APPENDIX

A. Proof of Lemma 1

Let us condition on $U_{[n]} = u_{[n]}$, then

$$
\begin{aligned}
\mathbb{E}_{W,\tilde{Z}_{[n+m]}|u_{[n]}}\left[\mathcal{E}(W,\tilde{Z}_{[n+m]},u_{[n]})\right] &= \mathbb{E}_{W,\tilde{Z}_{u_{[n]}},\tilde{Z}_{\bar{u}_{[m]}}}\left[\frac{1}{m}\sum_{i\in\bar{u}_{[m]}}\ell(W,\tilde{Z}_i)-\frac{1}{n}\sum_{i\in u_{[n]}}\ell(W,\tilde{Z}_i)\right] \\
&= \mathbb{E}_{W,\tilde{Z}_{\bar{u}_{[m]}}}\left[\frac{1}{m}\sum_{i\in\bar{u}_{[m]}}\ell(W,\tilde{Z}_i)\right]-\mathbb{E}_{W,\tilde{Z}_{u_{[n]}}}\left[\frac{1}{n}\sum_{i\in u_{[n]}}\ell(W,\tilde{Z}_i)\right] \\
&= \frac{1}{m}\sum_{i\in\bar{u}_{[m]}}\mathbb{E}_W\mathbb{E}_{\tilde{Z}_{\bar{u}_i}}\left[\ell(W,\tilde{Z}_i)\right]-\mathbb{E}_{W,\tilde{Z}_{u_{[n]}}}\left[\frac{1}{n}\sum_{i\in u_{[n]}}\ell(W,\tilde{Z}_i)\right] \\
&= \mathbb{E}_W\mathbb{E}_Z\left[\ell(W,Z)\right]-\mathbb{E}_{W,Z_{[n]}}\left[\frac{1}{n}\sum_{i=1}^{n}\ell(W,Z_i)\right] \\
&= \mathrm{gen}(\mathcal{A})
\end{aligned}
\tag{29}
$$

Taking expectation over $u_{[n]}$ and we get the result of the lemma.

B. Proof of Theorem 6

By the tower property of expectation, we rewrite (14) as

$$
\mathbb{E}_{W,\tilde{Z}_{[n+m]},U_{[n]}}\left[\mathcal{E}\left(W,\tilde{Z}_{[n+m]},U_{[n]}\right)\right] = \mathbb{E}_{P_{W,\tilde{Z}_{[n+m]}}}\mathbb{E}_{P_{U_{[n]}|W,\tilde{Z}_{[n+m]}}}\left[\mathcal{E}\left(W,\tilde{Z}_{[n+m]},U_{[n]}\right)\right] \tag{30}
$$

Then, by the Donsker-Varadhan variational representation of KL divergence, the change of measure is performed for $\lambda\mathcal{E}(w,\tilde{z}_{[n+m]},U_{[n]})$, from the distribution $P_{U_{[n]}|w,\tilde{z}_{[n+m]}}$ to a prior distribution $Q_{U_{[n]}}$, such that

$$
\begin{aligned}
\mathbb{E}_{P_{U_{[n]}|w,\tilde{z}_{[n+m]}}}\left[\lambda\mathcal{E}\left(w,\tilde{z}_{[n+m]},U_{[n]}\right)\right] &\le D\left(P_{U_{[n]}|w,\tilde{z}_{[n+m]}}\|Q_{U_{[n]}}\right)+\mathbb{E}_{Q_{U_{[n]}}}\left[\exp\left(\lambda\mathcal{E}\left(w,\tilde{z}_{[n+m]},U_{[n]}\right)\right)\right] \\
&= D\left(P_{U_{[n]}|w,\tilde{z}_{[n+m]}}\|Q_{U_{[n]}}\right)+\psi_{\mathcal{E}|w,\tilde{z}_{[n+m]},Q_{U_{[n]}}}(\lambda)+\mathbb{E}_{Q_{U_{[n]}}}\left[\lambda\mathcal{E}\left(w,\tilde{z}_{[n+m]},U_{[n]}\right)\right] \\
&= D\left(P_{U_{[n]}|w,\tilde{z}_{[n+m]}}\|Q_{U_{[n]}}\right)+\psi_{\mathcal{E}|w,\tilde{z}_{[n+m]},Q_{U_{[n]}}}(\lambda),
\end{aligned}
\tag{31}
$$

where the last equation holds due to the assumption that $\mathbb{E}_{Q_{U_{[n]}}}\left[\lambda\mathcal{E}\left(w,\tilde{z}_{[n+m]},U_{[n]}\right)\right]=0$ for any w and $\tilde{z}_{[n+m]}$. The RHS is a function of the free parameter $\lambda>0$, which can be optimized to obtain the tightest bound. After taking expectation on $(w,\tilde{z}_{[n+m]})$, we finally get

$$
\mathbb{E}_{W,\tilde{Z}_{[n+m]},U_{[n]}}\left[\mathcal{E}\left(W,\tilde{Z}_{[n+m]},U_{[n]}\right)\right] \le \mathbb{E}_{P_{W,\tilde{Z}_{[n+m]}}}\left[\inf_{\lambda>0}\frac{D\left(P_{U_{[n]}|W,\tilde{Z}_{[n+m]}}\|Q_{U_{[n]}}\right)+\psi_{\mathcal{E}|W,\tilde{Z}_{[n+m]},Q_{U_{[n]}}}(\lambda)}{\lambda}\right]. \tag{32}
$$

C. Proof of Theorem 7

We begin by stating two essential lemmas.

Lemma 3. *For all $\tilde{z}_{[n+m]}\in\mathcal{Z}^{n+m}$ and $w\in\mathcal{W}$, we have*

$$
\mathbb{E}_{U_{[n]}}\left[\mathcal{E}\left(w,\tilde{z}_{[n+m]},U_{[n]}\right)\right]=0. \tag{33}
$$

Proof.

$$
\begin{aligned}
&\mathbb{E}_{U_{[n]}}\left[\mathcal{E}\left(w,\tilde{z}_{[n+m]},U_{[n]}\right)\right] \\
&= \frac{1}{\binom{n+m}{n}}\sum_{u_{[n]}}\left(\frac{1}{m}\sum_{i\in\bar{u}_{[m]}}\ell(w,z_i)-\frac{1}{n}\sum_{i\in u_{[n]}}\ell(w,z_i)\right) \\
&= \frac{1}{\binom{n+m}{n}}\sum_{i=1}^{n+m}\left(\frac{1}{m}\binom{n+m-1}{n}\ell(w,z_i)-\frac{1}{n}\binom{n+m-1}{n-1}\ell(w,z_i)\right) \\
&= 0.
\end{aligned}
\tag{34}
$$

$\square$

Lemma 4. *If $\ell(w,z) \in [0,1]$ for all $\lambda > 0$, $w \in \mathcal{W}$, $z \in \mathcal{Z}$, then*

$$\log \mathbb{E}_{U_{[n]}}\left[\exp\left(\lambda \mathcal{E}\left(W, \tilde{Z}_{[n+m]}, U_{[n]}\right)\right)\right] \leq \frac{\lambda^2 C_{n,m}(n+m)}{8nm}. \tag{35}$$

Proof. We start by constructing a martingale difference sequence as

$$D_i = \mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)|U_1, U_2, \ldots, U_i\right] - \mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)|U_1, U_2, \ldots, U_{i-1}\right], \quad \text{for } i = 1, 2, \ldots, n. \tag{36}$$

Then we define

$$A_i := \inf_u \mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)|U_1, U_2, \ldots, U_{i-1}, U_i = u\right] - \mathbb{E}\left[\mathcal{E}(w, \tilde{z}_{[n+m]}, U_{[n]})|U_1, U_2, \ldots, U_{i-1}\right], \tag{37}$$

$$B_i := \sup_u \mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)|U_1, U_2, \ldots, U_{i-1}, U_i = u\right] - \mathbb{E}\left[\mathcal{E}(w, \tilde{z}_{[n+m]}, U_{[n]})|U_1, U_2, \ldots, U_{i-1}\right]. \tag{38}$$

Observe that $A_i \leq D_i \leq B_i$ and define $\Delta_i := B_i - A_i$, then we have

$$\begin{aligned}
\Delta_i =& \sup_u \mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)|U_1, U_2, \ldots, U_{i-1}, U_i = u\right] - \inf_u \mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)|U_1, U_2, \ldots, U_{i-1}, U_i = u\right] \\
=& \sup_{u,u'} \left(\mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)|U_1, U_2, \ldots, U_{i-1}, U_i = u\right] - \mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)|U_1, U_2, \ldots, U_{i-1}, U_i = u'\right]\right) \\
=& \frac{1}{n}\left(\ell(w, z_{u'}) - \ell(w, z_u)\right) + \frac{1}{\binom{n+m-i}{n-i}}\left(\frac{1}{m}\binom{n+m-i-1}{n-i}\ell(w, z_{u'}) - \frac{1}{n}\binom{n+m-i-1}{n-i-1}\ell(w, z_{u'})\right) \\
&- \frac{1}{\binom{n+m-i}{n-i}}\left(\frac{1}{m}\binom{n+m-i-1}{n-i}\ell(w, z_u) - \frac{1}{n}\binom{n+m-i-1}{n-i-1}\ell(w, z_u)\right) \\
=& \left(\frac{1}{n} + \frac{1}{m}\cdot\frac{m}{n+m-i} - \frac{1}{n}\cdot\frac{n-i}{n+m-i}\right)\left(\ell(w, z_{u'}) - \ell(w, z_u)\right) \\
=& \frac{n+m}{n(n+m-i)}\left(\ell(w, z_{u'}) - \ell(w, z_u)\right) \\
\leq& \frac{n+m}{n(n+m-i)}
\end{aligned} \tag{39}$$

By iteration, we decompose $\mathbb{E}\left[\exp\left(\lambda \mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)\right)\right]$ and apply Hoeffding's lemma to obtain

$$\begin{aligned}
\log \mathbb{E}\left[\exp\left(\lambda \mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)\right)\right] =& \log \mathbb{E}\left[\exp\left(\lambda\left(\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right) - \mathbb{E}\left[\mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)\right]\right)\right)\right] && \text{(Lemma 3)} \\
=& \log \mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n} D_i\right)\right] \\
=& \log \mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n-1} D_i\right)\mathbb{E}\left[\exp\left(\lambda D_n\right)|U_1, U_2, \ldots, U_{n-1}\right]\right] && \text{(tower expectation)} \\
\leq& \log \mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n-1} D_i\right)\right] + \frac{\lambda^2 \Delta_i^2}{8} && \text{(Hoeffding's lemma)} \\
&\vdots \\
\leq& \frac{\lambda^2 (n+m)^2}{8n^2} \sum_{i=1}^{n} \frac{1}{(n+m-i)^2} \\
=& \frac{\lambda^2 (n+m)^2}{8n^2} \sum_{i=m}^{n+m-1} \frac{1}{i^2}.
\end{aligned} \tag{40}$$

Now consider an equivalent setup where we use $U_{[m]}$ to pick the test set, then we can also obtain a upper bound that resembles (40). This thus gives a tightened upper bound as

$$\log \mathbb{E}\left[\exp\left(\lambda \mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)\right)\right] \leq \min\left(\frac{\lambda^2 (n+m)^2}{8n^2} \sum_{i=m}^{n+m-1} \frac{1}{i^2}, \frac{\lambda^2 (n+m)^2}{8m^2} \sum_{i=n}^{n+m-1} \frac{1}{i^2}\right). \tag{41}$$

Note that when $m=1$ or $n=1$, the RHS of (41) can be directly computed as

$$\log \mathbb{E}\left[\exp\left(\lambda \mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)\right)\right] \leq \frac{\lambda^2(n+m)^2}{8\max(n^2, m^2)}. \tag{42}$$

Otherwise, (41) is further upper bounded by

$$\begin{aligned}
\log \mathbb{E}\left[\exp\left(\lambda \mathcal{E}\left(w, \tilde{z}_{[n+m]}, U_{[n]}\right)\right)\right] \leq & \frac{\lambda^2(n+m)^2}{8} \min\left(\frac{1}{n^2}\sum_{i=m}^{n+m-1}\frac{1}{i(i-1)}, \frac{1}{m^2}\sum_{i=n}^{n+m-1}\frac{1}{i(i-1)}\right) \\
\leq & \frac{\lambda^2(n+m)^2}{8} \min\left(\frac{1}{n^2}\left(\frac{1}{m-1}-\frac{1}{n+m-1}\right), \frac{1}{m^2}\left(\frac{1}{n-1}-\frac{1}{n+m-1}\right)\right) \\
= & \frac{\lambda^2(n+m)^2}{8(n+m-1)} \cdot \frac{1}{nm-\min(n,m)}
\end{aligned} \tag{43}$$

Combining (42) and (43) together completes the proof. □

Now, by the Donsker–Varadhan variational representation of KL divergence, for any $\lambda \in \mathbb{R}$ we have

$$\begin{aligned}
&\mathbb{E}_{W,U_{[n]}|\tilde{z}_{[n+m]}}\left[\lambda \mathcal{E}(W, \tilde{z}_{[n+m]}, U_{[n]})\right] \\
&\leq I^{\tilde{z}_{[n+m]}}\left(W; U_{[n]}\right) + \log \mathbb{E}_{W|\tilde{z}_{[n+m]}}\mathbb{E}_{U_{[n]}|\tilde{z}_{[n+m]}}\left[\exp\left(\lambda \mathcal{E}\left(W, \tilde{z}_{[n+m]}, U_{[n]}\right)\right)\right] && \text{(Donsker-Varadhan)} \\
&= I^{\tilde{z}_{[n+m]}}\left(W; U_{[n]}\right) + \log \mathbb{E}_{W|\tilde{z}_{[n+m]}}\mathbb{E}_{U_{[n]}}\left[\exp\left(\lambda \mathcal{E}\left(W, \tilde{z}_{[n+m]}, U_{[n]}\right)\right)\right] && (U_{[n]} \perp\!\!\!\perp \tilde{Z}_{[n+m]}) \\
&\leq I^{\tilde{z}_{[n+m]}}\left(W; U_{[n]}\right) + \frac{\lambda^2 C_{n,m}(n+m)}{8mn}. && \text{(Lemma 4)}
\end{aligned} \tag{44}$$

After optimizing λ and taking expectations over $\tilde{Z}_{[n+m]}$ on both sides, we obtain

$$\begin{aligned}
\mathbb{E}_{W,U_{[n]},\tilde{Z}_{[n+m]}}\left[\mathcal{E}\left(W, \tilde{Z}_{[n+m]}, U_{[n]}\right)\right] &\leq \mathbb{E}_{\tilde{Z}_{[n+m]}}\left[\inf_{\lambda>0}\frac{1}{\lambda} I^{\tilde{Z}_{[n+m]}}\left(W; U_{[n]}\right) + \frac{\lambda C_{n,m}(n+m)}{8mn}\right] \\
&= \mathbb{E}_{\tilde{Z}_{[n+m]}}\left[\sqrt{\frac{C_{n,m}(n+m)}{2mn} I^{\tilde{Z}_{[n+m]}}\left(W; U_{[n]}\right)}\right] \\
&\leq \sqrt{\frac{C_{n,m}(n+m)}{2mn} I\left(W; U_{[n]}|\tilde{Z}_{[n+m]}\right)}. && \text{(Jensen's inequality)}
\end{aligned} \tag{45}$$

This concludes the proof.

D. Proof of Theorem 8

The corresponding expected generalization error with respect to (18) is

$$\begin{aligned}
\mathbb{E}_{W,\tilde{Z}_{[n+m]},U_{[n]}}\left[\lambda \mathcal{E}\left(W, \tilde{Z}_{[n+m]}, U_{[n]}\right)\right] =& \frac{1}{k}\sum_{i=1}^{k}\mathbb{E}_{W,\tilde{Z}_{[n+m]},U_{[n]}}\left[\lambda \mathcal{E}_i\left(w, \tilde{z}^{(i)}_{[\frac{n+m}{k}]}, u^{(i)}_{[\frac{n}{k}]}\right)\right] \\
=& \frac{1}{k}\sum_{i=1}^{k}\mathbb{E}_{W,\tilde{Z}_{[n+m]},U^{(i)}_{[\frac{n}{k}]}}\left[\lambda \mathcal{E}_i\left(w, \tilde{z}^{(i)}_{[\frac{n+m}{k}]}, u^{(i)}_{[\frac{n}{k}]}\right)\right].
\end{aligned} \tag{46}$$

Then the rest of the arguments closely follows the proof of Theorem 6, with the difference that we perform the change of measure method for $\mathcal{E}_i\left(w, \tilde{z}^{(i)}_{[\frac{n+m}{k}]}, u^{(i)}_{[\frac{n}{k}]}\right)$, from $P_{U^{(i)}_{[\frac{n}{k}]}|w,\tilde{z}_{[n+m]}}$ to $Q_{U^{(i)}_{[\frac{n}{k}]}}$.

E. Proof of Theorem 9

Again, by using the Donsker-Varadhan inequality, we upper bound each summand of the above expression, i.e.,

$$\mathbb{E}_{W,U^{(i)}_{[\frac{n}{k}]}|\tilde{z}_{[n+m]}}\left[\lambda \mathcal{E}_i\left(w, \tilde{z}^{(i)}_{[\frac{n+m}{k}]}, u^{(i)}_{[\frac{n}{k}]}\right)\right] \leq I^{\tilde{z}_{[n+m]}}\left(W; U^{(i)}_{[\frac{n}{k}]}\right) + \log \mathbb{E}_{W|\tilde{z}_{[n+m]}}\mathbb{E}_{U^{(i)}_{[\frac{n}{k}]}}\left[\exp\left(\lambda \mathcal{E}_i\left(w, \tilde{z}^{(i)}_{[\frac{n+m}{k}]}, u^{(i)}_{[\frac{n}{k}]}\right)\right)\right]. \tag{47}$$

Notice that Lemma 4 is also applicable to each individual partition, where n and m are scaled to $\frac{n}{k}$ and $\frac{m}{k}$, respectively, we have

$$\begin{aligned}
\log \mathbb{E}_{U^{(i)}_{[\frac{n}{k}]}}\left[\exp\left(\lambda \mathcal{E}_i\left(w, \tilde{z}^{(i)}_{[\frac{n+m}{k}]}, u^{(i)}_{[\frac{n}{k}]}\right)\right)\right] &\leq \min\left(\frac{\lambda^2(n+m)^2}{8n^2}\sum_{i=\frac{m}{k}}^{\frac{n+m}{k}-1}\frac{1}{i^2}, \frac{\lambda^2(n+m)^2}{8m^2}\sum_{i=\frac{n}{k}}^{\frac{n+m}{k}-1}\frac{1}{i^2}\right) \\
&= \frac{\lambda^2 C^k_{n,m} k(n+m)}{8mn}.
\end{aligned} \tag{48}$$

The result now follows by optimizing λ and averaging over $\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}$, as done in (45).

F. Proof of Theorem 10

Starting from (46), by marginalizing out partitions that do not appear in each summand of (46), we further simplify the expected generalization error as

$$\mathbb{E}_{W,\tilde{Z}_{[n+m]},U_{[n]}}\left[\lambda\mathcal{E}\left(W,\tilde{Z}_{[n+m]},U_{[n]}\right)\right]=\frac{1}{k}\sum_{i=1}^{k}\mathbb{E}_{W,\tilde{Z}^{(i)}_{[\frac{n+m}{k}]},U^{(i)}_{[\frac{n}{k}]}}\left[\lambda\mathcal{E}_i\left(w,\tilde{z}^{(i)}_{[\frac{n+m}{k}]},u^{(i)}_{[\frac{n}{k}]}\right)\right]. \tag{49}$$

Using the Donsker-Varadhan inequality gives

$$\begin{aligned}\mathbb{E}_{W,U^{(i)}_{[\frac{n}{k}]}|\tilde{z}^{(i)}_{[\frac{n+m}{k}]}}\left[\lambda\mathcal{E}_i\left(w,z^{(i)}_{[\frac{n+m}{k}]},u^{(i)}_{[\frac{n}{k}]}\right)\right]&\le I^{\tilde{z}^{(i)}_{[\frac{n+m}{k}]}}\left(W;U^{(i)}_{[\frac{n}{k}]}\right)+\log\mathbb{E}_{W|\tilde{z}^{(i)}_{[\frac{n+m}{k}]}}\mathbb{E}_{U^{(i)}_{[\frac{n}{k}]}}\left[\exp\left(\lambda\mathcal{E}_i\left(w,z^{(i)}_{[\frac{n+m}{k}]},u^{(i)}_{[\frac{n}{k}]}\right)\right)\right]\\&=I^{\tilde{z}^{(i)}_{[\frac{n+m}{k}]}}\left(W;U^{(i)}_{[\frac{n}{k}]}\right)+\frac{\lambda^2C^k_{n,m}k(n+m)}{8mn}.\end{aligned} \tag{50}$$

The rest of the argument also follows from (45).

G. Proof of Lemma 2

For brevity, we omit the subscript of each set (e.g., $U^{(i)}$ for $U^{(i)}_{[\frac{n}{k}]}$) as the set size is clear from the context. To prove $I\left(W;U^{(i)}|\tilde{Z}^{(i)}\right)\le I\left(W;U^{(i)}|\tilde{Z}\right)$, we let $\tilde{Z}^{\backslash(i)}=\tilde{Z}\backslash\tilde{Z}^{(i)}$, then

$$\begin{aligned}&I\left(W;U^{(i)}|\tilde{Z}\right)-I\left(W;U^{(i)}|\tilde{Z}^{(i)}\right)\\&=I\left(W;U^{(i)}|\tilde{Z}^{(i)},\tilde{Z}^{\backslash(i)}\right)-I\left(W;U^{(i)}|\tilde{Z}^{(i)}\right)\\&=I\left(W;U^{(i)}|\tilde{Z}^{(i)},\tilde{Z}^{\backslash(i)}\right)+I\left(U^{(i)};\tilde{Z}^{\backslash(i)}|\tilde{Z}^{(i)}\right)\\&\quad-I\left(W;U^{(i)}|\tilde{Z}^{(i)}\right)\\&=I\left(W,\tilde{Z}^{\backslash(i)};U^{(i)}|\tilde{Z}^{(i)}\right)-I\left(W;U^{(i)}|\tilde{Z}^{(i)}\right)\\&=I\left(U^{(i)};\tilde{Z}^{\backslash(i)}|\tilde{Z}^{(i)},W\right)\\&\ge 0\end{aligned} \tag{51}$$

H. Proof of Theorem 11

Lemma 5. *When m tends to infinity and k,n are constants, we have* $\lim_{m\to\infty}I\left(W;\tilde{Z}_{[n+m]}\right)=\lim_{m\to\infty}I\left(W;\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right)=0.$

Proof. Using the tower property of conditional expectation and the fact that $U_{[n]}\perp\!\!\!\perp\tilde{Z}_{[n+m]}$, we get

$$\begin{aligned}P_{W|\tilde{Z}_{[n+m]}}&=\mathbb{E}_{U_{[n]}}\left[P_{W|\tilde{Z}_{[n+m]},U_{[n]}}\right]\\&=\frac{1}{\binom{\frac{n+m}{k}}{\frac{n}{k}}^k}\sum_{u_{[n]}}P_{W|\tilde{Z}_{[n+m]},U_{[n]}}(w|\tilde{z}_{[n+m]},u_{[n]})\\&=\frac{1}{\binom{\frac{n+m}{k}}{\frac{n}{k}}^k}\sum_{u_{[n]}}P_{W|\tilde{Z}_{U_{[n]}}}(w|\tilde{z}_{u_{[n]}}).\end{aligned} \tag{52}$$

Let $g_w:\mathcal{Z}^{n+m}\to[0,1]$ be a function defined as $g_w(\tilde{Z}_{[n+m]}):=P_{W=w|\tilde{Z}_{[n+m]}}$. Given a supersample set $\tilde{z}_{[n+m]}\in\mathcal{Z}^{n+m}$, let $\tilde{z}'_{[n+m]}$ equal $\tilde{z}_{[n+m]}$ for all instances except the i-th, i.e., $\tilde{z}'_{[n+m]}=\tilde{z}_{[n+m]}\backslash\{\tilde{z}_i\}\cup\{\tilde{z}'_i\}$, where $\tilde{z}'_i$ is an independent copy of $\tilde{z}_i$. For each summand in (52), it is shown that $P_{W|\tilde{Z}_{U_{[n]}}}(w|\tilde{z}_{u_{[n]}})$ and $P_{W|\tilde{Z}_{U_{[n]}}}(w|\tilde{z}'_{u_{[n]}})$ only differ when the index i is included by $u_{[n]}$. Thus, the absolute difference between $g_w(\tilde{z}_{[n+m]})$ and $g_w(\tilde{z}'_{[n+m]})$ can be bounds as:

$$\begin{aligned}
\left|g_w(\tilde{z}_{[n+m]}) - g_w(\tilde{z}'_{[n+m]})\right| &= \left| \frac{1}{\binom{\frac{n+m}{k}}{\frac{n}{k}}^k} \sum_{u_{[n]}:\ i \in u_{[n]}} P_{W|\tilde{Z}_{U_{[n]}}}(w|\tilde{z}_{u_{[n]}}) - P_{W|\tilde{Z}_{U_{[n]}}}(w|\tilde{z}'_{u_{[n]}}) \right| \\
&\leq \frac{1}{\binom{\frac{n+m}{k}}{\frac{n}{k}}^k} \sum_{u_{[n]}:\ i \in u_{[n]}} \left| P_{W|\tilde{Z}_{U_{[n]}}}(w|\tilde{z}_{u_{[n]}}) - P_{W|\tilde{Z}_{U_{[n]}}}(w|\tilde{z}'_{u_{[n]}}) \right| \\
&\leq \frac{1}{\binom{\frac{n+m}{k}}{\frac{n}{k}}^k} \cdot \binom{\frac{n+m}{k}}{\frac{n}{k}}^{k-1} \cdot \binom{\frac{n+m}{k}-1}{\frac{n}{k}-1} \\
&= \frac{n}{n+m}.
\end{aligned} \tag{53}$$

With the difference property in (53), applying McDiarmid's inequality gives

$$\mathbb{P}\left\{ \left| g_w(\tilde{Z}_{[n+m]}) - \mathbb{E}\left[g_w(\tilde{Z}_{[n+m]}) \right] \right| \geq \epsilon \right\} \leq \exp\left(-\frac{2(n+m)\epsilon^2}{n^2} \right). \tag{54}$$

Since $\mathbb{E}\left[g_w(\tilde{Z}_{[n+m]})\right] = P_{W=w}$, in probability we have $\lim_{m\to\infty} P_{W=w|\tilde{Z}_{[n+m]}} = P_{W=w}$. Next, by definition of mutual information, we obtain

$$\begin{aligned}
\lim_{m\to\infty} I\left(W;\tilde{Z}_{[n+m]}\right) &= \lim_{m\to\infty} \mathbb{E}\left[D\left(P_{W|\tilde{Z}_{[n+m]}} \| P_W \right) \right] \\
&= \lim_{m\to\infty} \sum_w \mathbb{E}\left[P_{W=w|\tilde{Z}_{[n+m]}} \log \frac{P_{W=w|\tilde{Z}_{[n+m]}}}{P_{W=w}} \right] \\
&= 0.
\end{aligned} \tag{55}$$

By the chain rule of mutual information,

$$I\left(W;\tilde{Z}_{[n+m]}\right) = I\left(W;\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right) + I\left(W;\tilde{Z}^{\backslash(i)}_{[\frac{k-1}{k}(n+m)]} | \tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right), \tag{56}$$

where $\tilde{Z}^{\backslash(i)}_{[\frac{k-1}{k}(n+m)]} = \tilde{Z}_{[n+m]} \backslash \tilde{Z}^{(i)}_{[\frac{n+m}{k}]}$. Due to the non-negativity of conditional mutual information and take $\lim_{m\to\infty}$ on both sides, we finally get

$$\lim_{m\to\infty} I\left(W;\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right) = 0. \tag{57}$$

□

We first provide the proof for the IPCIMI bound. According to the proposed supersample setting, we have

$$\begin{aligned}
I\left(W;\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}, U^{(i)}_{[\frac{n}{k}]}\right) &= I\left(W;\tilde{Z}^{(i)}_{U^{(i)}_{[\frac{n}{k}]}}, \tilde{Z}^{(i)}_{\bar{U}^{(i)}_{[\frac{m}{k}]}}, U^{(i)}_{[\frac{n}{k}]}\right) \\
&= I\left(W;\tilde{Z}^{(i)}_{U^{(i)}_{[\frac{n}{k}]}}\right) + I\left(W;\tilde{Z}^{(i)}_{\bar{U}^{(i)}_{[\frac{m}{k}]}}, U^{(i)}_{[\frac{n}{k}]} | \tilde{Z}^{(i)}_{U^{(i)}_{[\frac{n}{k}]}}\right) \\
&\overset{(a)}{=} I\left(W;\tilde{Z}^{(i)}_{U^{(i)}_{[\frac{n}{k}]}}\right) \\
&\overset{(b)}{=} I\left(W;Z^{(i)}_{[\frac{n}{k}]}\right),
\end{aligned} \tag{58}$$

where (a) is due to the fact that W is independent of other random variables given $\tilde{Z}_{U_{[n]}}$, and (b) is according to the supersample set structure where $Z^{(i)}_{[\frac{n}{k}]}$ and $\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}$ lead to the same training set. Next, using the chain rule of mutual information, one can obtain

$$\begin{aligned}\lim_{m\to\infty} I\left(W;U^{(i)}_{[\frac{n}{k}]}|\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right) &= \lim_{m\to\infty} I\left(W;\tilde{Z}^{(i)}_{[\frac{n+m}{k}]},U^{(i)}_{[\frac{n}{k}]}\right) - I\left(W;\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right)\\ &= I\left(W;Z^{(i)}_{[\frac{n}{k}]}\right) - \lim_{m\to\infty} I\left(W;\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right)\\ &= I\left(W;Z^{(i)}_{[\frac{n}{k}]}\right),\end{aligned} \tag{59}$$

where the last step is by Lemma 5. Finally, it is simple to verify that

$$\begin{aligned}\frac{1}{k}\sum_{i=1}^{k}\lim_{m\to\infty}\sqrt{\frac{C^k_{n,m}k(n+m)}{2nm}I\left(W;U^{(i)}_{[\frac{n}{k}]}|\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right)} &= \frac{1}{k}\sum_{i=1}^{k}\sqrt{\lim_{m\to\infty}\frac{C^k_{n,m}k(n+m)}{2nm}\lim_{m\to\infty} I\left(W;U^{(i)}_{[\frac{n}{k}]}|\tilde{Z}^{(i)}_{[\frac{n+m}{k}]}\right)}\\ &= \frac{1}{k}\sum_{i=1}^{k}\sqrt{\frac{k}{2n}I\left(W;Z^{(i)}_{[\frac{n}{k}]}\right)}.\end{aligned} \tag{60}$$

Specially, when $k=1$ and $m\to\infty$, the $(\infty,1)$-IPCIMI bound reduces to the MI bound as

$$\lim_{m\to\infty}\sqrt{\frac{C^k_{n,m}(n+m)}{2nm}I\left(W;U_{[n]}|\tilde{Z}_{[n+m]}\right)} = \sqrt{\frac{1}{2n}I\left(W;Z_{[n]}\right)}, \tag{61}$$

while the (∞,n)-IPCIMI bound recovers the IMI bound:

$$\frac{1}{n}\sum_{i=1}^{n}\lim_{m\to\infty}\sqrt{\frac{C^k_{n,m}(n+m)}{2m}I\left(W;U^{(i)}_{[1]}|\tilde{Z}^{(i)}_{[\frac{m}{n}+1]}\right)} = \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{2}I\left(W;Z_i\right)}. \tag{62}$$

I. Discussions on LOFO Bounds

In this section, we present more elaborations for the proposed LOFO bounds. Under the LOFO setting, $U\sim \mathrm{Uni}([\frac{n}{m}+1])$ and W is the output of $\mathcal{A}(\tilde{Z}_{[n+m]},U)$, and thus the LOFO-IPCMI bound share the same expression with its CMI version, which is given by

$$\mathrm{gen}(\mathcal{A}) \le \frac{n+m}{n}\sqrt{\frac{1}{2}I\left(W;U|\tilde{Z}_{[n+m]}\right)}. \tag{63}$$

By the chain rule of mutual information, we have

$$\begin{aligned}I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U);U|\tilde{Z}_{[n+m]}\right) &= I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U);\tilde{Z}_{[n+m]},U\right) - I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U);\tilde{Z}_{[n+m]}\right)\\ &= I\left(\mathcal{A}(Z_{[n]});Z_{[n]}\right) - I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U);\tilde{Z}_{[n+m]}\right)\\ &\le I\left(\mathcal{A}(Z_{[n]});Z_{[n]}\right).\end{aligned} \tag{64}$$

We now compare the LOFO-IPCMI bound with the (m,m)-IPCMI bound. Under the (m,m)-IPCMI setting, $U_{[m]} = (U^{(1)},\dots,U^{(m)})$, where $U^{(i)}\sim \mathrm{Uni}([\frac{n}{m}+1])$, and W is replaced by $\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]})$. Then, it holds that

$$\begin{aligned}I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});U^{(i)}|\tilde{Z}_{[n+m]}\right) =& I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});\tilde{Z}_{[n+m]},U^{(i)}\right) - I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});\tilde{Z}_{[n+m]}\right)\\ =& I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});\tilde{Z}_{[n+m]},U_{[m]}\right) - I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});U_{[m]}\backslash U^{(i)}|\tilde{Z}_{[n+m]},U^{(i)}\right)\\ &- I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});\tilde{Z}_{[n+m]}\right)\\ =& I\left(\mathcal{A}(Z_{[n]});Z_{[n]}\right) - I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});U_{[m]}\backslash U^{(i)}|\tilde{Z}_{[n+m]},U^{(i)}\right)\\ &- I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});\tilde{Z}_{[n+m]}\right)\\ \le& I\left(\mathcal{A}(Z_{[n]});Z_{[n]}\right) - I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});U_{[m]}\backslash U^{(i)}|\tilde{Z}_{[n+m]},U^{(i)}\right).\end{aligned} \tag{65}$$

As can be observed, the (m,m)-IPCIMI bound has a tighter upper bound. However, since the relationship between $I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U);\tilde{Z}_{[n+m]}\right)$ and $I\left(\mathcal{A}(\tilde{Z}_{[n+m]},U_{[m]});\tilde{Z}_{[n+m]}\right)$ is not clear, the tightness of LOFO-IPCMI and (m,m)-IPCMI bound can not be compared directly.

Besides, the LOFO bounds benefit from a lower computational complexity. To see this, we consider the computation of the mutual information quantity in the (m, m)-IPCMI bound, where $P_{W|\tilde{Z}_{[n+m]}}$ is computed by

$$P_{W|\tilde{Z}_{[n+m]}}(w|\tilde{z}_{[n+m]}) = \sum_{u_{[m]}} P_{W|U_{[m]},\tilde{Z}_{[n+m]}}(w|u_{[m]}, \tilde{z}_{[n+m]}) P_{U_{[m]}}(u_{[m]}), \tag{66}$$

which needs to iterate over $(\frac{n}{m}+1)^m$ possible values of $U_{[m]}$. In contrast, in the LOFO-IPCMI bound, the domain size of U is $\frac{n}{m}+1$, inducing a significant reduction of computation. This benefit can also be obtained by the proposed LOFO-IPCIMI bound.

J. Calculation Details for the Bernoulli Example

In this section, we provide the calculations details when evaluating the considered bounds. To make it clear from the context, the realizations of capital random variables are represented with the corresponding lower-case letters. Before presenting the calculation details, we introduce some useful lemmas as below.

Lemma 6. *Let X and Y be two independent binomial random variables following $X \sim \mathrm{Bin}(n,p)$ and $Y \sim \mathrm{Bin}(m,p)$, respectively. If $\bar{X} = n - X$, then $\bar{X} \sim \mathrm{Bin}(n, 1-p)$. If $Z = X + Y$, then $Z \sim \mathrm{Bin}(n+m, p)$.*

Lemma 7. *If X is a binomial random variable following $X \sim \mathrm{Bin}(n,p)$ and $a \geq 0$, then*

$$\mathbb{E}\left[\log(X+a)\right] = \log(np+a) - \frac{np(1-p)}{2(np+a)^2 \ln 2} + O\left(\frac{1}{n^2}\right). \tag{67}$$

Proof. We apply Taylor series to the function $\log(x+a)$ at point $x_0 = np$, which is

$$T_{x_0=np}\left[\log(x+a)\right] = \log(np+a) + \sum_{i=1}^{\infty} c_i (x-np)^i, \tag{68}$$

where $c_i = \frac{(-1)^{i-1}}{i(np+a)^i \ln 2}$. The convergent radius around x_0 is given by

$$r = \lim_{i\to\infty} \left|\frac{c_{i+1}}{c_i}\right| = \lim_{i\to\infty} \left| \frac{\frac{(-1)^{i-1}}{i(np+a)^i \ln 2}}{\frac{(-1)^i}{(i+1)(np+a)^{i+1}\ln 2}} \right| = np + a, \tag{69}$$

which means that (68) only converges to $\log(x+a)$ when $x \in (-a, 2np+a]$. We now check if all realizations of X fall inside the convergence interval. If $p \geq \frac{1}{2}$, then $X \in [0,n] \subset (-a, 2np+a]$ and the convergence condition is satisfied. As such, we can take the expectation of (68) and get

$$\begin{aligned}
\mathbb{E}\left[\log(X+a)\right] &= \log(np+a) + \frac{\mathbb{E}\left[X-np\right]}{(np+a)\ln 2} - \frac{\mathbb{E}\left[(X-np)^2\right]}{2(np+a)^2 \ln 2} + \frac{\mathbb{E}\left[(X-np)^3\right]}{3(np+a)^3 \ln 2} + \dots \\
&= \log(np+a) + 0 - \frac{np(1-p)}{2(np+a)^2\ln 2} + \frac{np(1-p)(1-2p)}{3(np+a)^3 \ln 2} + \dots \\
&= \log(np+a) - \frac{np(1-p)}{2(np+a)^2 \ln 2} + O\left(\frac{1}{n^2}\right).
\end{aligned} \tag{70}$$

In the other case, i.e., $p < \frac{1}{2}$, $\log(x+a)$ can only expand between $x \in [0, 2np+a]$. Using Hoeffding's inequality, we can derive an upper bound for the tail probability that X falls outside the convergence interval, which is shown as follows:

$$\begin{aligned}
\Pr\left[X \geq 2np+a\right] &= \Pr\left[X - \mathbb{E}\left[X\right] \geq np+a\right] \\
&= \Pr\left[\sum_{i=1}^n X_i - \sum_{i=1}^n \mathbb{E}\left[X_i\right] \geq np+a\right] \\
&\leq e^{-2n(p+\frac{a}{n})^2} \\
&\leq e^{-2p^2 n},
\end{aligned} \tag{71}$$

where $X_i \sim \mathrm{Ber}(p)$ is a Bernoulli random variable for any $i \in [1, n]$. Then the expectation is calculated by

$$\begin{aligned}
\mathbb{E}\left[\log(X+a)\right] &= \mathbb{E}\left[\log(X+a)\mathbb{1}_{X<2np+a}\right] + \mathbb{E}\left[\log(X+a)\mathbb{1}_{X\geq 2np+a}\right] \\
&= \mathbb{E}\left[T_{x_0=np}\left[\log(X+a)\right]\mathbb{1}_{X<2np+a}\right] + \mathbb{E}\left[\log(X+a)\mathbb{1}_{X\geq 2np+a}\right] \\
&< \mathbb{E}\left[T_{x_0=np}\left[\log(X+a)\right]\right] + \mathbb{E}\left[\log(n+a)\mathbb{1}_{X\geq 2np+a}\right] \\
&= \log(np+a) - \frac{np(1-p)}{2(np+a)^2\ln 2} + O\left(\frac{1}{n^2}\right) + \log(n+a)\Pr\left[X \geq 2np+a\right] \\
&\leq \log(np+a) - \frac{np(1-p)}{2(np+a)^2\ln 2} + O\left(\frac{1}{n^2}\right) + \log(n+a)e^{-2p^2 n} \\
&= \log(np+a) - \frac{np(1-p)}{2(np+a)^2\ln 2} + O\left(\frac{1}{n^2}\right).
\end{aligned} \tag{72}$$

It is shown that both cases lead to the same result. This thus completes the proof. $\square$

Lemma 8. *If X is a binomial random variable following $X \sim \mathrm{Bin}(n,p)$, then*

$$\mathbb{E}\left[X\log X\right] = np\log np + \frac{1-p}{2\ln 2} + \frac{(1-p^2)}{12np\ln 2} + O\left(\frac{1}{n^2}\right). \tag{73}$$

Proof. Akin to the proof of Lemma 7, we first obtain the Taylor series of the function $x\log x$ at point $x_0 = np$:

$$T_{x_0=np}\left[x\log x\right] = np\log np + (x-np)\log enp + \sum_{i=2}^{\infty}\frac{(-1)^i}{i(i-1)(np)^{i-1}\ln 2}(x-np)^i, \tag{74}$$

where the convergence interval is $x \in (0, 2np]$. If $p \geq \frac{1}{2}$, the convergence condition is satisfied for all realizations of X and the expectation can be performed on the Taylor series to get

$$\begin{aligned}
\mathbb{E}\left[X\log X\right] &= np\log np + \left[X-np\right]\log enp + \frac{\mathbb{E}\left[(X-np)^2\right]}{2np\ln 2} - \frac{\mathbb{E}\left[(X-np)^3\right]}{6(np)^2\ln 2} + \frac{\mathbb{E}\left[(X-np)^4\right]}{12(np)^3\ln 2} + \ldots \\
&= \log(np+a) + 0 + \frac{1-p}{2\ln 2} - \frac{(1-p)(1-2p)}{6np\ln 2} + \frac{(1-p)((3n-6)p(1-p)+1)}{12(np)^2\ln 2} + \ldots \\
&= \log(np+a) + \frac{1-p}{2\ln 2} + \frac{1-p^2}{12np\ln 2} + O\left(\frac{1}{n^2}\right).
\end{aligned} \tag{75}$$

If $p < \frac{1}{2}$, then the series will diverge within the interval $[2np, n]$, the probability of which is given by

$$\begin{aligned}
\Pr\left[X \geq 2np\right] &= \Pr\left[X - \mathbb{E}\left[X\right] \geq np\right] \\
&\leq e^{-2p^2 n}.
\end{aligned} \tag{76}$$

Furthermore, since $x\log x$ is an increasing function on $[2np, n]$, the contributions of X outside $(0, 2np]$ is upper bounded by

$$\mathbb{E}\left[(X\log X)\mathbb{1}_{X\geq 2np}\right] \leq e^{-2p^2 n} n\log n, \tag{77}$$

which is order-wise negligible as compared to $O\left(\frac{1}{n^2}\right)$. Thus, (75) still holds for $p < \frac{1}{2}$, leading to the result of the lemma. $\square$

Lemma 9. *If a, b, c, d are all constants, where $a > 0, b \geq 0, c > 0, d \geq 0$ and $ad \neq bc$, then $\log\frac{an+b}{cn+d}$ scales as $\log\frac{a}{c} + O\left(\frac{1}{n}\right)$.*

Proof.

$$\begin{aligned}
\log\frac{an+b}{cn+d} &= \log\frac{a}{c} + \log\left(1 + \frac{\frac{b}{a} - \frac{d}{c}}{n + \frac{d}{c}}\right) \\
&= \log\frac{a}{c} + \frac{\frac{b}{a} - \frac{d}{c}}{n + \frac{d}{c}} + o\left(\frac{1}{n + \frac{d}{c}}\right) \\
&= \log\frac{a}{c} + O\left(\frac{1}{n}\right)
\end{aligned} \tag{78}$$

$\square$

Lemma 10 (Bounds for a binomial coefficient [17])**.** *If $1 \leq k \leq n-1$, then*

$$\sqrt{\frac{n}{8k(n-k)}}2^{nH\left(\frac{k}{n}\right)} \leq \binom{n}{k} \leq \sqrt{\frac{n}{2\pi k(n-k)}}2^{nH\left(\frac{k}{n}\right)}. \tag{79}$$

Lemma 11. *If X is a binomial random variable following $X \sim \mathrm{Bin}(n,p)$, then*

$$\binom{n}{X} = nH(p) - \frac{1}{2}\log n + O\left(C\right). \tag{80}$$

Proof. First, by Lemma 10, the lower bound of the expected binomial coefficient can be calculated as

$$\begin{aligned}
\mathbb{E}\left[\log\binom{n}{X}\right] \geq& \mathbb{E}\left[\frac{1}{2}\left(\log n - \log X - \log \bar{X} - 3\right) + nH(\frac{X}{n})\right] \\
=& \frac{1}{2}\log n - \frac{3}{2} - \mathbb{E}\left[\frac{1}{2}\left(\log X + \log \bar{X}\right) + X\log\frac{X}{n} + \bar{X}\log\frac{\bar{X}}{n}\right] \\
=& \frac{1}{2}\log n - \frac{3}{2} - \mathbb{E}\left[\frac{1}{2}\left(\log X + \log \bar{X}\right) + X\log X + \bar{X}\log\bar{X} - X\log n - \bar{X}\log n\right] \\
=& \frac{1}{2}\log n - \frac{3}{2} - \left(\frac{1}{2}\left(\log np - \frac{1-p}{2np\ln 2} + \log n(1-p) - \frac{p}{2n(1-p)\ln 2} + O\left(\frac{1}{n^2}\right)\right)\right. \\
& + np\log np + \frac{1-p}{2\ln 2} + \frac{1-p^2}{12np\ln 2} + O\left(\frac{1}{n^2}\right) \\
& + n(1-p)\log n(1-p) + \frac{p}{2\ln 2} + \frac{1-(1-p)^2}{12n(1-p)\ln 2} + O\left(\frac{1}{n^2}\right) \\
& \left. - np\log n - n(1-p)\log n\right) \\
=& -\frac{1}{2}\log np(1-p) - \frac{1}{2\ln 2} - \frac{3}{2} - np\log p - n(1-p)\log(1-p) + O\left(\frac{1}{n}\right) \\
=& nH(p) - \frac{1}{2}\log n - \frac{\log 8ep(1-p)}{2} + O\left(\frac{1}{n}\right).
\end{aligned} \tag{81}$$

Likewise, the upper bound can be obtained as

$$\mathbb{E}\left[\log\binom{n}{X}\right] \leq nH(p) - \frac{1}{2}\log n - \frac{\log 2\pi ep(1-p)}{2} + O\left(\frac{1}{n}\right). \tag{82}$$

Combining the lower and upper bounds completes the proof. $\square$

1) Standard MI Case: We first evaluate the standard MI bound in (5), where the learner is $W = \frac{1}{n}\sum_{i=1}^n Z_i$. Let w and $z_{[n]} = (z_1, \ldots, z_n)$ be the realizations of W and $Z_{[n]}$. Let $x = \sum_{i=1}^n z_i$, where x in an integer from $[0,n]$, then we can obtain the joint, marginal, conditional probabilities as follows:

$$P_{W,Z_{[n]}}(\frac{x}{n}, z_{[n]}) = p^x(1-p)^{n-x}, \tag{83}$$

$$P_W(\frac{x}{n}) = \binom{n}{x}p^x(1-p)^{n-x}, \tag{84}$$

$$P_{W|Z_{[n]}}(\frac{x}{n}|z_{[n]}) = 1. \tag{85}$$

By definition of mutual information, $I(W;Z_{[n]})$ can be calculated as

$$\begin{aligned}
I\left(W;Z_{[n]}\right) &= \sum_{w,z_{[n]}} P_{W,Z_{[n]}}(w,z_{[n]})\log\frac{P_{W|Z_{[n]}}(w|z_{[n]})}{P_W(w)} \\
&= -\sum_{x=0}^{n}\binom{n}{x}p^x(1-p)^{n-x}\log\binom{n}{x}p^x(1-p)^{n-x}.
\end{aligned} \tag{86}$$

Let x be the realization of a Binomial random variable $X \sim \mathrm{Bin}(n,p)$ and let $\bar{X} = n - X \sim \mathrm{Bin}(n, 1-p)$, we can rewritten $I(W;Z_{[n]})$ as

$$I\left(W;Z_{[n]}\right) = -\mathbb{E}\left[\log\binom{n}{X} + X\log p + \bar{X}\log(1-p)\right], \tag{87}$$

plugging which back to the standard MI bound yields

$$\begin{aligned}\sqrt{\frac{1}{2n}I\left(W;Z_{[n]}\right)}=&\sqrt{-\frac{1}{2n}\mathbb{E}\left[\log\binom{n}{X}+X\log p+\bar{X}\log(1-p)\right]}\\=&\sqrt{\frac{1}{2n}\left(-nH(p)+\frac{1}{2}\log n+O\left(C\right)-np\log p-n(1-p)\log(1-p)\right)}\\=&\sqrt{-\frac{1}{2}H(p)+\frac{1}{4}\frac{\log n}{n}+O\left(\frac{1}{n}\right)+\frac{1}{2}H(p)}\\=&\sqrt{\frac{1}{4}\frac{\log n}{n}+O\left(\frac{1}{n}\right)},\end{aligned}\tag{88}$$

where the second step is by Lemma 11. Thus, the MI bound scales as $O\left(\sqrt{\frac{\log n}{n}}\right)$.

2) IMI Case: This subsection evaluates the IMI bound in (7). For the realizations w and z_i, let $x=\sum_{j=1}^{n}z_j-z_i$, where x in an integer from $[0,n-1]$, then we have

$$P_{W,Z_i}(\frac{x+z_i}{n},z_i)=\binom{n-1}{x}p^{x+z_i}(1-p)^{n-x-z_i},\tag{89}$$

$$P_{W}(\frac{x+z_i}{n})=\binom{n}{x+z_i}p^{x+z_i}(1-p)^{n-x-z_i},\tag{90}$$

$$P_{W|Z_i}(\frac{x+z_i}{n}|z_i)=\binom{n-1}{x}p^{x}(1-p)^{n-1-x}.\tag{91}$$

With the above elements obtained, we can further calculate the mutual information according to its definition:

$$\begin{aligned}I\left(W;Z_i\right)=&\sum_{w,z_i}P_{W,Z_i}(w,z_i)\log\frac{P_{W|Z_i}(w|z_i)}{P_W(w)}\\=&\sum_{w}P_{W,Z_i}(w,0)\log\frac{P_{W|Z_i}(w|0)}{P_W(w)}+\sum_{w}P_{W,Z_i}(w,1)\log\frac{P_{W|Z_i}(w|1)}{P_W(w)}\\=&\sum_{x=0}^{n-1}\binom{n-1}{x}p^x(1-p)^{n-x}\log\frac{\binom{n-1}{x}p^x(1-p)^{n-1-x}}{\binom{n}{x}p^x(1-p)^{n-x}}\\&+\sum_{x=0}^{n-1}\binom{n-1}{x}p^{x+1}(1-p)^{n-1-x}\log\frac{\binom{n-1}{x}p^x(1-p)^{n-1-x}}{\binom{n}{x+1}p^{x+1}(1-p)^{n-1-x}}\\=&(1-p)\sum_{x=0}^{n-1}\binom{n-1}{x}p^x(1-p)^{n-1-x}\log\frac{n-x}{n(1-p)}+p\sum_{x=0}^{n-1}\binom{n-1}{x}p^x(1-p)^{n-1-x}\log\frac{x+1}{np}.\end{aligned}\tag{92}$$

Let $X\sim\mathrm{Bin}(n-1,p)$ and $\bar{X}=n-1-X\sim\mathrm{Bin}(n-1,1-p)$, after arrangement and simplification we arrive at

$$I\left(W;Z_i\right)=H(p)-\log n+p\mathbb{E}\left[\log(X+1)\right]+(1-p)\mathbb{E}\left[\log(\bar{X}+1)\right].\tag{93}$$

Next, applying the approximation introduced in Lemma 7 yields

$$\begin{aligned}I\left(W;Z_i\right)=&H(p)-\log n+p\log((n-1)p+1)-\frac{(n-1)(1-p)p}{2\left((n-1)p+1\right)^2\ln 2}\\&+(1-p)\log((n-1)(1-p)+1)-\frac{(n-1)(1-p)p}{2\left((n-1)(1-p)+1\right)^2\ln 2}+O\left(\frac{1}{n^2}\right)\\=&H(p)+p\log\frac{np+1-p}{n}+(1-p)\log\frac{(1-p)n+p}{n}+O\left(\frac{1}{n}\right).\end{aligned}\tag{94}$$

We further apply Lemma 9 to the above expression and obtain

$$\begin{aligned} I\left(W;Z_i\right) =& H(p)+p\log p+(1-p)\log(1-p)+O\left(\frac{1}{n}\right)\\ =& O\left(\frac{1}{n}\right), \end{aligned} \tag{95}$$

plugging which back to (7) implies that the IMI bound is of order $O\left(\frac{1}{\sqrt{n}}\right)$.

3) IPMI Case: We continue to evaluate the IPMI bound in (25), which is a general case of the MI and IMI bound. Under the partitioned setting, the learner $W=\frac{1}{n}\sum_{i=1}^{n}Z_i$ can be rewritten as $W=\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{\frac{n}{k}}Z_j^{(i)}$. For all $i\in[k]$ and $j\in[\frac{n}{k}]$, let w and $z_j^{(i)}$ be the realization of W and $Z_j^{(i)}$, respectively. Furthermore, let $x=\sum_{j=1}^{\frac{n}{k}}z_j^{(i)}$ and $y=nw-x$, with $x\in[0,\frac{n}{k}]$ and $y\in[0,n-\frac{n}{k}]$. Then, one can calculate the following elemental probabilities as

$$P_{W,Z^{(i)}_{[\frac{n}{k}]}}(\frac{x+y}{n},z^{(i)}_{[\frac{n}{k}]})=\binom{n-\frac{n}{k}}{y}p^{x+y}(1-p)^{n-x-y}, \tag{96}$$

$$P_{W}(\frac{x+y}{n})=\binom{n}{x+y}p^{x+y}(1-p)^{n-x-y}, \tag{97}$$

$$P_{W|Z^{(i)}_{[\frac{n}{k}]}}(\frac{x+y}{n}|z^{(i)}_{[\frac{n}{k}]})=\binom{n-\frac{n}{k}}{y}p^{y}(1-p)^{n-\frac{n}{k}-y}, \tag{98}$$

with which $I(W;Z^{(i)}_{[\frac{n}{k}]})$ can be further obtained as

$$\begin{aligned} I\left(W;Z^{(i)}_{[\frac{n}{k}]}\right)=&\sum_{y=0}^{n-\frac{n}{k}}\binom{n-\frac{n}{k}}{y}p^{y}(1-p)^{n-\frac{n}{k}-y}\sum_{x=0}^{\frac{n}{k}}\binom{\frac{n}{k}}{x}p^{x}(1-p)^{\frac{n}{k}-x}\log\frac{\binom{n-\frac{n}{k}}{y}}{\binom{n}{x+y}p^{x}(1-p)^{\frac{n}{k}-x}}\\ =&\mathbb{E}\left[\log\binom{n-\frac{n}{k}}{Y}-\log\binom{n}{X+Y}-X\log p-\bar{X}\log(1-p)\right], \end{aligned} \tag{99}$$

where $X\sim\mathrm{Bin}(\frac{n}{k},p)$, $Y\sim\mathrm{Bin}(n-\frac{n}{k},p)$ and $\bar{X}=\frac{n}{k}-X\sim\mathrm{Bin}(\frac{n}{k},1-p)$.

Using Lemmas 6 and 11, the above mutual information can be approximated as

$$\begin{aligned} I\left(W;Z^{(i)}_{[\frac{n}{k}]}\right)=&(n-\frac{n}{k})H(p)-\frac{1}{2}\log(n-\frac{n}{k})-nH(p)+\frac{1}{2}\log n+O\left(C\right)-\frac{n}{k}p\log p-\frac{n}{k}(1-p)\log(1-p)\\ =&-\frac{n}{k}H(p)+\frac{1}{2}\log\frac{k}{k-1}+O\left(C\right)+\frac{n}{k}H(p)\\ =&\frac{1}{2}\log\frac{k}{k-1}+O\left(C\right). \end{aligned} \tag{100}$$

Since k is a constant, $I(W;Z^{(i)}_{[\frac{n}{k}]})$ is of a constant order $O\left(C\right)$, and it is simple to further verify that the IPMI bound in (25) is of order $O\left(\frac{1}{\sqrt{n}}\right)$.

4) LOO-CMI Case: Under this case, the learner is simplified as $W=\frac{1}{n}(\sum_{i=1}^{n+1}\tilde{Z}_i-\tilde{Z}_U)$, and by Bayes rule we get

$$P_{U|W,\tilde{Z}_{[n+1]}}(u|w,\tilde{z}_{[n+1]})=\frac{P_{W|\tilde{Z}_{[n+1]},U}(w|\tilde{z}_{[n+1]},u)P_{\tilde{Z}_{[n+1]}}(\tilde{z}_{[n+1]})P_U(u)}{\sum_u P_{W,\tilde{Z}_{[n+1]},U}(w,\tilde{z}_{[n+1]},u)}. \tag{101}$$

For $w=\frac{x}{n}$, it is easy to verify that the number of 1's in $\tilde{z}_{[n+1]}$ is either x when $\tilde{z}_u=0$ or $x+1$ when $\tilde{z}_u=1$ since otherwise $P_{W|\tilde{Z}_{[n+1]},U}$ is zero. When $\tilde{z}_u=0$ and $\sum_{i=1}^{n+1}\tilde{z}_i=x$, we have

$$\begin{aligned} P_{U|W,\tilde{Z}_{[n+1]}}(u|\frac{x}{n},\tilde{z}_{[n+1]})=&\frac{\frac{1}{n+1}p^x(1-p)^{n+1-x}}{\frac{n+1-x}{n+1}p^x(1-p)^{n+1-x}}\\ =&\frac{1}{n+1-x}. \end{aligned} \tag{102}$$

Likewise, when $\tilde{z}_u=1$ and $\sum_{i=1}^{n+1}\tilde{z}_i=x+1$, it holds that:

$$P_{U|W,\tilde{Z}_{[n+1]}}(u|\frac{x}{n},\tilde{z}_{[n+1]})=\frac{1}{x+1}. \tag{103}$$

For $w=\frac{x}{n}$, combining (102) and (103) gives

$$P_{U|W,\tilde{Z}_{[n+1]}}(u|\frac{x}{n},\tilde{z}_{[n+1]})=\begin{cases}\frac{1}{n+1-x} & \text{If } \tilde{z}_u=0 \text{ and } \sum_{i=1}^{n+1}\tilde{z}_i=x\\ \frac{1}{x+1} & \text{If } \tilde{z}_u=1 \text{ and } \sum_{i=1}^{n+1}\tilde{z}_i=x+1\\ 0 & \text{Otherwise}\end{cases}. \tag{104}$$

By definition, we compute the mutual information as

$$\begin{aligned}I\left(W;U|\tilde{Z}_{[n+1]}\right)&=\mathbb{E}_{W,\tilde{Z}_{[n+1]},U}\left[\log\frac{P_{U|W,\tilde{Z}_{[n+1]}}}{P_U}\right]\\&=\sum_{x=0}^{n}\frac{n+1-x}{n+1}\binom{n+1}{x}p^x(1-p)^{n+1-x}\log\frac{n+1}{n+1-x}+\sum_{x=0}^{n}\frac{x+1}{n+1}\binom{n+1}{x+1}p^{x+1}(1-p)^{n-x}\log\frac{n+1}{x+1}\\&=\log(n+1)-\sum_{x=0}^{n}\binom{n}{x}p^x(1-p)^{n-x}\left((1-p)\log(n+1-x)+p\log(x+1)\right).\end{aligned} \tag{105}$$

Let X be a binomial random variable $\mathrm{Bin}(n,p)$, then the above mutual information can be rewritten as

$$I\left(W;U|\tilde{Z}_{[n+1]}\right)=\log(n+1)-\mathbb{E}\left[(1-p)\log(n+1-X)+p\log(X+1)\right]. \tag{106}$$

Using the approximations in Lemma 7, we get

$$\begin{aligned}I\left(W;U|\tilde{Z}_{[n+1]}\right)&=\log(n+1)-p\left(\log(np+1)-\frac{np(1-p)}{2(np+1)^2\ln 2}\right)-(1-p)\left(\log(n+1-np)-\frac{np(1-p)}{2(n+1-np)^2\ln 2}\right)\\&=p\log\frac{n+1}{np+1}+(1-p)\log\frac{n+1}{(1-p)n+1}+O\left(\frac{1}{n}\right)\\&=-p\log p-(1-p)\log(1-p)+O\left(\frac{1}{n}\right)\\&=H(p)+O\left(\frac{1}{n}\right),\end{aligned} \tag{107}$$

where we apply Lemma 9 in the third step. Finally, it is shown that the LOO bound converges to a constant $\sqrt{H(p)}$.

5) ICIMI Case: In this subsection, we evaluate the ICIMI bound in (10), where the learner is $W=\sum_{i=1}^{n}\tilde{Z}_{i+R_i n}$. Let $w=\frac{x}{n}$, $r_{[n]}=(r_1,\ldots,r_n)$ and $\tilde{z}_{[2n]}=(\tilde{z}_1,\ldots,\tilde{z}_{2n})$ be the realizations, where $x\in[0,n]$. First, the joint distribution between W, R_i, $\tilde{Z}_i$ and $\tilde{Z}_{i+n}$ is calculated by

$$\begin{aligned}P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(\frac{x}{n},r_i,\tilde{z}_i,\tilde{z}_{i+n})=&\sum_{\substack{r_{[n]}\setminus r_i\\ \tilde{z}_{[2n]}\setminus(\tilde{z}_i,\tilde{z}_{i+n})}}P_{W,R_{[n]},\tilde{Z}_{[2n]}}(\frac{x}{n},r_{[n]},\tilde{z}_{[2n]})\\=&\frac{1}{2}\binom{n-1}{x-\tilde{z}_{i+r_i n}}p^{x-\tilde{z}_{i+r_i n}}(1-p)^{n-1-x+\tilde{z}_{i+r_i n}}p^{\tilde{z}_i+\tilde{z}_{i+n}}(1-p)^{2-\tilde{z}_i-\tilde{z}_{i+n}}.\end{aligned} \tag{108}$$

We further calculate the conditional probability as

$$\begin{aligned}P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(\frac{x}{n}|r_i,\tilde{z}_i,\tilde{z}_{i+n})=&\frac{P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(\frac{x}{n},r_i,\tilde{z}_i,\tilde{z}_{i+n})}{P_{R_i}(r_i)P_{\tilde{Z}_i}(\tilde{z}_i)P_{\tilde{Z}_{i+n}}(\tilde{z}_{i+n})}\\=&\frac{\frac{1}{2}\binom{n-1}{x-\tilde{z}_{i+r_i n}}p^{x-\tilde{z}_{i+r_i n}}(1-p)^{n-1-x+\tilde{z}_{i+r_i n}}p^{\tilde{z}_i+\tilde{z}_{i+n}}(1-p)^{2-\tilde{z}_i-\tilde{z}_{i+n}}}{\frac{1}{2}p^{\tilde{z}_i}(1-p)^{1-\tilde{z}_i}p^{\tilde{z}_{i+n}}(1-p)^{1-\tilde{z}_{i+n}}}\\=&\binom{n-1}{x-\tilde{z}_{i+r_i n}}p^{x-\tilde{z}_{i+r_i n}}(1-p)^{n-1-x+\tilde{z}_{i+r_i n}}.\end{aligned} \tag{109}$$

By marginalizing over R_i, we get

$$\begin{aligned}P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(\frac{x}{n}|\tilde{z}_i,\tilde{z}_{i+n})=&P_{R_i}(0)P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(\frac{x}{n}|0,\tilde{z}_i,\tilde{z}_{i+n})+P_{R_i}(1)P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(\frac{x}{n}|1,\tilde{z}_i,\tilde{z}_{i+n})\\=&\frac{1}{2}\binom{n-1}{x-\tilde{z}_i}p^{x-\tilde{z}_i}(1-p)^{n-1-x+\tilde{z}_i}+\frac{1}{2}\binom{n-1}{x-\tilde{z}_{i+n}}p^{x-\tilde{z}_{i+n}}(1-p)^{n-1-x+\tilde{z}_{i+n}}.\end{aligned} \tag{110}$$

With the above elemental probabilities obtained, we continue to calculate the mutual information:

$$
\begin{aligned}
I\left(W;R_i|\tilde{Z}_i,\tilde{Z}_{i+n}\right) &= \sum_{w,r_i,\tilde{z}_i,\tilde{z}_{i+n}} P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,r_i,\tilde{z}_i,\tilde{z}_{i+n}) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|r_i,\tilde{z}_i,\tilde{z}_{i+n})}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|\tilde{z}_i,\tilde{z}_{i+n})} \\
&= \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,0,0,0) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,0,0)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,0)} + \dots \\
&\quad + \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,0,1,1) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,1,1)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,1)} \\
&\quad + \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,1,0,0) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,0,0)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,0)} + \dots \\
&\quad + \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,1,1,1) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,1,1)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,1)},
\end{aligned}
\tag{111}
$$

each summand of which is simplified according to the value of $(r_i,\tilde{z}_i,\tilde{z}_{i+n})$:

$$
\begin{aligned}
&\sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,0,0,0) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,0,0)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,0)} = \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,1,0,0) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,0,0)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,0)} \\
&= \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,0,1,1) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,1,1)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,1)} = \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,1,1,1) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,1,1)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,1)} \\
&= 0,
\end{aligned}
\tag{112}
$$

and similarly,

$$
\begin{aligned}
\sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,0,0,1) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,0,1)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,1)} &= \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,1,1,0) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,1,0)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,0)} \\
&= \frac{1}{2} \sum_{x=0}^{n-1} \binom{n-1}{x} p^{x+1}(1-p)^{n-x} \log \frac{2}{\frac{x}{n-x}\cdot\frac{1-p}{p}+1},
\end{aligned}
\tag{113}
$$

$$
\begin{aligned}
\sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,0,1,0) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,1,0)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,0)} &= \sum_w P_{W,R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w,1,0,1) \log \frac{P_{W|R_i,\tilde{Z}_i,\tilde{Z}_{i+n}}(w|1,0,1)}{P_{W|\tilde{Z}_i,\tilde{Z}_{i+n}}(w|0,1)} \\
&= \frac{1}{2} \sum_{x=1}^{n} \binom{n-1}{x-1} p^{x}(1-p)^{n+1-x} \log \frac{2}{\frac{n-x}{x}\cdot\frac{p}{1-p}+1} \\
&= \frac{1}{2} \sum_{x=0}^{n-1} \binom{n-1}{x} p^{x+1}(1-p)^{n-x} \log \frac{2}{\frac{n-1-x}{x+1}\cdot\frac{p}{1-p}+1}.
\end{aligned}
\tag{114}
$$

Plugging all the summand terms back to (111) and we get

$$
\begin{aligned}
I\left(W;R_i|\tilde{Z}_i,\tilde{Z}_{i+n}\right) &= \sum_{x=0}^{n-1} \binom{n-1}{x} p^{x+1}(1-p)^{n-x} \log \frac{2}{\frac{x}{n-x}\cdot\frac{1-p}{p}+1} + \sum_{x=0}^{n-1} \binom{n-1}{x} p^{x+1}(1-p)^{n-x} \log \frac{2}{\frac{n-1-x}{x+1}\cdot\frac{p}{1-p}+1} \\
&= 2p(1-p)\log 2 - \sum_{x=0}^{n-1} \binom{n-1}{x} p^{x+1}(1-p)^{n-x} \log \left(\frac{x}{n-x}\cdot\frac{1-p}{p}+1\right)\left(\frac{n-1-x}{x+1}\cdot\frac{p}{1-p}+1\right) \\
&= p(1-p)\left(2 - \mathbb{E}\left[\log\left(\frac{X}{n-X}\cdot\frac{1-p}{p}+1\right)\left(\frac{n-1-X}{X+1}\cdot\frac{p}{1-p}+1\right)\right]\right) \\
&= p(1-p)\left(2 - \mathbb{E}\left[\log\left(\frac{\alpha X}{\bar{X}+1}+1\right)\right] - \mathbb{E}\left[\left(\frac{\beta\bar{X}}{X+1}+1\right)\right]\right),
\end{aligned}
\tag{115}
$$

where we let $X \sim \mathrm{Bin}(n-1,p)$, $\bar{X}=n-1-X \sim \mathrm{Bin}(n-1,1-p)$, $\alpha=\frac{1-p}{p}$ and $\beta=\frac{p}{1-p}$. Without generality, we assume $p<\frac{1}{2}$ and $\alpha>1$, $\beta<1$. With a slight abuse of Lemma 7, the above expression can be approximated as

$$\begin{aligned}
I\left(W;R_i|\tilde{Z}_i,\tilde{Z}_{i+n}\right) =& p(1-p)\left(2-\mathbb{E}\left[\log\frac{(\alpha-1)X+n}{\bar{X}+1}+\log\frac{n-(1-\beta)\bar{X}}{X+1}\right]\right)\\
=& p(1-p)\left(2-\mathbb{E}\left[\log(\alpha-1)+\log\left(X+\frac{n}{\alpha-1}\right)-\log\left(\bar{X}+1\right)\right.\right.\\
&\left.\left.+\log(1-\beta)+\log\left(\frac{n}{1-\beta}-\bar{X}\right)-\log\left(X+1\right)\right]\right)\\
=& p(1-p)\left(2-\log(\alpha-1)(1-\beta)-\log\left((n-1)p+\frac{n}{\alpha-1}\right)+\frac{(n-1)(1-p)p}{2\left((n-1)p+\frac{n}{\alpha-1}\right)^2\ln 2}\right.\\
&+\log\left((n-1)(1-p)+1\right)-\frac{(n-1)(1-p)p}{2\left((n-1)(1-p)+1\right)^2\ln 2}\\
&-\log\left(\frac{n}{1-\beta}-(n-1)(1-p)\right)+\frac{(n-1)(1-p)p}{2\left(\frac{n}{1-\beta}-(n-1)(1-p)\right)^2\ln 2}\\
&\left.+\log\left((n-1)p+1\right)-\frac{(n-1)(1-p)p}{2\left((n-1)p+1\right)^2\ln 2}+O\left(\frac{1}{n^2}\right)\right)\\
=& p(1-p)\left(2-\log(\alpha-1)(1-\beta)+\log\frac{(n-1)(1-p)+1}{(n-1)p+\frac{n}{\alpha-1}}+\log\frac{(n-1)p+1}{\frac{n}{1-\beta}-(n-1)(1-p)}+O\left(\frac{1}{n}\right)\right)\\
=& p(1-p)\left(2-\log(\alpha-1)(1-\beta)+\log\frac{1-p}{p+\frac{1}{\alpha-1}}+\log\frac{p}{\frac{1}{1-\beta}-(1-p)}+O\left(\frac{1}{n}\right)\right)\\
=& p(1-p)\left(2+\log\frac{1-p}{p(\alpha-1)+1}+\frac{p}{1-(1-p)(1-\beta)}+O\left(\frac{1}{n}\right)\right)\\
=& p(1-p)\left(2+\log\frac{1-p}{2-2p}+\log\frac{p}{2p}+O\left(\frac{1}{n}\right)\right)\\
=& O\left(\frac{1}{n}\right).
\end{aligned} \tag{116}$$

After taking the square root of $I(W;R_i|\tilde{Z}_i,\tilde{Z}_{i+n})$, one can find that the ICIMI bound is of order $O\left(\frac{1}{\sqrt{n}}\right)$.

6) LmO-CMI Case: Generalized from the LOO case, the learner in the LmO case is $W=\frac{1}{n}\sum_{i=1}^{n}\tilde{Z}_{U_i}$. Consider realizations $w=\frac{1}{n}\sum_{i=1}^{n}\tilde{z}_{u_i}=\frac{x}{n}$ and $\sum_{i=1}^{n+m}\tilde{z}_i-\sum_{i=1}^{n}\tilde{z}_{u_i}=y$, then the joint probability is

$$\begin{aligned}
P_{W,\tilde{Z}_{[n+m]},U_{[n]}}(\frac{x}{n},\tilde{z}_{[n+m]},u_{[n]}) &= P_{W|\tilde{Z}_{[n+m]},U_{[n]}}(\frac{x}{n}|\tilde{z}_{[n+m]},u_{[n]})P_{U_{[n]}}(u_{[n]})\prod_{i=1}^{n+m}P_{\tilde{Z}_i}(\tilde{z}_i)\\
&= \frac{1}{\binom{n+m}{n}}p^{x+y}(1-p)^{n+m-x-y}.
\end{aligned} \tag{117}$$

For the conditional probability $P_{W|\tilde{Z}_{[n+m]}}(\frac{x}{n}|\tilde{z}_{[n+m]})$, we have

$$\begin{aligned}
P_{W|\tilde{Z}_{[n+m]}}(\frac{x}{n}|\tilde{z}_{[n+m]}) &= \sum_{u'_{[n]}}P_{W|\tilde{Z}_{[n+m]},U_{[n]}}(\frac{x}{n}|\tilde{z}_{[n+m]},u'_{[n]})P_{U_{[n]}}(u'_{[n]})\\
&= \frac{1}{\binom{n+m}{n}}\sum_{u'_{[n]}}P_{W|\tilde{Z}_{[n+m]},U_{[n]}}(\frac{x}{n}|\tilde{z}_{[n+m]},u'_{[n]}).
\end{aligned} \tag{118}$$

We now consider the number of feasible $u_{[n]}$'s such that $P_{W|\tilde{Z}_{[n+m]},U_{[n]}}(\frac{x}{n}|\tilde{z}_{[n+m]},u_{[n]})=1$, which means $\sum_{i=1}^{n}\tilde{z}_{u_i}=x$ is still satisfied. For clarity, we say that $\tilde{z}_i=1$ is a "successful trial" while the opposite case, i.e., $\tilde{z}_i=0$, is referred to as a "failed trial". Consider that there are $x+y$ successful trials and $n+m-x-y$ failed trials among all $n+m$ trials, then a feasible $u_{[n]}$ should choose x successful trials, which contributes to $\binom{x+y}{x}$ free choices, along with $n-x$ failed trials, which provides $\binom{n+m-x-y}{n-x}$ choices. Thus,

$$P_{W|\tilde{Z}_{[n+m]}}(\frac{x}{n}|\tilde{z}_{[n+m]}) = \frac{\binom{x+y}{x}\binom{n+m-x-y}{n-x}}{\binom{n+m}{n}}. \tag{119}$$

Then, the mutual information is given by

$$\begin{aligned}I\left(W;U_{[n]}|\tilde{Z}_{[n+m]}\right) &= \sum_{w,\tilde{z}_{[n+m]},u_{[n]}} P_{W,\tilde{Z}_{[n+m]},U_{[n]}}(w,\tilde{z}_{[n+m]},u_{[n]}) \log \frac{P_{W|\tilde{Z}_{[n+m]},U_{[n]}}(w|\tilde{z}_{[n+m]},u_{[n]})}{P_{W|\tilde{Z}_{[n+m]}}(w|\tilde{z}_{[n+m]})} \\ &= \sum_{x=0}^{n}\sum_{y=0}^{m} \frac{\binom{x+y}{x}\binom{n+m-x-y}{n-x}}{\binom{n+m}{n}} \binom{n+m}{x+y} p^{x+y}(1-p)^{n+m-x-y} \log \frac{\binom{n+m}{n}}{\binom{x+y}{x}\binom{n+m-x-y}{n-x}} \\ &= \log\binom{n+m}{n} - \sum_{x=0}^{n}\sum_{y=0}^{m}\binom{n}{x}p^x(1-p)^{n-x}\binom{m}{y}p^y(1-p)^{m-y}\log\binom{x+y}{x}\binom{n+m-x-y}{n-x}.\end{aligned} \tag{120}$$

Let X and Y be two binomial random variables following $X \sim \mathrm{Bin}(n,p)$ and $Y \sim \mathrm{Bin}(m,p)$, respectively, then the above mutual information can be rewritten as

$$I\left(W;U_{[n]}|\tilde{Z}_{[n+m]}\right) = \log\binom{n+m}{n} - \mathbb{E}\left[\log\binom{X+Y}{X}\binom{\bar{X}+\bar{Y}}{\bar{X}}\right], \tag{121}$$

where $\bar{X} = n-X$ and $\bar{Y} = m-Y$. Note that one can verify that (121) will reduce to (106) with $m=1$, while the detailed calculations are omitted.

Next, we will approximate each binomial coefficient term with a slight abuse of Lemma 10. Specifically, we have

$$\log\binom{n+m}{n} \le \frac{1}{2}\log\frac{n+m}{nm} - \frac{1}{2}\log 2\pi + (n+m)H\left(\frac{n}{n+m}\right). \tag{122}$$

Then, for the expected binomial coefficients, we get

$$\begin{aligned}\mathbb{E}\left[\log\binom{X+Y}{X}\right] \ge& \mathbb{E}\left[\frac{1}{2}\left(\log(X+Y) - \log X - \log Y - c\right) + X\log\frac{X+Y}{X} + Y\log\frac{X+Y}{Y}\right] \\ =& \mathbb{E}\left[\frac{1}{2}\left(\log Z - \log X - \log Y - c\right) + Z\log Z - X\log X - Y\log Y\right] \\ =& \frac{1}{2}\Bigg(\log(n+m)p - \frac{1-p}{2(n+m)p\ln 2} + O\left(\frac{1}{(n+m)^2}\right) \\ & - \log np + \frac{1-p}{2np\ln 2} + O\left(\frac{1}{n^2}\right) - \log mp + \frac{1-p}{2mp\ln 2} + O\left(\frac{1}{m^2}\right)\Bigg) \\ & + (n+m)p\log(n+m)p + \frac{1-p}{2\ln 2} + \frac{1-p^2}{12(n+m)p\ln 2} + O\left(\frac{1}{(n+m)^2}\right) \\ & - np\log np - \frac{1-p}{2\ln 2} - \frac{1-p^2}{12np\ln 2} + O\left(\frac{1}{n^2}\right) - mp\log np - \frac{1-p}{2\ln 2} - \frac{1-p^2}{12mp\ln 2} + O\left(\frac{1}{m^2}\right) \\ =& \frac{1}{2}\log\frac{n+m}{nm} + np\log\frac{n+m}{n} + mp\log\frac{n+m}{m} - \frac{\log p + (1-p)\log e}{2} \\ & + \frac{(p-1)(p-2)}{12p\ln 2}\left(\frac{1}{n} + \frac{1}{m} - \frac{1}{n+m}\right) + O\left(\frac{1}{(n+m)^2}\right) + O\left(\frac{1}{n^2}\right) + O\left(\frac{1}{m^2}\right).\end{aligned} \tag{123}$$

Likewise, it holds that

$$\begin{aligned}\mathbb{E}\left[\log\binom{\bar{X}+\bar{Y}}{\bar{X}}\right] \ge& \frac{1}{2}\log\frac{n+m}{nm} + n(1-p)\log\frac{n+m}{n} + m(1-p)\log\frac{n+m}{m} - \frac{\log(1-p) + p\log e}{2} \\ & + \frac{p(p+1)}{12(1-p)\ln 2}\left(\frac{1}{n} + \frac{1}{m} - \frac{1}{n+m}\right) + O\left(\frac{1}{(n+m)^2}\right) + O\left(\frac{1}{n^2}\right) + O\left(\frac{1}{m^2}\right).\end{aligned} \tag{124}$$

Then combine all the results, it follows that

$$
\begin{aligned}
I\left(W;U_{[n]}|\tilde{Z}_{[n+m]}\right) \leq& \frac{1}{2}\log\frac{n+m}{nm}-\frac{1}{2}\log 2\pi+(n+m)H\left(\frac{n}{n+m}\right)\\
&-\frac{1}{2}\log\frac{n+m}{nm}-np\log\frac{n+m}{n}-mp\log\frac{n+m}{m}+\frac{\log p+(1-p)\log e}{2}\\
&-\frac{(p-1)(p-2)}{12p\ln 2}\left(\frac{1}{n}+\frac{1}{m}-\frac{1}{n+m}\right)+O\left(\frac{1}{(n+m)^2}\right)+O\left(\frac{1}{n^2}\right)+O\left(\frac{1}{m^2}\right)\\
&-\frac{1}{2}\log\frac{n+m}{nm}-n(1-p)\log\frac{n+m}{n}-m(1-p)\log\frac{n+m}{m}+\frac{\log(1-p)+p\log e}{2}\\
&-\frac{p(p+1)}{12(1-p)\ln 2}\left(\frac{1}{n}+\frac{1}{m}-\frac{1}{n+m}\right)+O\left(\frac{1}{(n+m)^2}\right)+O\left(\frac{1}{n^2}\right)+O\left(\frac{1}{m^2}\right)\\
=&\frac{1}{2}\log\frac{nm}{n+m}+\frac{\log\frac{p(1-p)e}{\pi}-1}{2}-\frac{5p^2-5p+2}{12p(1-p)\ln 2}\left(\frac{1}{n}+\frac{1}{m}-\frac{1}{n+m}\right)\\
&+O\left(\frac{1}{(n+m)^2}\right)+O\left(\frac{1}{n^2}\right)+O\left(\frac{1}{m^2}\right).
\end{aligned}
\tag{125}
$$

Then we consider two cases of m. First, if m is a constant, we have $\log\frac{nm}{n+m}=\log m-\frac{m}{n+m}+O(\frac{1}{n^2})$, and thus the whole L$m$O-CMI bound will converge to a constant. Otherwise, if $m\propto n$, then the above expression is of order $O(\log n)$. By resorting to (17) and taking into consideration the factor before the mutual information quantity, it can be verified that the LmO-CMI bound is of order $O(\frac{\log n}{n})$.

7) (m,m)-IPCIMI and LOFO-IPCIMI Cases: We now evaluate the (m,m)-IPCIMI and LOFO-IPCIMI bounds. The learner is $\frac{1}{n}\sum_{i=1}^{m}\left(\sum_{j=1}^{\frac{n}{m}+1}\tilde{Z}_j^{(i)}-\tilde{Z}_U^{(i)}\right)$. Let $w=\frac{x+y}{n}$, $\tilde{z}^{(i)}_{u^{(i)}}=t$ and $\sum_{j=1}^{\frac{n}{m}+1}\tilde{z}_j^{(i)}-t=y$, where $x\in[0,n-\frac{n}{m}]$, $y\in[0,\frac{n}{m}]$ and $t\in\{0,1\}$. Then,

$$
P_{W,U^{(i)},\tilde{Z}^{(i)}_{[\frac{n}{m}+1]}}(\frac{x+y}{n},u^{(i)},\tilde{z}^{(i)}_{[\frac{n}{m}+1]})=\frac{1}{\frac{n}{m}+1}\binom{n-\frac{n}{m}}{x}p^{x+y}(1-p)^{n-x-y}p^t(1-p)^t. \tag{126}
$$

$$
P_{W|U^{(i)},\tilde{Z}^{(i)}_{[\frac{n}{m}+1]}}(\frac{x+y}{n}|u^{(i)},\tilde{z}^{(i)}_{[\frac{n}{m}+1]})=\binom{n-\frac{n}{m}}{x}p^{x}(1-p)^{n-\frac{n}{m}-x}. \tag{127}
$$

$$
\begin{aligned}
P_{W|\tilde{Z}^{(i)}_{[\frac{n}{m}+1]}}(\frac{x+y}{n}|\tilde{z}^{(i)}_{[\frac{n}{m}+1]})=&\sum_{u'}P_{W|U^{(i)},\tilde{Z}^{(i)}_{[\frac{n}{m}+1]}}(\frac{x+y}{n}|u',\tilde{z}^{(i)}_{[\frac{n}{m}+1]})P_{U^{(i)}}(u')\\
=&\frac{y+t}{\frac{n}{m}+1}\binom{n-\frac{n}{m}}{x+1-t}p^{x+1-t}(1-p)^{n-\frac{n}{m}-x-1+t}\\
&+\frac{\frac{n}{m}+1-y-t}{\frac{n}{m}+1}\binom{n-\frac{n}{m}}{x-t}p^{x-t}(1-p)^{n-\frac{n}{m}-x+t}
\end{aligned}
\tag{128}
$$

Using the above expressions, we calculate the mutual information as

$$
\begin{aligned}
&I\left(W;U^{(i)}|\tilde{Z}^{(i)}_{[\frac{n}{m}+1]}\right)\\
=&\sum_{x=0}^{n-\frac{n}{m}}\sum_{y=0}^{\frac{n}{m}}\frac{y+1}{\frac{n}{m}+1}\binom{\frac{n}{m}+1}{y+1}\binom{n-\frac{n}{m}}{x}p^{x+y}(1-p)^{n-x-y}p\log\frac{\frac{n}{m}+1}{y+1+\frac{1-p}{p}\cdot\frac{(\frac{n}{m}-y)x}{n-\frac{n}{m}+1-x}}\\
&+\sum_{x=0}^{n-\frac{n}{m}}\sum_{y=0}^{\frac{n}{m}}\frac{\frac{n}{m}+1-y}{\frac{n}{m}+1}\binom{\frac{n}{m}+1}{y}\binom{n-\frac{n}{m}}{x}p^{x+y}(1-p)^{n-x-y}(1-p)\log\frac{\frac{n}{m}+1}{\frac{p}{1-p}\cdot\frac{(n-\frac{n}{m}-x)y}{x+1}+\frac{n}{m}+1-y}\\
=&\sum_{x=0}^{n-\frac{n}{m}}\sum_{y=0}^{\frac{n}{m}}\binom{\frac{n}{m}}{y}p^y(1-p)^{\frac{n}{m}-y}\binom{n-\frac{n}{m}}{x}p^x(1-p)^{n-\frac{n}{m}-x}\Bigg(p\log\frac{\frac{n}{m}+1}{y+1+\frac{1-p}{p}\cdot\frac{(\frac{n}{m}-y)x}{n-\frac{n}{m}+1-x}}\\
&+(1-p)\log\frac{\frac{n}{m}+1}{\frac{p}{1-p}\cdot\frac{(n-\frac{n}{m}-x)y}{x+1}+\frac{n}{m}+1-y}\Bigg)
\end{aligned}
\tag{129}
$$

which can be rewritten as

$$
I\left(W;U^{(i)}|\tilde{Z}^{(i)}_{[\frac{n}{m}+1]}\right)=\log\left(\frac{n}{m}+1\right)-p\mathbb{E}\left[\log\left(Y+1+\frac{\alpha X\bar{Y}}{\bar{X}+1}\right)\right]-(1-p)\mathbb{E}\left[\log\left(\bar{Y}+1+\frac{\beta\bar{X}Y}{X+1}\right)\right], \tag{130}
$$

where $X \sim \mathrm{Bin}\left(n-\frac{n}{m},p\right)$, $Y \sim \mathrm{Bin}\left(\frac{n}{m},p\right)$, $\bar{X}=n-\frac{n}{m}-X \sim \mathrm{Bin}\left(n-\frac{n}{m},1-p\right)$, $\bar{Y}=\frac{n}{m}-Y \sim \mathrm{Bin}\left(\frac{n}{m},1-p\right)$, $\alpha=\frac{1-p}{p}$ and $\beta=\frac{p}{1-p}$. Through the above expression, one can validate the argument that the (m,m)-IPCIMI bound is a general case of the LOO-CMI and ICIMI bounds. With detailed calculations omitted, we note that letting $m=1$ reduces (130) to the LOO-CMI bound in (106), whereas setting $m=n$ leads to the ICIMI bound shown in (115). Furthermore, it can be checked that (130) also holds for the LOFO-IPCIMI bound. In the following, we continue to derive the order of (130).

Let $f(X,Y)$ be any deterministic function of X and Y. By multivariate Taylor expansion, we approximate $f(X,Y)$ at points $X=\mathbb{E}[X]$ and $Y=\mathbb{E}[Y]$:

$$\begin{aligned} f(X,Y) \approx & f\left(\mathbb{E}[X],\mathbb{E}[Y]\right)+\frac{\partial f}{\partial X}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\left(X-\mathbb{E}[X]\right)+\frac{\partial f}{\partial Y}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\left(Y-\mathbb{E}[Y]\right)+\frac{\partial^2 f}{\partial X^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\frac{\left(X-\mathbb{E}[X]\right)^2}{2} \\ & +\frac{\partial^2 f}{\partial X \partial Y}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\left(X-\mathbb{E}[X]\right)\left(Y-\mathbb{E}[Y]\right)+\frac{\partial^2 f}{\partial Y^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\frac{\left(Y-\mathbb{E}[Y]\right)^2}{2}. \end{aligned} \tag{131}$$

Take expectation on both sides and use the independence of X and Y, it follows that

$$\begin{aligned} \mathbb{E}\left[f(X,Y)\right] \approx & f\left(\mathbb{E}[X],\mathbb{E}[Y]\right)+\frac{\partial^2 f}{\partial X^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\frac{\mathbb{E}\left[\left(X-\mathbb{E}[X]\right)^2\right]}{2}+\frac{\partial^2 f}{\partial Y^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\frac{\mathbb{E}\left[\left(Y-\mathbb{E}[Y]\right)^2\right]}{2} \\ = & f\left(\mathbb{E}[X],\mathbb{E}[Y]\right)+\frac{\partial^2 f}{\partial X^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\frac{m-1}{2m}np(1-p)+\frac{\partial^2 f}{\partial Y^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)\frac{n}{2m}p(1-p). \end{aligned} \tag{132}$$

On the one hand, by letting $f_1(X,Y)=\log\left(Y+1+\frac{\alpha X\bar{Y}}{\bar{X}+1}\right)$ (where $\bar{X}$ and $\bar{Y}$ are deterministic functions of X and Y, respectively), we now compute $f_1\left(\mathbb{E}[X],\mathbb{E}[Y]\right)$ and prove that both $\frac{\partial^2 f_1}{\partial X^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)$ and $\frac{\partial^2 f_1}{\partial Y^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)$ are of order $O\left(\frac{1}{n^2}\right)$. First, plugging $X=\mathbb{E}[X]=\frac{m-1}{m}np$ and $Y=\mathbb{E}[Y]=\frac{n}{m}p$ into $f_1\left(X,Y\right)$ gives

$$\begin{aligned} f_1\left(\mathbb{E}[X],\mathbb{E}[Y]\right) = & \log\left(\frac{n}{m}p+1+\frac{(1-p)^2\frac{n}{m}\frac{m-1}{m}n}{(1-p)\frac{m-1}{m}n+1}\right) \\ = & \log\frac{(1-p)(m-1)n^2+((1-p)(m-1)+p)mn+m^2}{(1-p)(m-1)mn+k^2}. \end{aligned} \tag{133}$$

Next, we compute the following partial derivatives of f_1 as

$$\frac{\partial f_1}{\partial X}\left(X,Y\right)=\frac{\alpha\left(\frac{m-1}{m}n+1\right)Y}{\left(\bar{X}+1\right)^2\left(Y+1\right)+\alpha\left(\bar{X}+1\right)XY}, \tag{134}$$

$$\frac{\partial^2 f_1}{\partial X^2}\left(X,Y\right)=\frac{\alpha\left(\frac{m-1}{m}n+1\right)\left(\alpha XY+(2-\alpha)\bar{X}Y+2\bar{X}+(2-\alpha)Y+2\right)Y}{\left(\left(\bar{X}+1\right)^2\left(Y+1\right)+\alpha\left(\bar{X}+1\right)XY\right)^2}, \tag{135}$$

$$\frac{\partial f_1}{\partial Y}\left(X,Y\right)=\frac{\bar{X}-\alpha X+1}{\left(\bar{X}+1\right)\left(Y+1\right)+\alpha X\bar{Y}}, \tag{136}$$

$$\frac{\partial^2 f_1}{\partial Y^2}\left(X,Y\right)=-\frac{\left(\bar{X}-\alpha X+1\right)^2}{\left(\left(\bar{X}+1\right)\left(Y+1\right)+\alpha X\bar{Y}\right)^2}. \tag{137}$$

At points $X=\frac{m-1}{m}np$, $\bar{X}=\frac{m-1}{m}n(1-p)$, $Y=\frac{n}{m}p$ and $\bar{Y}=\frac{n}{m}(1-p)$, one can verify that

$$\frac{\partial f_1}{\partial X}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)=O\left(\frac{1}{n}\right), \qquad \frac{\partial^2 f_1}{\partial X^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)=O\left(\frac{1}{n^2}\right), \tag{138}$$

$$\frac{\partial f_1}{\partial X}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)=O\left(\frac{1}{n}\right), \qquad \frac{\partial^2 f_1}{\partial X^2}\left(\mathbb{E}[X],\mathbb{E}[Y]\right)=O\left(\frac{1}{n^2}\right). \tag{139}$$

Plugging (133), (138) and (139) back into (132) yields

$$\mathbb{E}\left[f_1(X,Y)\right]=\log\frac{(1-p)(m-1)n^2+((1-p)(m-1)+p)mn+m^2}{(1-p)(m-1)mn+m^2}+O\left(\frac{1}{n}\right). \tag{140}$$

On the other, we let $f_2(X,Y)=\log\left(\bar{Y}+1+\frac{\beta\bar{X}Y}{X+1}\right)$ and follow a similar procedure to obtain

$$\mathbb{E}\left[f_2(X,Y)\right]=\log\frac{p(m-1)n^2+(p(m-1)+1-p)mn+m^2}{p(m-1)mn+m^2}+O\left(\frac{1}{n}\right). \tag{141}$$

Combining all the results yields

$$
\begin{aligned}
I\left(W;U^{(i)}|\tilde{Z}^{(i)}_{[\frac{n}{m}+1]}\right) =&\log\left(\frac{n}{m}+1\right)-p\mathbb{E}\left[f_1(X,Y)\right]-(1-p)\mathbb{E}\left[f_2(X,Y)\right]\\
=&p\left(\log\left(\frac{n}{m}+1\right)-\log\frac{(1-p)(m-1)n^2+((1-p)(m-1)+p)mn+m^2}{(1-p)(m-1)mn+m^2}\right)\\
&+(1-p)\left(\log\left(\frac{n}{m}+1\right)-\log\frac{p(m-1)n^2+(p(m-1)+1-p)mn+m^2}{p(m-1)mn+m^2}\right)+O\left(\frac{1}{n}\right)\\
=&p\log\left(1+\frac{(1-p)mn}{(1-p)(m-1)n^2+((1-p)(m-1)+p)mn+m^2}\right)\\
&+(1-p)\log\left(1+\frac{pmn}{p(m-1)n^2+(p(m-1)+1-p)mn+m^2}\right)+O\left(\frac{1}{n}\right)\\
=&\frac{p(1-p)mn}{(1-p)(m-1)n^2+((1-p)(m-1)+p)mn+m^2}\\
&+\frac{p(1-p)mn}{p(m-1)n^2+(p(m-1)+1-p)mn+m^2}+O\left(\frac{1}{n}\right)\\
=&O\left(\frac{1}{n}\right),
\end{aligned}
\tag{142}
$$

plugging which into (26) implies that the (m,m)-IPCIMI bound scale as $O\left(\frac{1}{\sqrt{n}}\right)$. Also, LOFO-IPCIMI shares the same order with the (m,m)-IPCIMI bound.

8) (m,n)-IPCIMI Case: In Section IV, the asymptotic performance of the (m,n)-IPCIMI bound is compared with the IMI bound to validate Corollary 4. This subsection provides the calculation result of the (m,n)-IPCIMI bound under the Bernoulli example, without evaluating its decay order. As derived from the original (m,k)-IPCIMI bound in (23), the (m,n)-IPCIMI bound is given by

$$
\mathrm{gen}(\mathcal{A})\leq\frac{n+m}{nm}\sum_{i=1}^{n}\sqrt{\frac{1}{2}I\left(W;U^{(i)}|\tilde{Z}^{(i)}_{[\frac{m}{n}+1]}\right)}. \tag{143}
$$

Following a similar procedure in the (m,m)-IPCIMI case, we calculate $I(W;U^{(i)}|\tilde{Z}^{(i)}_{[\frac{m}{n}+1]})$ and the result is presented as follows:

$$
I\left(W;U^{(i)}|\tilde{Z}^{(i)}_{[\frac{m}{n}+1]}\right)=\log\left(\frac{m}{n}+1\right)-p\mathbb{E}\left[\log\left(Y+1+\frac{\beta\bar{X}\bar{Y}}{X+1}\right)\right]-(1-p)\mathbb{E}\left[\log\left(\bar{Y}+1+\frac{\alpha XY}{\bar{X}+1}\right)\right], \tag{144}
$$

where $X\sim\mathrm{Bin}\left(n-1,p\right)$, $Y\sim\mathrm{Bin}\left(\frac{m}{n},p\right)$, $\bar{X}=n-1-X\sim\mathrm{Bin}\left(n-1,1-p\right)$, $\bar{Y}=\frac{m}{n}-Y\sim\mathrm{Bin}\left(\frac{m}{n},1-p\right)$, $\alpha=\frac{1-p}{p}$ and $\beta=\frac{p}{1-p}$.