$$
$$\ldots, F(X_t) \in S_t, X_{t\downarrow} = z] + \delta_\downarrow$$

with simulator $sim_{F\downarrow}$, we know for all database $X \in \mathcal{U}^n$, for all $v \in \mathbb{R}^d$ in the range of $F(X_\downarrow)$ and for all set $S \subseteq \mathbb{R}^d$

$$Pr[F(x_1)+F(X_\downarrow) \in S | X_\downarrow = z, F(X_\downarrow) = v] = Pr[F(x_1)+F(X_\downarrow) \in S | X_\downarrow = z]$$

$$\leq Pr[sim_{F\downarrow}(\mathbf{alt}(X)) \in S | X_\downarrow = z]$$

$$= Pr[(F(x_{11}, \ldots, x_{1t}, X_{1\downarrow(t)}), \ldots, F(x_{(t-1)1}, x_{(t-1)2} X_{(t-1)\downarrow 2}), , F(x_{t1}, X_{t\downarrow})) \in S_t \mid X_{t\downarrow} = z]$$

Notice that block $x_{ij} = x_{i'j'}$ iff $i + j = i' + j'$.

$$= \Pi_{j=1}^t Pr[F(x_{j1}, \ldots, x_{j(t+1-j)}, X_{t\downarrow(t+1-j)}) \mid]$$

Let

$$\mathcal{H} = (F(x_{11}, \ldots, x_{1(t-1)}, X_{1\downarrow(t-1)}), \ldots, F(x_{(t-2)1}, x_{(t-2)2}, X_{(t-1)\downarrow 2}), F(x_{(t-1)1}, X_{(t-1)\downarrow})).$$

Given $X_{t\uparrow} = z$, $\mathcal{H}$ is a function on $(x_{11}, \ldots, x_{1(t-1)})$ and $F(x_{(t-1)1}, X_{(t-1)\downarrow}))$ is a function on $x_{(t-1)1}$. Since $x_{(t-1)1}$ is independent from $x_{11}, \ldots, x_{1(t-1)}$, we have

$$= Pr[\mathcal{H}(x_{11}, \ldots, x_{1(t-1)}) \in S_{-t} | X_{t\uparrow} = z] \cdot Pr[F(X_{t\uparrow}, x_{t\frac{1}{c}}) \in S_t \mid X_{t\uparrow} = z]$$

Proof (2). For any set $S = (S_1, \ldots, S_{t-1}, S_t)$, we have

$$Pr[G(X) \in S \mid X_{t\uparrow} = z]$$

$$= Pr[(F(X_1, \ldots, F(X_{t-1})) \in (S_1, \ldots, S_{t-1}), F(X_t) \in S_t \mid X_{t\uparrow} = z]$$

For ease of analysis, assume $t - 1 < n$. That is, X_1 and X_t still have at least one overlapped row.

$$= Pr[(F(x_{11}, \ldots, x_{1(t-1)}, X_{1\downarrow(t-1)}), \ldots, F(x_{(t-2)1}, x_{(t-2)2}, X_{(t-1)\downarrow 2}), F(x_{(t-1)1}, X_{(t-1)\downarrow})) \in S_{-t}$$

$$, F(X_{t\uparrow}, x_{t\frac{1}{c}}) \in S_t \mid X_{t\uparrow} = z]$$

Notice that block $x_{ij} = x_{i'j'}$ iff $i + j = i' + j'$. Let

$$\mathcal{H} = (F(x_{11}, \ldots, x_{1(t-1)}, X_{1\downarrow(t-1)}), \ldots, F(x_{(t-2)1}, x_{(t-2)2}, X_{(t-1)\downarrow 2}), F(x_{(t-1)1}, X_{(t-1)\downarrow})).$$

Given $X_{t\uparrow} = z$, $\mathcal{H}$ is a function on $(x_{11}, \ldots, x_{1(t-1)})$ and $F(x_{(t-1)1}, X_{(t-1)\downarrow}))$ is a function on $x_{(t-1)1}$. Since $x_{(t-1)1}$ is independent from $x_{11}, \ldots, x_{1(t-1)}$, we have

$$= Pr[\mathcal{H}(x_{11}, \ldots, x_{1(t-1)}) \in S_{-t} | X_{t\uparrow} = z] \cdot Pr[F(X_{t\uparrow}, x_{t\frac{1}{c}}) \in S_t \mid X_{t\uparrow} = z]$$

Lemma. For $t = 1, \ldots, n$ and all sets $S_{-t} = (S_1, \ldots, S_{t-1})$,

$$Pr[\mathcal{H}(x_{11}, \ldots, x_{1(t-1)}) \in S_{-t}| \, X_{t\uparrow} = z]$$

$$\leq e^{\downarrow} Pr[(sim_{F\downarrow}(\mathbf{alt}(X_1)), \ldots, sim_{F\downarrow}(\mathbf{alt}(X_{t-1}))) \in S_{-t} \,|\, X_{t\uparrow} = z] + \delta_{\downarrow}.$$

Proof. We prove this lemma inductively. The base case when $t = 1$ is already shown above.

Assume for all sets $S_{-(t-1)} = (S_1, \ldots, S_{t-2})$,

$$Pr[\mathcal{H}(x_{11}, \ldots, x_{1(t-2)}) \in S_{-(t-1)}| \, X_{(t-1)\uparrow} = z']$$

$$\leq e^{\downarrow} Pr[(sim_{F\downarrow}(\mathbf{alt}(X_1)), \ldots, sim_{F\downarrow}(\mathbf{alt}(X_{t-2}))) \in S_{-(t-1)} \,|\, X_{(t-1)\uparrow} = z'] + \delta_{\downarrow}.$$

We have

$$Pr[\mathcal{H}(x_{11}, \ldots, x_{1(t-1)}) \in S_{-t}| \, X_{t\uparrow} = z]$$

$$= \Sigma_{v \in S_{t-1}} Pr[\mathcal{H}(x_{11}, \ldots, x_{1(t-2)}) \in S_{-(t-1)}(v)| \, x_{1(t-1)} = v, X_{t\uparrow} = z]$$

$$\cdot Pr[x_{1(t-1)} = v \,|\, X_{t\uparrow} = z]$$

Notice that $X_{t-1} = [x_{(t-1)1}, X_{t\uparrow}]^\top = [x_{1(t-1)}, X_{t\uparrow}]^\top = [X_{(t-1)\uparrow}, x_{(t-1)\frac{1}{c}}]^\top$. Let $[X_{(t-1)\uparrow}, x_{(t-1)\frac{1}{c}}]^\top = [z', v']^\top$, we have

$$= \Sigma_{v \in S_{t-1}} Pr[\mathcal{H}(x_{11}, \ldots, x_{1(t-2)}) \in S_{-(t-1)}(v)| \, X_{(t-1)\uparrow} = z', x_{(t-1)\frac{1}{c}} = v']$$

$$\cdot Pr[x_{1(t-1)} = v \,|\, X_{t\uparrow} = z]$$

Since $\mathcal{H}(x_{11}, \ldots, x_{1(t-2)})$ is independent with $x_{(t-1)\frac{1}{c}}$, we have

$$= \Sigma_{v \in S_{t-1}} Pr[\mathcal{H}(x_{11}, \ldots, x_{1(t-2)}) \in S_{-(t-1)}(v) \,|\, X_{(t-1)\uparrow} = z']$$

$$\cdot Pr[x_{1(t-1)} = v \,|\, X_{t\uparrow} = z]$$

By the inductive assumption,

$$\leq \Sigma_{v \in S_{t-1}} (e^{\epsilon_\downarrow} Pr[(sim_{F\downarrow}(\mathbf{alt}(X_1)), \ldots, sim_{F\downarrow}(\mathbf{alt}(X_{t-2}))) \in S_{-(t-1)}(v) \,|\, X_{(t-1)\uparrow} = z'] + \delta_{\downarrow})$$

$$\cdot Pr[x_{1(t-1)} = v \,|\, X_{t\uparrow} = z]$$

Again, since none of $sim_{F\downarrow}(\mathbf{alt}(X_i))$ depends on $x_{(t-1)\frac{1}{c}}$, the condition on $X_{(t-1)\uparrow} = z'$ can be substituted with $[X_{(t-1)\uparrow} = z', x_{(t-1)\frac{1}{c}} = v']^\top = [x_{1(t-1)} = v, X_{(t)\uparrow} = z]^\top$.

$$= e^{\epsilon_\downarrow} (\Sigma_{v \in S_{t-1}} Pr[(sim_{F\downarrow}(\mathbf{alt}(X_1)), \ldots, sim_{F\downarrow}(\mathbf{alt}(X_{t-2}))) \in S_{-(t-1)}(v) \,|\, x_{1(t-1)} = v, X_{(t)\uparrow} = z]$$

$$\cdot Pr[x_{1(t-1)} = v \,|\, X_{t\uparrow} = z]) + \delta_{\downarrow}.$$