Measures of Information Leakage for Incomplete Statistical Information: Application to a Binary Privacy Mechanism

SHAHNEWAZ KARIM SAKIB, Iowa State University, USA
GEORGE T AMARIUCAI, Kansas State University, USA
YONG GUAN, Iowa State University, USA

Information leakage is usually defined as the logarithmic increment in the adversary's probability of correctly guessing the legitimate user's private data or some arbitrary function of the private data when presented with the legitimate user's publicly disclosed information. However, this definition of information leakage implicitly assumes that both the privacy mechanism and the prior probability of the original data are entirely known to the attacker. In reality, the assumption of complete knowledge of the privacy mechanism for an attacker is often impractical. The attacker can usually have access to only an approximate version of the correct privacy mechanism, computed from a limited set of the disclosed data, for which they can access the corresponding un-distorted data. In this scenario, the conventional definition of leakage no longer has an operational meaning. To address this problem, in this article, we propose novel meaningful information-theoretic metrics for information leakage when the attacker has *incomplete information* about the privacy mechanism—we call them *average subjective leakage*, *average confidence boost*, and *average objective leakage*, respectively. For the simplest, binary scenario, we demonstrate how to find an optimized privacy mechanism that minimizes the worst-case value of either of these leakages.

CCS Concepts: • **Mathematics of computing** → **Information theory**; **Numerical analysis**; • **Security and privacy** → **Information flow control**; *Trust frameworks;*

Additional Key Words and Phrases: Average subjective leakage, average objective leakage, average confidence boost, incomplete information, data privacy, information leakage

ACM Reference format:
Shahnewaz Karim Sakib, George T Amariucai, and Yong Guan. 2023. Measures of Information Leakage for Incomplete Statistical Information: Application to a Binary Privacy Mechanism. *ACM Trans. Priv. Sec.* 26, 4, Article 47 (November 2023), 31 pages.
https://doi.org/10.1145/3624982

This work was partially supported by NIST CSAFE under Cooperative Agreement Nos. 70NANB15H176 and 70NANB20H019, NSF under grant Nos. CNS-1527579, CNS-1619201, CNS-1730275, DEB-1924178, ECCS-2030249, NSF 2130889, 2232461, 2229654, NIFA 2021-67021-33775, and Boeing Company. This publication was made possible by NPRP grant #12C-33905-SP-165 from the Qatar National Research Fund (a member of Qatar Foundation).
Authors' addresses: S. K. Sakib and Y. Guan, Iowa State University, Ames, Iowa; e-mails: {ssakib, guan}@iastate.edu; G. T. Amariucai, Kansas State University, Manhattan, Kansas; e-mail: amariucai@ksu.edu.

2471-2566/2023/11-ART47 $15.00
https://doi.org/10.1145/3624982

1 INTRODUCTION

Due to the recent surge in popularity of social media platforms, users all over the world are disclosing increasing amounts of personal information, in exchange for certain gratifications, like those derived from social interactions [58], or directly measurable utilities, such as receiving more accurate recommendations. Heretofore, Facebook has 2.80 billion monthly active users, and more than half of those users log onto this networking site on a daily basis [40]. Even though there are various levels of privacy protections in these social networking sites, it is nevertheless possible for users' information to extend beyond their intended privacy bounds, hence leaking information about the users' sensitive personal data. It is of utmost importance to design measures to minimize such information leakage.

Much research work has been dedicated to measuring privacy and the development of privacy-preserving solutions. The measurement of privacy spans a wide range of fields, from data science to information theory. Differential privacy, introduced in Reference [18], emerged as a consensus definition for publishing data in a privacy-preserving manner. This metric is formulated around two databases that differ in a single entry. Later, Barthe and Kopf [6] extended this metric to deal with binary random variables masked by the addition of noise. In a more recent work, Desfontaines et al. [16] performed an in-depth analysis of the extensions and variations of differential privacy.

In the field of information theory, various privacy metrics have been proposed utilizing Shannon's mutual information or Fisher information [23]. *Shannon entropy* and *mutual information* based information-theoretic measures were proposed to define information leakage in References [27, 36, 44, 49]. Divergence-based quantities, such as total variation distance between the prior and posterior distribution, have also been adopted as a measure of leakage [43]. Issa et al. [30] introduced one-shot measures such as maximal leakage, maximal realizable leakage, maximal correlation, or local differential privacy [31] for quantifying information leakage.

Each of the metrics mentioned above only provides operational meaning if it is assumed that the stochastic privacy mechanism is completely known to the adversary, and only the specific realizations remain unknown. For example, if the mechanism relies on the addition of noise, then the statistics of the noise are perfectly known, but the samples drawn according to these statistics are not. However, this assumption is not compatible with practice. Consider a data aggregator that discloses data through such a privacy mechanism. As the original data is not available to the attacker, its statistics are not available either. So, even if the privacy mechanism is derived from an optimization problem as a function of the original statistics, it remains beyond the reach of the attacker. The attacker can collect few original data samples—perhaps from some friends—and estimate the privacy mechanism by comparing them to the disclosed data. But it is unlikely that the attacker's group of friends is so vast to render this estimation error insignificant. Hence, identifying metrics to precisely quantify privacy and privacy leakage in such scenarios is essential, as the traditional metrics fail to measure this leakage appropriately.

In this article, we have dealt with precisely this problem and provided meaningful measures of information leakage when the complete statistical information of either the privacy mechanism or the original data or both are unknown. It is to be noted that the privacy measures, in this scenario, rely on the attacker's *acquired* understanding of the privacy mechanism. For example, to represent how much confidence boost the attacker *expects* to get through the disclosure of information, we have proposed the ***average subjective leakage (ASL)***. This metric will help the attacker decide whether the cost incurred in the process of acquiring the information is worth the effort. To capture the *true* confidence boost of the attacker, we have proposed the metric ***average confidence boost (ACB)***. Finally, to represent the true probability that the attacker made a correct guess after observing the disclosed information, we have defined ***average objective leakage (AOL)***. This metric quantifies how much the disclosed information has really helped (or hurt!) the attacker.

We would like to point out here that imposing such uncertainty assumptions on the attacker is by no means tantamount to *security through obscurity*. This is because the attacker's imperfect statistical information does not in general constitute an advantage for the privacy-conscious data owner. Rather, at times, such imperfect information can actually lead the attacker to take actions that are more detrimental to the data owner than any actions taken in the presence of complete information, as should become clear in Section 2 below.

In the most related work to our framework, Chatzikokolakis et al. [11] also considered a scenario where the passive attacker lacks prior knowledge about the system. Eventually, the authors estimated the channel capacity based on the transition matrix generated from the collection of samples. Wang et al. [56] considered private data S and correlated data X and designed a privacy mechanism so the disclosed information Y—a noisy version of X—optimizes the privacy-utility tradeoff. The part related to our current article is their assumption that the joint probability distribution of (S, X) is not perfectly known, but rather approximated, by the adversary. Privacy is quantified as the χ^2 information between S and Y (i.e., $\chi^2(S, Y)$) and utility as χ^2 information between X and Y (i.e., $\chi^2(X, Y)$). These definitions require the knowledge of the joint distribution of (S, X), which means that the attacker achieves only an approximate version of $\chi^2(S, Y)$—say, $\hat{\chi}^2(S, Y)$. Afterward, the authors provide a bound for the error $\chi^2(S, Y) - \hat{\chi}^2(S, Y)$. Even though the authors correctly point out that there is some difference between these two measurements, they do not provide a measure that captures the leakage that the adversary can achieve from their approximated mechanism. The proposed metrics, therefore, have a definite advantage over Reference [56], as these measures are explicitly related to the privacy mechanism that is approximated by the adversary.

Now, to demonstrate the behavior of the proposed metrics, we have considered the simplest possible scenario. In our analysis, we disclose a binary private variable through a binary privacy mechanism and determine the optimal mechanism that minimizes the worst-case leakages under the constraint of both the utility of the revealed data and the deviation of the approximated mechanism from the true distribution of the privacy mechanism.

The contributions of our work can be summarized as follows: (1) We introduce novel notions of information-theoretic metrics when the complete statistical information of the privacy mechanism is unknown—these are *average subjective leakage*, *average confidence boost*, and *average objective leakage*. (2) We formulate several optimization problems, based on binary random variables and privacy mechanisms, to illustrate the applicability of our privacy leakage metrics. In each of those optimization problems, we analytically compute the parameters that minimize the worst-case realization of each of the metrics. (3) We numerically compute the optimal probabilities that result in the optimized privacy mechanism for each optimization problem and compare the obtained leakages to the worst-case leakage of the same metric that will result from other optimization problems.

We presented the preliminary results in Reference [47]. In this manuscript, we have renamed the metrics that were introduced in Reference [47] to better reflect their operational meaning. This current version further enhances Reference [47] in the following manner:

- We have discussed the motivation of the problem formulation for the binary privacy mechanism to demonstrate the applicability of the metrics in practical scenarios.
- We have introduced a new metric to compute the *true* confidence boost that an attacker will achieve through the disclosure of information. We have termed the metric as *average confidence boost* and formulated an optimization problem that minimizes the worst-case value of the metric. In Reference [47], we formulated optimization problems for minimizing the worst-case value of *average subjective leakage* (referred to as *maximal subjective leakage* in Reference [47]).

— We have also presented and proved several properties of the proposed metrics.
— We have also introduced the notion of an advanced utility provider and formulated optimization problems for this type of utility provider. We have only dealt with a utility provider who takes the output at face value (termed as a generic utility provider in this manuscript) in Reference [47].
— We have numerically computed the worst-case leakage values for advanced utility provider and compared the value with the value achieved by optimizing the min-entropy leakage.
— Finally, we have outlined the formulation of the optimization problem in the general scenario and discussed the approach of finding a solution.

The rest of the article is organized as follows: The motivation of the article is discussed in Section 2. In Section 3, we explain the system setup, define the proposed metrics, formulate the optimization problems, and explain several properties of the proposed metrics. Section 4 evaluates the feasibility of all possible combinations of probabilities of actual and approximated privacy mechanisms under various input distribution probabilities. The details of the *ACB* optimization problem are explained in Section 5. The simulation result is illustrated in Section 6. A brief overview of the optimization problem for the general case and their solution is discussed in Section 7. Several related works have been discussed in Section 8. Finally, Section 9 summarizes our article and presents some directions for future works.

2 MOTIVATION

According to a study by Ohio State University psychologist Terri Fisher, students are likely to skew information regarding their sexual behavior to match society's cultural norms [24]. The study was performed on 293 participants, consisting of general psychology students, who were asked to complete a questionnaire asking how often they engage in "typical" gender behaviors. Almost half of the participants were attached to a polygraph machine and were told that the machine would detect the lies, whereas, in reality, the machine was not actually functioning. The study found that the male participants, who were not attached to the polygraph machine, tended to report having more sexual partners than other male participants who were connected to the machine. Similarly, the reported number of female participants is lower for those who were not attached to the machine than those who believed the machine would detect their lies. Such a discrepancy was not found for questions related to non-sexual behaviors. Therefore, it is highly plausible that people may not always provide an honest answer when presented with a sensitive question.

Such a conclusion, naturally, leads us to Warner's randomized response model [59]. In the randomized response model, a respondent answers "Yes" or "No" to either the sensitive question of interest or the complementary question. The respondent uses a chance device to determine which question to answer. Even though the interviewer may know *a priori* the distribution of the device, the output of the device for each trial is unknown to the interviewer. Consequently, the interviewer does not know the actual question to which each respondent answered.

Let us denote θ as the probability that the chance device selects the sensitive question (Q_1), π as the proportion of the sampled population that belongs to the sensitive group, and λ as the probability that any randomly selected person will answer "Yes." Here, π is our parameter of interest, as we want to determine the proportion of population having the sensitive attribute. A pictorial depiction of the privacy mechanism of the randomized model is shown in Figure 1. The relation among θ, π, and λ is given below ($\theta \neq 0.5$) [59]:

$$\begin{aligned} \lambda = P(\text{Ans=Yes}) &= P(\text{Ans=Yes} \mid Q_1)P(Q_1) + P(\text{Ans=Yes} \mid Q_1^c)P(Q_1^c) \\ &= \pi\theta + (1-\pi)(1-\theta) = (2\theta - 1)\pi + (1-\theta). \end{aligned} \tag{1}$$

Generally, it is assumed that both the respondent and the interviewer know θ, and λ can be computed from the survey data. Therefore, from Equation (1), we can easily compute π as $\pi = \frac{\lambda-(1-\theta)}{2\theta-1}$. Afterward, the interviewer can utilize the values of π and λ to compute the number of people belonging to the group of people having sensitive attributes.

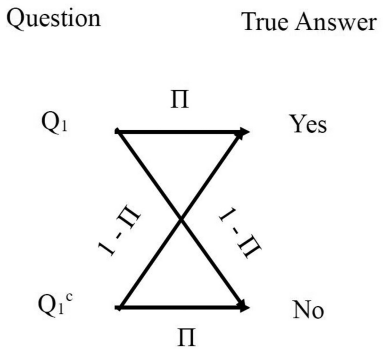


Fig. 1. Privacy mechanism in Warner's randomized response model.

Note that the underlying rationale of the randomized response model is that even though the respondents may feel uncomfortable for being perceived as having a sensitive attribute, as long as the answer is not too revealing, the respondents will follow the procedure. Nonetheless, such an assumption does not match with the survey results conducted by Fisher et al. [24], where the participants intentionally tweaked their answers to conform to the typical gender norms. To accommodate the scenarios where the participants may not always respond honestly, we need to extend the privacy mechanism of Warner's randomized response model. Such an extension results in the privacy mechanism of Figure 2. Here, γ_1 indicates the probability with which the participant answers honestly when the true answer is "Yes," and γ_2 indicates the probability with which the participant answers honestly when the true answer is "No."

Therefore, the sensitive question of interest can be asking a student if they have ever consumed an illegal substance. Similarly to the previous discussion, a student may skew their answer, depending on their understanding of the confidentiality preservation of the information collection procedure. Moreover, the probability that a student replies honestly when the true answer is "Yes" can be different from that of replying honestly when the true answer is "No." Observe that both γ_1 and γ_2 depend on the confidence of the respondent in the information collection procedure. Naturally, the respondent will want to ensure their own well-being. Suppose the respondent is not confident enough that their information will be kept confidential, and thus, it is unlikely that such a respondent will answer truthfully to the sensitive question of interest, and therefore, γ_1 will be small.

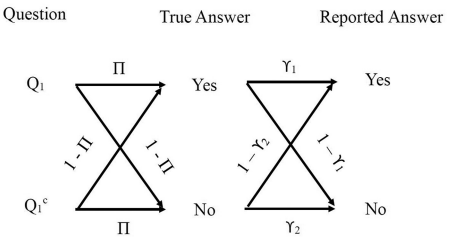


Fig. 2. Extension of privacy mechanism in Warner's randomized response model.

The students choose γ_1 and γ_2, depending on their own *confidence* in the information collection procedure and their desired privacy guarantees. In essence, they are concerned that a malicious entity may try to infer private information about a specific user (i.e., whether that specific student consumed any illegal substance or not). But the students should also be aware that a sophisticated adversary can estimate γ_1 and γ_2—call these estimates γ_1', γ_2'. Interestingly, if γ_1', γ_2' are far from the original values γ_1, γ_2, then such an adversary can reach the wrong conclusion about the targeted student, potentially blaming an innocent participant. Therefore, it is not only the true leakage of information that should concern each student, but also the potential mismatch leading to adversaries that are wrong, but confident in their estimations. The following optimization problem

captures these situations in the general case:

$$\min_{(\gamma_1,\gamma_2)} \max_{(\gamma_1',\gamma_2')} \text{privacy leakage metric,}$$
$$\text{such that } (\gamma_1', \gamma_2') \text{ is within a certain distance of } (\gamma_1, \gamma_2),$$
$$\text{and a utility constraint is maintained.}$$

The next important step is to define appropriate privacy leakage metrics. It is apparent from the context that the metrics should capture both the *confidence boost* an individual gets upon observing the disclosed information and how much actual information can be gained by collecting it. In this article, we have proposed distinct metrics that properly compute these various aspects of leakages. We have proposed several metrics that compute the confidence of the adversary after the information disclosure (average confidence boost and average subjective leakage) and the true leakage of the disclosed information (average objective leakage). Afterward, we have presented an optimization problem that will ensure the minimization of information leakage even in the worst-case scenario. Therefore, we believe the proposed extension of the privacy mechanism of Warner's response model has more usability in practice. The proposed metrics, along with the optimization problem formulation, will enable the user to design the privacy mechanism effectively such that the confidentiality of each user is preserved while maintaining the desired utility of the gathered information. The aforementioned situation can arise in various real-life instances, such as ensuring privacy during election audits, maintaining secure medical diagnoses, and protecting privacy in targeted advertising, among other examples.

3 PRELIMINARIES

3.1 Problem Setup

Let X be our private random variable and Y be the disclosed information about X. Throughout our article, we shall consider all the random variables are discrete with binary support sets.

We shall consider a three-party system for our model. The data owner, henceforth referred to as Alice, will design a privacy mechanism, $P_{Y|X}$, such that the randomized mechanism minimizes the information leakage between X and Y, while achieving a desired utility. However, neither the utility provider nor the adversary possesses the knowledge of the correct joint distribution, P_{XY}. They, nonetheless, know the initial distribution of X, P_X. Thus, their lack of knowledge of the joint distribution arises from the incomplete knowledge of the privacy mechanism, $P_{Y|X}$. Hereafter, we will refer to the utility provider as Bob and the adversary as Eve.

Therefore, our system setup is delineated as follows: Alice will disclose X through the privacy mechanism $P_{Y|X}$. We have represented the disclosed information as Y. Both Eve and Bob will collect several (X, Y) pairs and approximate the privacy mechanism as $Q_{Y|X}$ and $Q'_{Y|X}$, respectively. Afterward, Eve will try to guess X based on $Q_{Y|X}$.

3.2 System Matrices

The correct privacy mechanism, $P_{Y|X}$, and both the approximated privacy mechanisms, $Q_{Y|X}$ and $Q'_{Y|X}$, are given in Table 1.

The initial distribution of X, P_X, the marginal distribution of Y, P_Y, and the conditional probability $P_{X|Y}$ are given in Tables 2–4, respectively. Note that $P_Y = \sum_x P_{Y|X}.P_X$ and $P_{X|Y} = \dfrac{P_{Y|X}.P_X}{P_Y}$.

The marginal distribution of Y, approximated by Eve, Q_Y, and the corresponding conditional distribution, $Q_{X|Y}$, are given in Tables 5 and 7, respectively. Bob will perform a similar computation, and the resultant marginal distribution Q'_Y and the conditional distribution $Q'_{X|Y}$ are given in Tables 6 and 8, respectively.

Table 1. Privacy Mechanisms

(a) Original privacy mechanism, $P_{Y|X}$

$P_{Y\|X}$	Y = 0	Y = 1
X = 0	p_1	$1-p_1$
X = 1	$1-p_2$	p_2

(b) Approximated privacy mechanism of Eve, $Q_{Y|X}$

$Q_{Y\|X}$	Y = 0	Y = 1
X = 0	q_1	$1-q_1$
X = 1	$1-q_2$	q_2

(c) Approximated privacy mechanism of Bob, $Q'_{Y|X}$

$Q'_{Y\|X}$	Y = 0	Y = 1
X = 0	q'_1	$1-q'_1$
X = 1	$1-q'_2$	q'_2

Table 2. Initial Distribution of Private Variable, P_X

P_X	value
X = 0	p_0
X = 1	1 - p_0

Table 3. Marginal Distribution of Public Variable, P_Y

P_Y	value
Y = 0	$p_1p_0 + (1-p_2)(1-p_0)$
Y = 1	$p_0(1-p_1) + p_2(1-p_0)$

Table 4. Matrix for Conditional Distribution, $P_{X|Y}$

$P_{X\|Y}$	Y = 0	Y = 1
X = 0	$\frac{p_1p_0}{p_1p_0 + (1-p_2)(1-p_0)}$	$\frac{p_0(1-p_1)}{p_0(1-p_1) + p_2(1-p_0)}$
X = 1	$\frac{(1-p_2)(1-p_0)}{p_1p_0 + (1-p_2)(1-p_0)}$	$\frac{p_2(1-p_0)}{p_0(1-p_1) + p_2(1-p_0)}$

Table 5. Approximated Marginal Distribution of Public Variable by Eve, Q_Y

Q_Y	value
Y = 0	$q_1p_0 + (1-q_2)(1-p_0)$
Y = 1	$p_0(1-q_1) + q_2(1-p_0)$

Table 6. Approximated Marginal Distribution of Public Variable by Bob, Q'_Y

Q'_Y	value
Y = 0	$q'_1p_0 + (1-q'_2)(1-p_0)$
Y = 1	$p_0(1-q'_1) + q'_2(1-p_0)$

Table 7. Matrix for Conditional Distribution, $Q_{X|Y}$

$Q_{X\|Y}$	Y = 0	Y = 1
X = 0	$\frac{q_1p_0}{q_1p_0 + (1-q_2)(1-p_0)}$	$\frac{p_0(1-q_1)}{p_0(1-q_1) + q_2(1-p_0)}$
X = 1	$\frac{(1-q_2)(1-p_0)}{q_1p_0 + (1-q_2)(1-p_0)}$	$\frac{q_2(1-p_0)}{p_0(1-q_1) + q_2(1-p_0)}$

Table 8. Matrix for Conditional Distribution, $Q'_{X|Y}$

$Q'_{X\|Y}$	Y = 0	Y = 1
X = 0	$\frac{q'_1p_0}{q'_1p_0 + (1-q'_2)(1-p_0)}$	$\frac{p_0(1-q'_1)}{p_0(1-q'_1) + q'_2(1-p_0)}$
X = 1	$\frac{(1-q'_2)(1-p_0)}{q'_1p_0 + (1-q'_2)(1-p_0)}$	$\frac{q'_2(1-p_0)}{p_0(1-q'_1) + q'_2(1-p_0)}$

3.3 System Definitions

In this article, we are interested in an adversary who is interested in guessing the correct value of X in a single try (i.e., a one-try attack). Without any disclosed variable, Eve will analyze the initial distribution of X (P_X) to make a blind guess. The uncertainty of such a measure is represented by $H_\infty(X)$. Once Eve has access to Y, the uncertainty in guessing X is reduced, and the reduced uncertainty is represented by $H_\infty(X|Y)$. Therefore, an appropriate leakage measure should capture the difference between these two uncertainties, and *min-entropy leakage* metric correctly captures such a leakage. Thus, we shall begin by analyzing the definition of the *min-entropy leakage* metric and, afterward, modify the definition to represent the various measures of information leakage that may arise due to the incomplete statistical information of the privacy mechanism.

The definition of *min-entropy leakage (L)*, as given in Reference [3], is depicted by Equation (2). Here, we denote the support of random variable X as $\mathcal{X} = \{x_0, ..., x_{n-1}\}$, support of Y as $\mathcal{Y} = \{y_0, ..., y_{n-1}\}$, and $x_1^*(y) = \arg\max_{x \in \mathcal{X}} P_{X|Y}(x, y)$. We should note here that for our case $\mathcal{Y} = \mathcal{X} = \{0, 1\}$.

$$\begin{aligned} \mathrm{L}(P_{Y|X}) = I_\infty(X;Y) = H_\infty(X) - H_\infty(X|Y) &= H_\infty(X) + \log_2 \sum_{y \in \mathcal{Y}} P_Y(y) \max_{x \in \mathcal{X}} P_{X|Y}(x, y) \\ &= H_\infty(X) + \log_2 \sum_{y \in \mathcal{Y}} P_Y(y) P_{X|Y}(x_1^*(y)|y). \end{aligned} \tag{2}$$

Now, we shall modify Equation (2) to represent the confidence boost the adversary *expects* to achieve through the disclosure of the information. We have proposed the metric ***average subjective leakage (ASL)*** to compute this expected boost. As *ASL* is related to the expectation of the attacker, the definition considers both $Q_{Y|X}$ and Q_Y, as given by Equation (3). Here, $x_2^*(y) = \arg\max_{x \in \mathcal{X}} Q_{X|Y}(x, y)$. The adversary knows she should choose the most likely $x \in \mathcal{X}$, given her own understanding of the statistics. Such a choice is represented by $x_2^*(y)$, for each $y \in \mathcal{Y}$. Afterward, she computes $Q_{X|Y}(x_2^*(y)|y)$, and this measure represents her *belief* that she has made a correct guess regarding the value of X. Finally, she averages the gain with respect to her statistics, and we have called the resultant measure as *average subjective leakage*. Note that we explicitly include the dependence of *ASL* on the attacker's perceived privacy mechanism $Q_{Y|X}$:

$$\mathrm{ASL}(Q_{Y|X}) = H_\infty(X) + \log_2 \sum_{y \in \mathcal{Y}} Q_Y(y) \max_{x \in \mathcal{X}} Q_{X|Y}(x|y) = H_\infty(X) + \log_2 \sum_{y \in \mathcal{Y}} Q_Y(y) Q_{X|Y}(x_2^*(y)|y). \tag{3}$$

The ***Average confidence boost (ACB)*** measures the *actual* boost in confidence of the attacker due to disclosure of information. Accordingly, this metric relates $Q_{Y|X}$ to the correct marginal distribution P_Y, and consequently, we have:

$$\begin{aligned} \mathrm{ACB}(P_{Y|X}, Q_{Y|X}) &= H_\infty(X) + \log_2 \sum_{y \in \mathcal{Y}} P_Y(y) \max_{x \in \mathcal{X}} Q_{X|Y}(x|y) \\ &= H_\infty(X) + \log_2 \sum_{y \in \mathcal{Y}} P_Y(y) Q_{X|Y}(x_2^*(y)|y). \end{aligned} \tag{4}$$

Similar to *ASL*, the adversary again chooses the most likely $x \in \mathcal{X}$ from her own realization of the privacy mechanism. However, for computing *ACB*, we average such a gain with respect to the *true* statistics of the privacy mechanism. Therefore, *ACB* correctly measures the *true* confidence boost of the adversary upon observing Y.

Table 9. Summary of Notations

Symbol	Meaning
P_X	Original distribution of X
P_Y	Original distribution of Y
$P_{Y\|X}$	Original privacy mechanism
$Q_{Y\|X}$	Approximated privacy mechanism for attacker
Q_Y	Approximated distribution of Y for attacker
$Q'_{Y\|X}$	Approximated privacy mechanism for utility provider
Q'_Y	Approximated distribution of Y for utility provider
$x_1^*(y)$	$\arg\max_{x\in\mathcal{X}} P_{X\|Y}$
$x_2^*(y)$	$\arg\max_{x\in\mathcal{X}} Q_{X\|Y}$
$x_3^*(y)$	$\arg\max_{x\in\mathcal{X}} Q'_{X\|Y}$
u_{min}	Minimum utility requirement of the privacy mechanism $P_{Y\|X}$
δ	Maximum distance between $P_{Y\|X}$ and $Q_{Y\|X}$ (or $Q'_{Y\|X}$)

Finally, we have proposed the ***average objective leakage (AOL)*** to indicate how much information has *actually* been leaked by the system. As we are considering a three-party system, we shall have *AOL* for both the attacker and the utility provider. This metric can be measured by computing the leakage of the original privacy mechanism (i.e., $P_{Y|X}$) at the index in which the attacker/utility provider believes the system leaks maximum information. We know from the previous discussion that for the attacker, this index is $x_2^*(y)$. Similarly, for utility provider, we can denote $x_3^*(y) = \arg\max_{x\in\mathcal{X}} Q'_{X|Y}(x, y)$. Thus, the mathematical formulation of *AOL* is given by

$$\mathrm{AOL}(P_{Y|X}, Q_{Y|X}) = H_\infty(X) + \log_2 \sum_y P_Y(y) P_{X|Y}(x_i^*(y)|y). \tag{5}$$

(For indicating the attacker, we shall use $i = 2$, while $i = 3$ will be used to indicate the utility provider.)

Finally, Table 9 presents the summary of the notations used throughout the article. Note that, while defining each metric, we have considered the average case. Each of the definitions can easily be extended to represent the worst-case value where we take the maximum over all realizations of Y instead of summing over them. However, optimization becomes more challenging due to the discontinuous character of the function. For ease of explanation, we have only presented our analysis of the average case in this article.

3.4 Problem Statement

The primary objective of the privacy mechanism is to minimize the amount of information leakage to potential adversaries when they gain access to the disclosed information. However, we need to consider how much the *disclosed information* helps the utility provider as well. It is thus crucial to strike a delicate balance between safeguarding user privacy and providing the necessary data for the utility provider's analysis, and therefore the design of the privacy mechanism should ensure that the privatization of the sensitive data does not prevent the utility provider from achieving their desired utility. Consequently, the purpose of the design of the privacy mechanism is twofold: (i) *maximize* the utility of the disclosed data; and (ii) *minimize* the information leakage between X and Y.

To capture the information leakage minimization, we need to specify that $Q_{Y|X}$ is an approximate representation of $P_{Y|X}$, and accordingly, it is within a certain distance of $P_{Y|X}$. We have characterized this constraint by ensuring the maximum total variation distance between $P_{Y|X}$ and $Q_{Y|X}$ is δ, i.e., $d_{TV}(P_{Y|X}, Q_{Y|X}) \leq \delta$. This model corresponds to an attacker who successfully collects a set of (X, Y) pairs (say, by enlisting all of her friends) and estimates the privacy mechanism based on the collected set.

In our model, we have defined utility in two ways. A *generic utility provider* believes Y reflects an undisturbed value of X. We have also considered an *advanced utility provider* who, instead of taking Y at face value, performs statistical analysis to maximize the probability of having a correct guess.

Let us define utility for the *generic utility provider* first. If $X = 0$, then $P_{Y=0|X=0} = p_1$, and the probability that $X = Y$ can be computed as $p_1 p_0$. Similarly, when $X = 1$, the probability that $X = Y$ is $(1 - p_0)p_2$. To ensure that the utility is above a certain threshold, represented by u_{min}, we impose the utility constraint $\mathcal{U}(p_1, p_2) = p_1 p_0 + (1 - p_0)p_2 \geq u_{min}$. This is to be interpreted from the perspective of a utility provider that takes the disclosed information at face value. Such scenarios are currently the most common practice—for instance, with online social networks [2] or with the current recommender systems [38].

An *advanced utility provider* makes an educated guess, denoted by $x_3^*(y)$ (see Table 9), based on his approximated channel $Q'_{X|Y}$. From the perspective of utility, the data owner needs to make sure the *objective* leakage, computed at the index $x_3^*(y)$, is higher than a minimum threshold. It is to be noted that for an *advanced utility provider*, we take the worst-utility case as the minimum over $Q'(Y|X)$ and ensure that this worst utility is higher than u_{min}. Thus, the utility constraint is expressed as $\mathcal{U}(p_1, p_2) = \min_{Q'(Y|X)} AOL(Q'(Y|X)) \geq u_{min}$. As $Q'_{Y|X}$ is also an approximation of $P_{Y|X}$, we have $d_{TV}(P_{Y|X}, Q'_{Y|X}) \leq \delta$.

Now, let us formulate the optimization problem. For any of the proposed metrics, Alice never knows beforehand if the q_1 and q_2 values, chosen by Eve, are the ones that maximize that specific metric. She always needs to consider the worst-case scenario—which is usually the maximum with respect to q_1 and q_2—and find the parameters p_1 and p_2 that minimize this worst-case value of the metric. Thus, Alice needs to solve a minimax optimization problem.

For example, if Alice wants to devise a privacy mechanism that minimizes the worst-case value of *ACB*, then she will have the following optimization problem:

$$\begin{gathered}\min_{p_1,p_2} \max_{q_1,q_2} \mathrm{ACB}(P_{Y|X}, Q_{Y|X}),\\ \text{such that } \mathcal{U}(p_1, p_2) \geq u_{min},\\ \text{and } d_{TV}(P_{Y|X}, Q_{Y|X}) \leq \delta.\end{gathered}$$

3.5 Properties of the Proposed Metrics

Previously, we have introduced several metrics, defined each of them, and formulated the problem statement. Now, we shall present several properties of our proposed metrics.

Property 1. *$\max_{Q_{X|Y}} AOL$, where $Q_{X|Y}$ belongs to a probability simplex that includes $P_{X|Y}$, is always smaller than min-entropy leakage (L).*

Proof. Recall that $x_1^*(y)$ indicates the value of $x \in \mathcal{X}$ that maximizes $P_{X|Y}$ and $x_2^*(y)$ represents $x \in \mathcal{X}$ that maximizes $Q_{X|Y}$. When $\delta > 0$, $x_1^*(y)$ and $x_2^*(y)$ will refer to different values of $x \in \mathcal{X}$. As $x_1^*(y)$ always maximizes $P_{X|Y}$, $P_{X|Y}(x_1^*(y)|y)$ is always higher than $P_{X|Y}(x_2^*(y)|y)$. Therefore, by analyzing both Equations (2) and (5), we can say that maximum of average objective leakage will be lower than the min-entropy leakage.

Both $P_{Y|X}$ and $Q_{Y|X}$ refer to the same distribution when $\delta = 0$. Consequently, both $x_1^*(y)$ and $x_2^*(y)$ will represent the same value of $x \in \mathcal{X}$, and thus, average objective leakage will be same as the min-entropy leakage. □

Property 2. *Average objective leakage (AOL) can be negative.*

Proof. We know, $H_\infty(X) = -\log_2 \max_x P_X(x)$. Also, from Equation (5), we see that we need to specify whether the average objective leakage is computed for the adversary or the utility provider. The following proof shows the computation of the objective leakage for an adversary, and therefore, we use $i = 2$. Thus, we shall have the following:

$$\begin{aligned}\mathrm{AOL}(P_{Y|X}, Q_{Y|X}) &= -\log_2 \max_x P_X(x) + \log_2 \sum_y P_Y(y) P_{X|Y}(x_2^*(y)|y) \\ &= \log_2 \frac{\sum_y P_Y(y) P_{X|Y}(x_2^*(y)|y)}{\max_x P_X(x)} = \log_2 \frac{\sum_y P_Y(y) P_{X|Y}(x_2^*(y)|y)}{\max_x \sum_y P_{XY}(x,y)} \\ &\overset{(*)}{\geq} \log_2 \frac{\sum_y P_Y(y) P_{X|Y}(x_2^*(y)|y)}{\sum_y \max_x P_{XY}(x,y)} = \log_2 \frac{\sum_y P_Y(y) P_{X|Y}(x_2^*(y)|y)}{\sum_y P_Y(y) P_{X|Y}(x_1^*(y))}.\end{aligned}$$

Here, (*) holds as $\max_x \sum_y P_{XY}(x,y) \leq \sum_y \max_x P_{XY}(x,y)$. As $x_1^*(y)$ always maximizes $P_{X|Y}$, the value of the ratio $\frac{\sum_y P_Y(y) P_{X|Y}(x_2^*(y)|y)}{\sum_y P_Y(y) P_{X|Y}(x_1^*(y))}$ is less than 1. Therefore, the lower limit of the average objective leakage is negative.

Note that the lower limit can be achieved when the inequality $\max_x \sum_y P_{XY}(x,y) \leq \sum_y \max_x P_{XY}(x,y)$ holds with equality. In this case, we get the following:

$$\max_x \sum_y P_{XY}(x,y) = \sum_y \max_x P_{XY}(x,y) \Leftrightarrow \max_x \sum_y P_Y(y) P_{X|Y}(x|y) = \sum_y P_Y(y) \max_x P_{X|Y}(x|y). \tag{6}$$

Notice the left-hand side of Equation (6). The maximization can occur either at $X = 0$ or $X = 1$. Let us assume that the maximization occurs at $X = 0$. In that case, the left-hand side of Equation (6) would be $P_Y(0)\, P_{X|Y}(x = 0|y = 0) + P_Y(1)\, P_{X|Y}(x = 0|y = 1)$. Thus, if we assume that $P_{X=0|Y} > P_{X=1|Y}$ for both $Y = 0$ and $Y = 1$, then we see that Equation (6) holds. Similarly, Equation (6) will also hold when $P_{X=1|Y} > P_{X=0|Y}$ for both $Y = 0$ and $Y = 1$. Thus, when the same value of $x \in \mathcal{X}$ maximizes both $P_{X|Y=0}$ and $P_{X|Y=1}$, we achieve Equation (6), and consequently, the average objective leakage becomes negative. □

Property 3. $\max_{Q_{X|Y}}$ *ACB, where* $Q_{X|Y}$ *belongs to a probability simplex that includes* $P_{X|Y}$*, is always larger than min-entropy leakage (L).*

Proof. From the definitions of both average confidence boost (shown in Equation (4)) and min-entropy leakage (shown in Equation (2)), we get the following:

$$\mathrm{ACB}(P_{Y|X}, Q_{Y|X}) - \mathrm{L}(P_{Y|X}) = \log_2 \frac{\sum_y P_Y(y) Q_{X|Y}(x_2^*(y)|y)}{\sum_y P_Y(y) P_{X|Y}(x_1^*(y)|y)}. \tag{7}$$

When $\delta > 0$, the measure space of $Q_{X|Y}$ is larger than the measure space of $P_{X|Y}$ and also includes the measure space of $P_{X|Y}$. Moreover, $x_2^*(y)$ maximizes $Q_{X|Y}$ and $x_1^*(y)$ maximizes $P_{X|Y}$. As the measure space is larger, we have $\max_{Q_{X|Y}} Q_{X|Y}(x_2^*(y)|y) > P_{X|Y}(x_1^*(y)|y)$. Thus, the difference will be greater than zero. The difference between average confidence boost and min-entropy leakage will be zero when both $P_{Y|X}$ and $Q_{Y|X}$ refer to the same distribution ($\delta = 0$).

We know that min-entropy leakage is always non-negative [20]. Thus, $\max_{Q_{X|Y}}$ *ACB is always non-negative.* □

Property 4. $\max_{Q_{X|Y}}$ *ASL, where* $Q_{X|Y}$ *belongs to a probability simplex that includes* $P_{X|Y}$*, is always larger than min-entropy leakage (L).*

Proof. From the definitions of both average subjective leakage (shown in Equation (3)) and min-entropy leakage (shown in Equation (2)), we get the following:

$$\mathrm{ASL}(Q_{Y|X}) - \mathrm{L}(P_{Y|X}) = \log_2 \frac{\sum_y Q_Y(y) Q_{X|Y}(x_2^*(y)|y)}{\sum_y P_Y(y) P_{X|Y}(x_1^*(y)|y)} = \log_2 \frac{\sum_y \max_{x\in\mathcal{X}} Q_{XY}(x,y)}{\sum_y \max_{x\in\mathcal{X}} P_{XY}(x,y)}. \quad (8)$$

When $\delta > 0$, $\max_{Q_{XY}} \max_{x\in\mathcal{X}} Q_{XY}(x,y)$ is higher than the value of $\max_{x\in\mathcal{X}} P_{XY}(x,y)$, as the measure space of Q_{XY} is larger and includes the measure space of P_{XY}. Therefore, the difference in Equation (8) is higher than zero. We shall have ASL = L when $\delta = 0$, as both $P_{XY}(x,y)$ and $Q_{XY}(x,y)$ will then indicate the same distribution.

As min-entropy leakage is non-negative [20], $\max_{Q_{X|Y}}$ *ASL is always non-negative.* □

4 FEASIBLE REGION ANALYSIS

4.1 Feasible Region of Probability Tuples (p_1, p_2) and (q_1, q_2)

We know that $x_1^*(y) = \arg\max_{x\in\mathcal{X}} P_{X|Y}$, and therefore, various combinations of (p_1, p_2) will result in different $P_{X|Y}$ (distribution is shown in Table 4) and consequently lead to different values of $x_1^*(y)$. Similarly, different combinations of (q_1, q_2) will result in different values of $x_2^*(y)$.

Let us assume $p_0 = 0.5$. When we claim that the value of $x_1^*(y)$ at a specific index is 0, it means that for a given value $Y = y$, $P_{X=0|Y=y}$ is higher than $P_{X=1|Y=y}$. Suppose for a specific combination of (p_1, p_2), we have $p_1 > (1 - p_2)$. As $p_0 = 0.5$, this inequality also means $p_1 p_0 > (1 - p_2)(1 - p_0)$. From Table 4, we see that the inequality $p_1 p_0 > (1 - p_2)(1 - p_0)$ means $P_{X=0|Y=0}$ is higher than $P_{X=1|Y=0}$. Similarly, $p_1 > (1 - p_2)$ also conveys $p_0(1 - p_1) < p_2(1 - p_0)$ when $p_0 = 0.5$, and as a result, points out that $P_{X=1|Y=1}$ is higher than $P_{X=0|Y=1}$. Thus, for $Y = 0$, $P_{X=0|Y=0}$ is higher than $P_{X=1|Y=0}$, and when $Y = 1$, $P_{X=1|Y=1}$ is higher than $P_{X=0|Y=1}$. These two conditions on $P_{X|Y}$ are equivalent to $x_1^*(y) = [0\ 1]$—henceforth, when writing $x_i^*(y) = [a\ b]$, we mean that $x_i^*(y = 0) = a$ and $x_i^*(y = 1) = b$, for any $a, b \in \{0, 1\}$.

If we consider $x_1^*(y) = [0\ 0]$, then we shall have conditions (9) and (10).

$$p_1 p_0 > (1 - p_2)(1 - p_0), \quad (9)$$

$$p_0(1 - p_1) > p_2(1 - p_0). \quad (10)$$

When we use the value of $p_0 = 0.5$, condition (9) results in $p_1 + p_2 > 1$ and condition (10) corresponds to $p_1 + p_2 < 1$. Therefore, when $p_0 = 0.5$, $x_1^*(y) = [0\ 0]$ is not feasible.

Setting $x_1^*(y) = [0\ 1]$ corresponds to the two inequalities shown in conditions (11) and (12).

$$p_1 p_0 > (1 - p_2)(1 - p_0), \quad (11)$$

$$p_2(1 - p_0) > p_0(1 - p_1). \quad (12)$$

When we put $p_0 = 0.5$, both conditions (11) and (12) result in $p_1 + p_2 > 1$. Thus, $x_1^*(y) = [0\ 1]$ is feasible for this value of p_0. Using an identical reasoning, it is possible to show that $x_1^*(y) = [1\ 0]$

Table 10. Difference between $P_{Y|X}$ and $Q_{Y|X}$

$P_{Y\|X}$ - $Q_{Y\|X}$	Y = 0	Y = 1
X = 0	$p_1 - q_1$	$q_1 - p_1$
X = 1	$q_2 - p_2$	$p_2 - q_2$

is also feasible, whereas $x_1^*(y) = [1\ 1]$ is not. Similar analysis can be performed to show that $x_1^*(y) = [1\ 1]$ is not feasible when $p_0 > 0.5$, and $x_1^*(y) = [0\ 0]$ is not feasible when $p_0 < 0.5$.

Similarly, when $p_0 = 0.5$ and $x_2^*(y) = [0\ 0]$, we get the following two conditions (based on Table 7):

$$q_1 p_0 > (1 - q_2)(1 - p_0), \tag{13}$$

$$p_0(1 - q_1) > q_2(1 - p_0). \tag{14}$$

As $p_0 = 0.5$, condition (13) results in $q_1 + q_2 > 1$, and condition (14) leads to $q_1 + q_2 < 1$. Therefore, $p_0 = 0.5$ and $x_2^*(y) = [0\ 0]$ is not feasible. Doing further analysis shows that feasible regions for various possible values of (q_1, q_2) and p_0 are identical to those feasible regions for various possible values of (p_1, p_2) and p_0.

4.2 Feasible Region for Combinations of (p_1, p_2) and (q_1, q_2)

With feasible regions for both (p_1, p_2) and (q_1, q_2) already explained, it is now time to analyze the feasible regions for the combination of (p_1, p_2) and (q_1, q_2). Let us presume that a specific value of (p_1, p_2) results in $x_1^*(y) = [1\ 0]$, and a particular (q_1, q_2) leads to $x_2^*(y) = [0\ 1]$. Now, we want to find for which values of p_0 both $x_1^*(y) = [1\ 0]$ and $x_2^*(y) = [0\ 1]$ hold simultaneously.

Here, δ will play a significant role. As $d_{TV}(P_{Y|X}, Q_{Y|X}) \leq \delta$, we get $||P_{Y|X} - Q_{Y|X}||_1 \leq 2\delta$. Based on the privacy mechanisms of Table 1, we get the difference in Table 10.

We know that for any given matrix A, the ℓ_1 norm can be calculated as $||A||_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|$ [42]. If we use this formula to compute the ℓ_1 norm of $P_{Y|X} - Q_{Y|X}$, then we get both $(p_1 - q_1) - (p_2 - q_2) \leq 2\delta$ and $(p_2 - q_2) - (p_1 - q_1) \leq 2\delta$. These two inequalities result in the following bounds for both q_1 and q_2:

$$p_1 - \delta \leq q_1 \leq p_1 + \delta,$$
$$p_2 - \delta \leq q_2 \leq p_2 + \delta.$$

Similarly, we get $(1 - p_1) - \delta \leq 1 - q_1 \leq (1 - p_1) + \delta$ and $(1 - p_2) - \delta \leq 1 - q_2 \leq (1 - p_2) + \delta$.

Now, let us assume a certain combination of (p_1, p_2) results in $x_1^*(y) = [1\ 0]$ when $p_0 = 0.5$, represented by the *Green* region in Figure 3, and the (q_1, q_2) tuple can lie anywhere within the δ ball of this value of (p_1, p_2) tuple. If (p_1, p_2) lies far away from the *Green-Cyan* boundary and the value of δ is small, then (q_1, q_2) will also be in the *Green* region. Thus, $x_1^*(y) = [1\ 0]$ will also convey $x_2^*(y) = [1\ 0]$. However, if (p_1, p_2) lies close to the boundary, then it is possible for (q_1, q_2) to fall into the *Cyan* region as, in this particular case, the δ ball will cover the *Cyan* region as well. Thus, it would be possible to have $x_2^*(y) = [0\ 1]$ even though $x_1^*(y) = [1\ 0]$. Therefore, we need to divide the *Green* region into two sub-regions. The first sub-region (*Green*) will indicate when $x_1^*(y) = [1\ 0]$ and $x_2^*(y) = [1\ 0]$. The second sub-region (*Magenta*) will indicate $x_1^*(y) = [1\ 0]$, and $x_2^*(y)$ can either be $[1\ 0]$ or $[0\ 1]$. Figure 4 shows the possible sub-regions for various combinations of (p_1, p_2) and (q_1, q_2) under possible values of p_0. The meaning of each sub-region is given in Table 11. *Note that we achieve several sub-regions by analyzing several inequalities, such as* $p_1 p_0 > (1 - p_2)(1 - p_0)$*, along with others. Whenever we replace the inequalities with an equal constraint, for example,* $p_1 p_0 = (1 - p_2)(1 - p_0)$*, we get the boundary between two different sub-regions.*

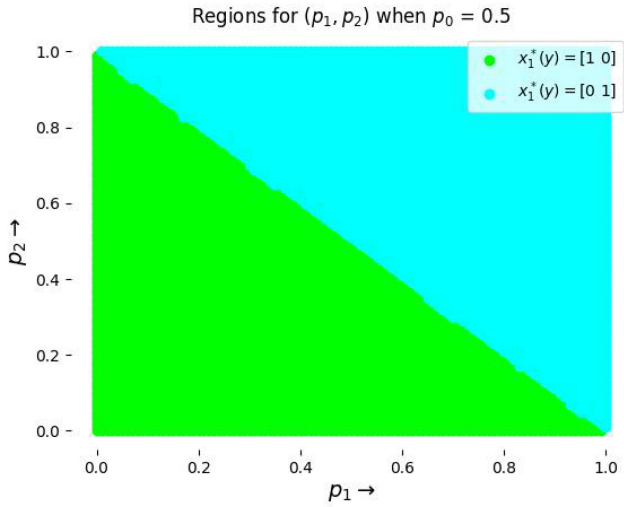
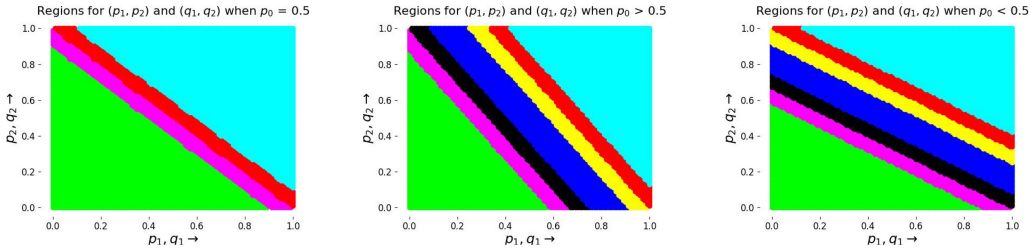


Fig. 3. Possible regions of (p_1, p_2) when $p_0 = 0.5$.

Table 11. Possible Sub-regions for Combinations of (p_1, p_2) and (q_1, q_2)

Sub-region	$p_0 = 0.5$	$p_0 > 0.5$	$p_0 < 0.5$
Green	$x_1^*(y) = [1\ 0], x_2^*(y) = [1\ 0]$	$x_1^*(y) = [1\ 0], x_2^*(y) = [1\ 0]$	$x_1^*(y) = [1\ 0], x_2^*(y) = [1\ 0]$
Magenta	$x_1^*(y) = [1\ 0], x_2^*(y) = [1\ 0]$ or $[0\ 1]$	$x_1^*(y) = [1\ 0], x_2^*(y) = [1\ 0]$ or $[0\ 0]$	$x_1^*(y) = [1\ 0], x_2^*(y) = [1\ 0]$ or $[1\ 1]$
Black	Not Feasible	$x_1^*(y) = [0\ 0], x_2^*(y) = [1\ 0]$ or $[0\ 0]$	$x_1^*(y) = [1\ 1], x_2^*(y) = [1\ 0]$ or $[1\ 1]$
Blue	Not Feasible	$x_1^*(y) = [0\ 0], x_2^*(y) = [0\ 0]$	$x_1^*(y) = [1\ 1], x_2^*(y) = [1\ 1]$
Yellow	Not Feasible	$x_1^*(y) = [0\ 0], x_2^*(y) = [0\ 1]$ or $[0\ 0]$	$x_1^*(y) = [1\ 1], x_2^*(y) = [0\ 1]$ or $[1\ 1]$
Red	$x_1^*(y) = [0\ 1], x_2^*(y) = [1\ 0]$ or $[0\ 1]$	$x_1^*(y) = [0\ 1], x_2^*(y) = [0\ 0]$ or $[0\ 1]$	$x_1^*(y) = [0\ 1], x_2^*(y) = [0\ 1]$ or $[1\ 1]$
Cyan	$x_1^*(y) = [0\ 1]$ and $x_2^*(y) = [0\ 1]$	$x_1^*(y) = [0\ 1]$ and $x_2^*(y) = [0\ 1]$	$x_1^*(y) = [0\ 1]$ and $x_2^*(y) = [0\ 1]$



(a) $p_0 = 0.5$ (b) $p_0 > 0.5 (p_0 = 0.6)$ (c) $p_0 < 0.5 (p_0 = 0.4)$

Fig. 4. Possible sub-regions for various combinations of (p_1, p_2) and (q_1, q_2) under possible p_0.

5 ACB OPTIMIZATION

We have discussed the optimization of *ASL* in Reference [47], albeit referring to the measure as ***maximal subjective leakage (MSL)***. In this article, we shall review the optimization of *ACB*, which is formulated as follows:

$$\begin{gathered}\min_{p_1,p_2} \max_{q_1,q_2} \mathrm{ACB}(P_{Y|X}, Q_{Y|X}),\\ \text{such that } \mathcal{U}(p_1,p_2) \geq u_{min},\\ \text{and } d_{TV}(P_{Y|X}, Q_{Y|X}) \leq \delta.\end{gathered}$$

We can define the utility measure in two ways. If Bob (the utility provider) wants to take the output at the face value, then the utility constraint will be expressed as $\mathcal{U}(p_1,\ p_2) = p_1p_0 + (1 - p_0)p_2 \geq u_{min}$. Alternatively, Bob can also perform some statistical analysis to make an educated guess, and in that scenario, the utility constraint will be $\mathcal{U}(p_1, p_2) = \min_{Q'(Y|X)} \text{AOL}(Q'(Y|X)) \geq u_{min}$. The details can be found in Section 3.4.

Straightaway, this optimization problem can be divided into two components: (i) finding q_1 and q_2 that maximize *ACB* (a maximization problem) and (ii) finding p_1 and p_2 that minimize this maximized value of *ACB* (a minimization problem).

5.1 The Maximization Problem

In this subsection, we shall determine q_1 and q_2, in terms of p_1 and p_2, that maximize the *ACB*. Thus, we shall review the following problem:

$$\begin{gathered}\max_{q_1,q_2} \text{ACB}(P_{Y|X}, Q_{Y|X}),\\ \text{such that } \mathcal{U}(p_1,\ p_2) \geq u_{min},\\ \text{and } d_{TV}(P_{Y|X}, Q_{Y|X}) \leq \delta.\end{gathered}$$

Let us consider the *Cyan* sub-region of Figure 4(a) first, i.e., we have $x_1^*(y) = [0\ 1]$, $x_2^*(y) = [0\ 1]$, and $p_0 = 0.5$. The value of ACB, for this value of $x_2^*(y)$, is given by Equation (15). Here, $c_1 = p_1p_0 + (1-p_2)(1-p_0)$ and $c_2 = p_0(1-p_1) + p_2(1-p_0)$. Note that the following analysis is applicable for the cyan sub-region of both Figures 4(b) and 4(c) as well:

$$\text{ACB}(P_{Y|X}, Q_{Y|X}) = H_\infty(X) + \log_2\left(\frac{c_1q_1p_0}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2q_2(1-p_0)}{p_0(1-q_1)+q_2(1-p_0)}\right). \tag{15}$$

Let us denote $f = \frac{c_1q_1p_0}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2q_2(1-p_0)}{p_0(1-q_1)+q_2(1-p_0)}$. Now, differentiating f, with respect to both q_1 and q_2, we get Equations (16) and (17), respectively.

$$\frac{\partial f}{\partial q_1} = \frac{c_1p_0(1-q_2)(1-p_0)}{(q_1p_0+(1-q_2)(1-p_0))^2} + \frac{c_2p_0q_2(1-p_0)}{(p_0(1-q_1)+q_2(1-p_0))^2} \tag{16}$$

$$\frac{\partial f}{\partial q_2} = \frac{c_1q_1p_0(1-p_0)}{(q_1p_0+(1-q_2)(1-p_0))^2} + \frac{c_2(1-p_0)p_0(1-q_1)}{(p_0(1-q_1)+q_2(1-p_0))^2} \tag{17}$$

Moreover, we know $p_1 - \delta \leq q_1 \leq p_1 + \delta$ and $p_2 - \delta \leq q_2 \leq p_2 + \delta$. It is clear that both the partial derivatives are positive for all values of q_1 and q_2 in these intervals. Thus, ACB is an increasing function of both q_1 and q_2, and the maximization will occur when $q_1 = p_1 + \delta$, and $q_2 = p_2 + \delta$.

From Table 11, we can see that if $x_2^*(y) = [0\ 1]$, then either the *Red* or the *Cyan* sub-region is feasible. Let us analyze Figure 4 and consider any random (p_1, p_2) tuple in the *Red* sub-region. Depending on the value of δ, the resultant tuple, that maximize ACB ($q_1 = p_1 + \delta, q_2 = p_2 + \delta$), can either be in the *Red* or the *Cyan* sub-region. Gradual increase in the value of δ will push the (q_1, q_2) tuple from the *Red* to the *Cyan* sub-region. Similarly, if the initial (p_1, p_2) tuple is in the *Cyan* sub-region, then the resultant (q_1, q_2) tuple will be further in the *Cyan* sub-region. Thus, for all possible values of δ, (q_1, q_2) tuple, which maximizes ACB, will either occupy the *Red* or the *Cyan* sub-region for which we have $x_2^*(y) = [0\ 1]$.

Afterward, let us analyze the *Green* sub-region of Figure 4. For this sub-region, we have $x_2^*(y) = [1\ 0]$ and Equation (18) depicts the value of ACB for this value of $x_2^*(y)$. If we denote the term inside the logarithm as g (i.e., $g = \frac{c_1(1-q_2)(1-p_0)}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)}$) and compute partial derivatives with respect to q_1 and q_2, then we get Equations (19) and (20), respectively.

$$\text{ACB}(P_{Y|X}, Q_{Y|X}) = H_\infty(X) + \log_2\left(\frac{c_1(1-q_2)(1-p_0)}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)}\right) \tag{18}$$

$$\frac{\partial g}{\partial q_1} = -\frac{c_1p_0(1-q_2)(1-p_0)}{(q_1p_0+(1-q_2)(1-p_0))^2} - \frac{c_2p_0q_2(1-p_0)}{(p_0(1-q_1)+q_2(1-p_0))^2} \tag{19}$$

$$\frac{\partial g}{\partial q_2} = -\frac{c_1 p_0(1-p_0)q_1}{(q_1 p_0+(1-q_2)(1-p_0))^2} - \frac{c_2 p_0(1-q_1)(1-p_0)}{(p_0(1-q_1)+q_2(1-p_0))^2} \tag{20}$$

As both the partial derivatives are negative for all possible values of q_1 and q_2, the maximization will occur when $q_1 = p_1 - \delta$ and $q_2 = p_2 - \delta$. Further examination of the feasible region shows that sub-regions *Green* and *Magenta* result in $x_2^*(y) = [1\ 0]$. For any value of tuple (p_1, p_2), which resides in either the *Green* or the *Magenta* sub-region, choice of $(q_1 = p_1 - \delta, q_2 = p_2 - \delta)$ tuple will always result in $x_2^*(y) = [1\ 0]$. Therefore, $q_1 = p_1 - \delta$, and $q_2 = p_2 - \delta$ maximize ACB when $x_2^*(y) = [1\ 0]$.

5.2 The Minimization Problem

Previously, we have found the values of q_1 and q_2 that maximize *ACB*. Now, we want to compute p_1 and p_2 that minimize this maximization of *ACB*. As a demonstration, we will explain the optimization problem formulation when $p_0 = 0.5$. The details of the optimization problems, for each possible sub-region, can be found in Table 12.

The case of $p_0 = 0.5$. We have seen from Figure 4 that there are four possible sub-regions when $p_0 = 0.5$. Straightaway, we will formulate the constraints that need to hold for each of the sub-regions.

Let us consider the *Green* sub-region first. For this sub-region, we have $x_1^*(y) = [1\ 0]$ and $x_2^*(y) = [1\ 0]$. The inequalities, corresponding to $x_1^*(y) = [1\ 0]$, are given by Equations (21) and (22).

$$p_1 p_0 < (1 - p_2)(1 - p_0) \tag{21}$$

$$p_2(1 - p_0) < p_0(1 - p_1) \tag{22}$$

We have seen previously that $q_1 = p_1 - \delta$ and $q_2 = p_2 - \delta$ maximize *ACB* when $x_2^*(y) = [1\ 0]$. Using these values of q_1 and q_2, we get Equations (23) and (24) for $x_2^*(y) = [1\ 0]$.

$$\begin{aligned} & q_1 p_0 < (1 - q_2)(1 - p_0) \\ \implies & p_0(p_1 - \delta) < (1 - p_0)(1 - p_2 + \delta) \\ \implies & p_0 p_1 < (1 - p_2)(1 - p_0) + \delta \end{aligned} \tag{23}$$

$$\begin{aligned} & q_2(1 - p_0) < p_0(1 - q_1) \\ \implies & (p_2 - \delta)(1 - p_0) < p_0(1 - p_1 + \delta) \\ \implies & p_2(1 - p_0) < p_0(1 - p_1) + \delta \end{aligned} \tag{24}$$

It is clear that if Equation (21) is maintained, then Equation (23) will also hold. Similarly, fulfilling Equation (22) will imply Equation (24) is also met. As a result, to obtain $x_1^*(y) = [1\ 0]$ and $x_2^*(y) = [1\ 0]$, we need to ensure that Equations (21) and (22) hold simultaneously.

Note that for the *Green* sub-region, we need to make sure that no (p_1, p_2) tuple results in $x_2^*(y) = [0\ 1]$. We know that $q_1 = p_1 + \delta$ and $q_2 = p_2 + \delta$ maximize *ACB* when $x_2^*(y) = [0\ 1]$. Using these values of q_1 and q_2, we get Equations (25) and (26) for $x_2^*(y) = [0\ 1]$.

$$\begin{aligned} & p_0 q_1 > (1 - q_2)(1 - p_0) \\ \implies & p_0(p_1 + \delta) > (1 - p_2 - \delta)(1 - p_0) \\ \implies & p_0 p_1 > (1 - p_2)(1 - p_0) - \delta \end{aligned} \tag{25}$$

$$\begin{aligned} & q_2(1 - p_0) > p_0(1 - q_1) \\ \implies & (p_2 + \delta)(1 - p_0) > p_0(1 - p_1 - \delta) \\ \implies & p_2(1 - p_0) > p_0(1 - p_1) - \delta \end{aligned} \tag{26}$$

Table 12. Optimization Problems of Worst-case *ACB* Minimization

Sub-region	$p_0 = 0.5$	$p_0 > 0.5$	$p_0 < 0.5$
Green	$\min_{p_1,p_2} \frac{c_1(1-p_0)(1-p_2+\delta)}{p_0(p_1-\delta)+(1-p_0)(1-p_2+\delta)} + \frac{c_2p_0(1-p_1+\delta)}{p_0(1-p_1+\delta)+(1-p_0)(p_2-\delta)}$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_0p_1 < (1-p_2)(1-p_0)-\delta$ $p_2(1-p_0) < p_0(1-p_1)-\delta$	$\min_{p_1,p_2} \frac{c_1(1-p_0)(1-p_2+\delta)}{p_0(p_1-\delta)+(1-p_0)(1-p_2+\delta)} + \frac{c_2p_0(1-p_1+\delta)}{p_0(1-p_1+\delta)+(1-p_0)(p_2-\delta)}$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_0p_1 < (1-p_2)(1-p_0)-\delta$ $p_2(1-p_0) < p_0(1-p_1)-\delta$	$\min_{p_1,p_2} \frac{c_1(1-p_0)(1-p_2+\delta)}{p_0(p_1-\delta)+(1-p_0)(1-p_2+\delta)} + \frac{c_2p_0(1-p_1+\delta)}{p_0(1-p_1+\delta)+(1-p_0)(p_2-\delta)}$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_0p_1 < (1-p_2)(1-p_0)-\delta$ $p_2(1-p_0) < p_0(1-p_1)-\delta$
Magenta	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1(1-p_0)(1-q_2)}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2(1-p_0)q_2}{p_0(1-q_1)+(1-p_0)q_2})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_1p_0 < (1-p_2)(1-p_0)$ $p_2(1-p_0) < p_0(1-p_1)$ $p_0p_1 > (1-p_2)(1-p_0)-\delta$ $p_2(1-p_0) > p_0(1-p_1)-\delta$	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1(1-p_0)(1-q_2)}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1q_1p_0}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_1p_0 < (1-p_2)(1-p_0)$ $p_0p_1 > (1-p_2)(1-p_0)-\delta$ $p_2(1-p_0) < p_0(1-p_1)-\delta$	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1(1-p_0)(1-q_2)}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1(1-q_2)(1-p_0)}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2q_2(1-p_0)}{p_0(1-q_1)+q_2(1-p_0)})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_1p_0 < (1-p_2)(1-p_0)$ $p_2(1-p_0) < p_0(1-p_1)$ $p_2(1-p_0) > p_0(1-p_1)-\delta$
Black	Not Feasible	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1(1-p_0)(1-q_2)}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1q_1p_0}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_1p_0 > (1-p_2)(1-p_0)$ $p_0p_1 < (1-p_2)(1-p_0)+\delta$ $p_2(1-p_0) < p_0(1-p_1)-\delta$	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1(1-p_0)(1-q_2)}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1(1-q_2)(1-p_0)}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2q_2(1-p_0)}{p_0(1-q_1)+q_2(1-p_0)})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $(1-p_2)(1-p_0) > p_1p_0$ $p_2(1-p_0) > p_0(1-p_1)$ $p_2(1-p_0) < p_0(1-p_1)+\delta$
Blue	Not feasible	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_1p_0 > (1-p_2)(1-p_0)+\delta$ $p_0(1-p_1) > p_2(1-p_0)+\delta$	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_1p_0 < (1-p_2)(1-p_0)-\delta$ $p_2(1-p_0) > p_0(1-p_1)+\delta$
Yellow	Not Feasible	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2(1-p_0)q_2}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1q_1p_0}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_0(1-p_1) > p_2(1-p_0)$ $p_0(1-p_1) < p_2(1-p_0)+\delta$ $p_0p_1 > (1-p_2)(1-p_0)$	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2(1-p_0)q_2}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1(1-q_2)(1-p_0)}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2q_2(1-p_0)}{p_0(1-q_1)+q_2(1-p_0)})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $(1-p_2)(1-p_0) > p_1p_0$ $p_2(1-p_0) > p_0(1-p_1)$ $p_0p_1 > (1-p_2)(1-p_0)-\delta$
Red	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1(1-p_0)(1-q_2)}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2(1-p_0)q_2}{p_0(1-q_1)+(1-p_0)q_2})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $(1-p_2)(1-p_0) < p_1p_0$ $p_0(1-p_1) < p_2(1-p_0)$ $p_0p_1 < (1-p_2)(1-p_0)+\delta$ $p_2(1-p_0) < p_0(1-p_1)+\delta$	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2(1-p_0)q_2}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1q_1p_0}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_0p_1 > (1-p_2)(1-p_0)+\delta$ $p_2(1-p_0) > p_0(1-p_1)$ $p_2(1-p_0) < p_0(1-p_1)+\delta$	$\min_{p_1,p_2}\max_{q_1,q_2}(\frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2(1-p_0)q_2}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1(1-q_2)(1-p_0)}{q_1p_0+(1-q_2)(1-p_0)} + \frac{c_2q_2(1-p_0)}{p_0(1-q_1)+q_2(1-p_0)})$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_1p_0 > (1-p_2)(1-p_0)$ $p_2(1-p_0) > p_0(1-p_1)+\delta$ $p_1p_0 < (1-p_2)(1-p_0)+\delta$
Cyan	$\min_{p_1,p_2} \frac{c_1p_0(p_1+\delta)}{p_0(p_1+\delta)+(1-p_0)(1-p_2-\delta)} + \frac{c_2(1-p_0)(p_2+\delta)}{p_0(1-p_1-\delta)+(1-p_0)(p_2+\delta)}$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_0p_1 > (1-p_2)(1-p_0)+\delta$ $p_2(1-p_0) > p_0(1-p_1)+\delta$	$\min_{p_1,p_2} \frac{c_1p_0(p_1+\delta)}{p_0(p_1+\delta)+(1-p_0)(1-p_2-\delta)} + \frac{c_2(1-p_0)(p_2+\delta)}{p_0(1-p_1-\delta)+(1-p_0)(p_2+\delta)}$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_0p_1 > (1-p_2)(1-p_0)+\delta$ $p_2(1-p_0) > p_0(1-p_1)+\delta$	$\min_{p_1,p_2} \frac{c_1p_0(p_1+\delta)}{p_0(p_1+\delta)+(1-p_0)(1-p_2-\delta)} + \frac{c_2(1-p_0)(p_2+\delta)}{p_0(1-p_1-\delta)+(1-p_0)(p_2+\delta)}$ such that $\mathcal{U}(p_1,\ p_2) \geq u_{min}$ $p_0p_1 > (1-p_2)(1-p_0)+\delta$ $p_2(1-p_0) > p_0(1-p_1)+\delta$

Here, $c_1 = p_1p_0 + (1-p_2)(1-p_0)$, and $c_2 = p_0(1-p_1) + p_2(1-p_0)$.

Observe that Equation (25) reveals if $p_0p_1 > (1-p_2)(1-p_0) - \delta$ holds, then there is a possibility that $x_2^*(y)$ can be [0 1]. Similarly, the inequality $p_2(1-p_0) > p_0(1-p_1)-\delta$ may result in $x_2^*(y) = [0\ 1]$. Therefore, for the *Green* sub-region, for which we have $x_2^*(y) = [1\ 0]$, we need to make sure that both inequalities $p_0p_1 < (1-p_2)(1-p_0) - \delta$ and $p_2(1-p_0) < p_0(1-p_1) - \delta$ hold simultaneously.

Now, we will explain the method of finding optimum p_1 and p_2 for the minimization problem. When $x_2^*(y) = [1\ 0]$, *ACB* is given by Equation (18). Recall that $c_1 = p_1p_0 + (1-p_2)(1-p_0)$, and $c_2 = p_0(1-p_1) + p_2(1-p_0)$. Using $q_1 = p_1 - \delta$ and $q_2 = p_2 - \delta$ in Equation (18), we get Equation (27).

$$\text{ACB}(P_{Y|X}, Q_{Y|X}) = H_\infty(X) + \log_2\left(\frac{c_1(1-p_0)(1-p_2+\delta)}{p_0(p_1-\delta)+(1-p_0)(1-p_2+\delta)} + \frac{c_2p_0(1-p_1+\delta)}{p_0(1-p_1+\delta)+(1-p_0)(p_2-\delta)}\right) \tag{27}$$

Thus, for the *Green* sub-region, the minimization problem will be as follows:

$$\begin{gathered}\min_{p_1,p_2} \frac{c_1(1-p_0)(1-p_2+\delta)}{p_0(p_1-\delta)+(1-p_0)(1-p_2+\delta)} + \frac{c_2p_0(1-p_1+\delta)}{p_0(1-p_1+\delta)+(1-p_0)(p_2-\delta)},\\ \text{such that } \mathcal{U}(p_1,\ p_2) \geq u_{min},\\ p_0p_1 < (1-p_2)(1-p_0) - \delta,\\ \text{and } p_2(1-p_0) < p_0(1-p_1) - \delta.\end{gathered}$$

Let us discuss the formulation of constraints for the *Magenta* sub-region. For this sub-region, $x_2^*(y)$ can be either $[1\ 0]$ or $[0\ 1]$, and $x_1^*(y)$ should be $[1\ 0]$. We know that Equations (21) and (22) correspond to both $x_1^*(y) = [1\ 0]$ and $x_2^*(y) = [1\ 0]$. However, as $x_2^*(y)$ can be $[0\ 1]$ for the *Magenta* sub-region, we need to make sure Equations (25) and (26) hold as well. Therefore, for the *Magenta* sub-region, Equations (21), (22), (25), and (26) need to hold simultaneously to ensure $x_2^*(y)$ can take the value of either of $[1\ 0]$ or $[0\ 1]$ when $x_1^*(y) = [1\ 0]$.

To derive the optimization problem, we need to identify that $x_2^*(y)$ has two possible values (i.e., $[0\ 1]$ and $[1\ 0]$) for the *Magenta* sub-region. The formulas for ACB, when $x_2^*(y) = [0\ 1]$ and $x_2^*(y) = [1\ 0]$, are given by Equations (15) and (18), respectively. Note that *ACB* will be maximized if the term inside the logarithm is maximized. Thus, at first, we need to maximize each of the terms $(\frac{c_1(1-p_0)(1-q_2)}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2})$ and $(\frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2(1-p_0)q_2}{p_0(1-q_1)+(1-p_0)q_2})$ with respect to q_1 and q_2. Then, we will consider the maximum between these two terms and compute p_1, p_2 that minimize this maximum value. Thus, the optimization problem for the *Magenta* sub-region would be the following:

$$\begin{gathered}\min_{p_1,p_2} \max_{q_1,q_2} \left(\frac{c_1(1-p_0)(1-q_2)}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1p_0q_1}{p_0q_1+(1-p_0)(1-q_2)} + \frac{c_2(1-p_0)q_2}{p_0(1-q_1)+(1-p_0)q_2}\right),\\ \text{such that } \mathcal{U}(p_1,\ p_2) \geq u_{min},\\ p_1p_0 < (1-p_2)(1-p_0),\\ p_2(1-p_0) < p_0(1-p_1),\\ p_0p_1 > (1-p_2)(1-p_0) - \delta,\\ \text{and } p_2(1-p_0) > p_0(1-p_1) - \delta.\end{gathered}$$

For the *Red* sub-region, we have $x_1^*(y) = [0\ 1]$ and $x_2^*(y)$ can either be $[0\ 1]$ or $[1\ 0]$. The conditions for $x_1^*(y) = [0\ 1]$ are given by Equations (28) and (29), respectively.

$$(1-p_2)(1-p_0) < p_1p_0 \tag{28}$$

$$p_0(1-p_1) < p_2(1-p_0) \tag{29}$$

Using the analysis shown before, we can conclude that Equations (23), (24), (28), and (29) need to hold together for the *Red* sub-region. As $x_2^*(y)$ has two possible values here as well, the cost function will be similar to that of the *Magenta* sub-region.

Finally, for the *Cyan* sub-region, $x_2^*(y)$ is always $[0\ 1]$. Therefore, both the inequalities $p_0p_1 > (1-p_2)(1-p_0) + \delta$ and $p_2(1-p_0) > p_0(1-p_1) + \delta$ need to hold together. Putting the value of q_1 and q_2, which maximizes *ACB* in Equation (15), we get Equation (30).

$$\text{ACB}(P_{Y|X}, Q_{Y|X}) = H_\infty(X) + \log_2\left(\frac{c_1p_0(p_1+\delta)}{p_0(p_1+\delta)+(1-p_0)(1-p_2-\delta)} + \frac{c_2(1-p_0)(p_2+\delta)}{p_0(1-p_1-\delta)+(1-p_0)(p_2+\delta)}\right) \tag{30}$$

Thus, for the *Cyan* sub-region, optimum p_1 and p_2 can be computed from the following optimization problem:

$$\begin{gathered}\min_{p_1,p_2} \frac{c_1 p_0(p_1+\delta)}{p_0(p_1+\delta)+(1-p_0)(1-p_2-\delta)} + \frac{c_2(1-p_0)(p_2+\delta)}{p_0(1-p_1-\delta)+(1-p_0)(p_2+\delta)},\\ \text{such that } \mathcal{U}(p_1,\ p_2) \geq u_{min},\\ p_0 p_1 > (1-p_2)(1-p_0) + \delta,\\ \text{and } p_2(1-p_0) > p_0(1-p_1) + \delta.\end{gathered}$$

The case of $p_0 \neq 0.5$. Here, we will briefly explain the formulation of the cost function of optimization problems when $p_0 \neq 0.5$. Figures 4(b) and 4(c) show the possible sub-regions when $p_0 > 0.5$ and $p_0 < 0.5$, respectively. The details of the constraints, for each sub-region, can be found in Table 12.

To illustrate, let us consider the *Magenta* sub-region where $x_2^*(y)$ can be both [1 0] and [0 0] when $p_0 > 0.5$, and [1 0] and [1 1] when $p_0 < 0.5$. If $x_2^*(y) = [0\ 0]$, then *ACB* is given by Equation (31).

$$\text{ACB}(P_{Y|X}, Q_{Y|X}) = H_\infty(X) + \log_2\left(\frac{c_1 q_1 p_0}{q_1 p_0 + (1-q_2)(1-p_0)} + \frac{c_2 p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)}\right) \tag{31}$$

Now, we know the term inside the logarithm, when $x_2^*(y) = [1\ 0]$, is $(\frac{c_1(1-q_2)(1-p_0)}{q_1 p_0+(1-q_2)(1-p_0)} + \frac{c_2 p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)})$. Thus, the optimization problem of the *Magenta* sub-region, when $p_0 > 0.5$ is given below:

$$\begin{gathered}\min_{p_1,p_2} \max_{q_1,q_2} \left(\left(\frac{c_1(1-q_2)(1-p_0)}{q_1 p_0+(1-q_2)(1-p_0)} + \frac{c_2 p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)}\right), \frac{c_1 q_1 p_0}{q_1 p_0+(1-q_2)(1-p_0)} + \frac{c_2 p_0(1-q_1)}{p_0(1-q_1)+q_2(1-p_0)}\right),\\ \text{such that } \mathcal{U}(p_1,\ p_2) \geq u_{min},\\ p_1 p_0 < (1-p_2)(1-p_0),\\ p_0 p_1 > (1-p_2)(1-p_0) - \delta,\\ \text{and } p_2(1-p_0) < p_0(1-p_1) - \delta.\end{gathered}$$

Similarly, Equation (32) represents *ACB* when $x_2^*(y) = [1\ 1]$.

$$\text{ACB}(P_{Y|X}, Q_{Y|X}) = H_\infty(X) + \log_2\left(\frac{c_1(1-q_2)(1-p_0)}{q_1 p_0+(1-q_2)(1-p_0)} + \frac{c_2 q_2(1-p_0)}{p_0(1-q_1)+q_2(1-p_0)}\right) \tag{32}$$

Thus, the optimization problem of the *Magenta* sub-region, when $p_0 < 0.5$, is shown below:

$$\begin{gathered}\min_{p_1,p_2} \max_{q_1,q_2} \left(\frac{c_1(1-p_0)(1-q_2)}{p_0 q_1+(1-p_0)(1-q_2)} + \frac{c_2 p_0(1-q_1)}{p_0(1-q_1)+(1-p_0)q_2}, \frac{c_1(1-q_2)(1-p_0)}{q_1 p_0+(1-q_2)(1-p_0)} + \frac{c_2 q_2(1-p_0)}{p_0(1-q_1)+q_2(1-p_0)}\right),\\ \text{such that } \mathcal{U}(p_1,\ p_2) \geq u_{min},\\ p_1 p_0 < (1-p_2)(1-p_0),\\ p_2(1-p_0) < p_0(1-p_1),\\ \text{and } p_2(1-p_0) > p_0(1-p_1) - \delta.\end{gathered}$$

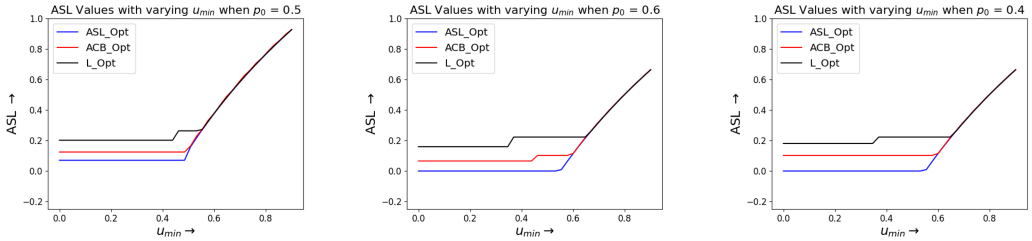
The cost function for the rest of the sub-regions are computed using the same technique.

6 SIMULATION RESULTS

Formerly, we have formulated various optimization problems and discussed how to solve *ACB* optimization problem. The details of *ASL* optimization problem can be found in Reference [47]. The codes, associated with these simulation results, can be found in Reference [50].

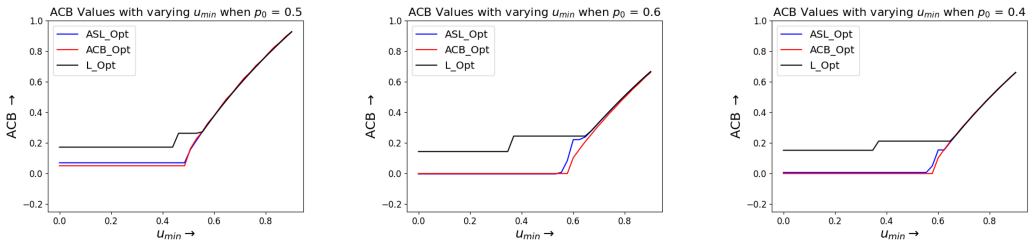
Once the data owner solves any of the optimization problems, she will have p_1 and p_2 that minimize the worst-case maximization of that metric. For example, the solution of the *ACB* optimization problem will result in p_1 and p_2 that will minimize the worst-case *ACB* whereas *ASL* optimization problem will result in the parameters that will minimize the worst-case *ASL*.

Depending on the system requirements, the data owner may wish to solve different optimization problems. For example, she may be interested in designing a system that minimizes the true



(a) $p_0 = 0.5$ (b) $p_0 > 0.5$ ($p_0 = 0.6$) (c) $p_0 < 0.5$ ($p_0 = 0.4$)

Fig. 5. Variation of worst-case ASL for different optimization problems when varying u_{min}, for different values of p_0 when $\delta = 0.05$ (generic utility provider).



(a) $p_0 = 0.5$ (b) $p_0 > 0.5$ ($p_0 = 0.6$) (c) $p_0 < 0.5$ ($p_0 = 0.4$)

Fig. 6. Variation of worst-case ACB for different optimization problems when varying u_{min}, for different values of p_0 when $\delta = 0.05$ (generic utility provider).

confidence boost of the attacker (i.e., *ACB*). Now, the data owner knows that minimizing the worst-case *ACB* does not necessarily indicate that worst-case *ASL* is also minimized. Nevertheless, she may be interested in finding out what is the worst-case *ASL* that may result from the designed system and how this value of *ASL* compares to the minimization of worst-case *ASL*. A comparative plot between these worst-case *ASL* values will enable her to have such information.

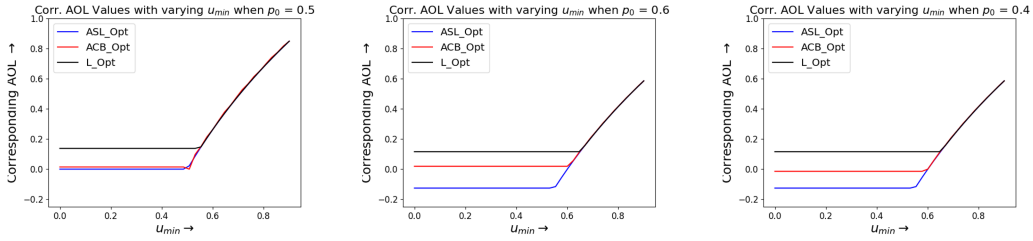
Thus, in this section, we shall compare the worst-case *ASL*, which results from *ACB* optimization, to the minimized worst-case *ASL* values for different u_{min}. Moreover, for completeness, we have also calculated p_1 and p_2, which optimize min-entropy leakage L, and compared the resultant worst-case *ASL* with the previously specified worst-case *ASL* values. We have obtained a similar plot for *ACB* as well. And finally, we have also plotted the *AOL* values corresponding to the p_1, p_2, q_1, q_2 values that constitute the solution of each of the optimization problems.

6.1 Generic Utility Provider

For a generic utility provider, the utility constraint is $\mathcal{U}(p_1, p_2) = p_1 p_0 + (1 - p_0) p_2 \geq u_{min}$. For our simulation, we have kept δ fixed to a small value ($\delta = 0.05$).

Figure 5 shows the variation of worst-case ASL, with varying u_{min}, for different optimization problems and different possible values of p_0. Similarly, Figures 6 and 7 show the variation of worst-case ACB and corresponding AOL, respectively, with varying u_{min}, for different optimization problems and different possible values of p_0.

We shall explain these graphs when $p_0 = 0.5$. Similar explanations hold when $p_0 \neq 0.5$. Moreover, for ease of the explanation, we shall consider three different values of u_{min} to reflect the three separate zones of utility. For indicating lower utility region, we have used $u_{min} = 0.2$, whereas



(a) $p_0 = 0.5$ (b) $p_0 > 0.5$ ($p_0 = 0.6$) (c) $p_0 < 0.5$ ($p_0 = 0.4$)

Fig. 7. Variation of corresponding AOL for different optimization problems when varying u_{min}, for different values of p_0 when $\delta = 0.05$ (generic utility provider).

Table 13. Optimized (p_1, p_2) Tuple for Various Optimization Problems ($p_0 = 0.5$)

Optimization	$u_{min} = 0.2$	$u_{min} = 0.5$	$u_{min} = 0.8$
ASL	$p_1 = 1$	$p_1 = 0.93$	$p_1 = 1$
	$p_2 = 0$	$p_2 = 0.025$	$p_2 = 0.6$
ACB	$p_1 = 0$	$p_1 = 0.05$	$p_1 = 0.6$
	$p_2 = 0.99$	$p_2 = 0.95$	$p_2 = 0.99$
L	$p_1 = 1$	$p_1 = 1$	$p_1 = 1$
	$p_2 = 0.1$	$p_2 = 0.1$	$p_2 = 0.6$

$u_{min} = 0.5$ and $u_{min} = 0.8$ indicate moderate and higher utility region, respectively. Table 13 shows the optimized (p_1, p_2) tuple for each value of u_{min} and each of the optimization problems.

Let us explain Figure 5(a) first. From Table 13, we see that $p_{1,ASL} = 1$ and $p_{2,ASL} = 0$ for $u_{min} = 0.2$. From Figure 4(a), we can see that these values of p_1 and p_2 correspond to the *Red* sub-region. For the *Red* sub-region, $x_2^*(y)$ can either be [1 0] or [0 1] (see Table 11). If $x_2^*(y) = [1\ 0]$, then we shall have $q_{1,ASL} = 0.95$ and $q_{2,ASL} = 0$, which results in ASL $= -\log_2(0.5) + \log_2(1 - (0.5 \times 0.95)) = \log_2(\frac{0.525}{0.5}) = 0.07$. Otherwise, when $x_2^*(y) = [0\ 1]$, we shall have $q_{1,ASL} = 1$ and $q_{2,ASL} = 0.05$, and ASL $= -\log_2(0.5) + \log_2(0.5 + 0.05 - (0.5 \times 0.05)) = \log_2(\frac{0.525}{0.5}) = 0.07$.

Now, we shall compare this value of ASL with the worst-case ASL, which will result from other optimization problems. When $u_{min} = 0.2$, we know that $p_{1,ACB} = 0$ and $p_{2,ACB} = 0.99$. This value of the (p_1, p_2) tuple indicates that the solution lies in the Red sub-region, more specifically, close to the Magenta-Red boundary. We already know that the optimization problem results in a solution that lies in the *Red* sub-region. Therefore, the worst-case (q_1, q_2) tuple will switch to the *Magenta* sub-region. As a result, we shall have $q_{1,ACB,worst} = 0$ and $q_{2,ACB,worst} = 0.94$. This value of (q_1, q_2) tuple will result in $x_2^*(y) = [1\ 0]$, and consequently $ASL = -\log_2(0.5) + \log_2(1 - 0.94 + (0.5 \times 0.94)) = \log_2(0.53/0.5) = 0.084$.

For min-entropy leakage (L), we need to identify that there are only *Green* and *Cyan* sub-regions. Recall that both the *Magenta* and the *Red* sub-regions are related to $x_2^*(y)$, and ergo, $Q_{X|Y}$. However, for min-entropy leakage (L), we do not have the notion of $Q_{X|Y}$. Thus, we shall only have two possible sub-regions, namely, *Green* and *Cyan*. When $u_{min} = 0.2$, the L optimization results in $p_{1,L} = 1$ and $p_{2,L} = 0.1$. This value of (p_1, p_2) tuple lies in the *Cyan* sub-region. And we already know that for the *Cyan* sub-region, we have $x_2^*(y) = [0\ 1]$, and thus, we shall have $q_{1,L,worst} = 1$ and $q_{2,L,worst} = 0.15$. Using this value of (q_1, q_2) tuple, we get ASL $= -\log_2(0.5) + \log_2(0.5 + 0.15 - (0.5 \times 0.15)) = \log_2(0.575/0.5) = 0.20$. Thus, as expected, *ASL* optimization results in the minimum value of the worst-case *ASL*. Similar observations hold when $u_{min} = 0.5$. When $u_{min} = 0.8$, each

Table 14. Worst-case ASL Values for Different Optimization Problems ($p_0 = 0.5$)

Optimization	$u_{min} = 0.2$	$u_{min} = 0.5$	$u_{min} = 0.8$
ASL	$p_1 = 1$	$p_1 = 0.93$	$p_1 = 1$
	$p_2 = 0$	$p_2 = 0.025$	$p_2 = 0.6$
	ASL = 0.07	ASL = 0.077	ASL = 0.72
ACB	$q_1 = 0$	$q_1 = 0$	$q_1 = 0.65$
	$q_2 = 0.94$	$q_2 = 0.9$	$q_2 = 1$
	ASL = 0.084	ASL = 0.1375	ASL = 0.72
L	$q_1 = 1$	$q_1 = 1$	$q_1 = 1$
	$q_2 = 0.15$	$q_2 = 0.15$	$q_2 = 0.65$
	ASL = 0.20	ASL = 0.20	ASL = 0.72

optimization problem results in such $(p_1,\ p_2)$ tuple that lies in the *Cyan* sub-region so the possible region of $(q_1,\ q_2)$ tuple only encompasses the *Cyan* sub-region. Therefore, for this value of u_{min}, we have the same worst-case *ASL* for each optimization problem. These results are summarized in Table 14.

Furthermore, Table 15 shows the ACB values for different values of u_{min} and different optimization problems. When $p_0 \neq 0.5$, solving *ACB* optimization results in zero *ACB* for lower u_{min}, implying that the disclosed information did not enhance the attacker's confidence at all. For higher utility region, both optimized $(p_1,\ p_2)$ and the possible region of $(q_1,\ q_2)$ lie in the *Cyan* sub-region (as explained before), and consequently, have the same worst-case *ACB*.

Finally, Table 16 shows the corresponding *AOL* values that result from the optimization problems. A lower value of *AOL* indicates that the disclosed information does not leak much information about the private variable. From Figure 7(a) and Table 16, we see that *L* optimization results in higher objective leakage, compared to *ASL* and *ACB* optimization, for low and moderate values of u_{min}. Moreover, note that when $p_0 \neq 0.5$, *ASL* optimization (and in some cases, *ACB* optimization) results in $(p_1,\ p_2)$ tuple for which the corresponding *AOL* is negative. Recall that it is possible for the average objective leakage to be negative when either $P_{X=0|Y} > P_{X=1|Y}$ or $P_{X=1|Y} > P_{X=0|Y}$ holds for both $Y = 0$ and $Y = 1$ (see property 2). When $p_0 > 0.5$, we see that $P_{X=0|Y=0} > P_{X=1|Y=0}$ and $P_{X=0|Y=1} > P_{X=1|Y=1}$ result in $x_1^*(y) = [0\ 0]$. Note from Table 11 that $x_1^*(y) = [0\ 0]$ is only possible for *Blue* and *Yellow* sub-regions when $p_0 > 0.5$. Our analysis shows that when the corresponding objective leakage gets negative, we have $x_1^*(y) = [0\ 0]$ and $x_2^*(y) = [0\ 1]$, and therefore, the solution lies in the *Yellow* sub-region. Similarly, when $p_0 < 0.5$, the solution corresponds to a $(p_1,\ p_2)$ tuple that results in $x_1^*(y) = [1\ 1]$, and the corresponding $(q_1,\ q_2)$ tuple results in $x_2^*(y) = [0\ 1]$. Therefore, the solution, again, lies in the *Yellow* sub-region. Accordingly, we can conclude that when the solution of the optimization problem lies in the *Yellow* sub-region, we get the negative average objective leakage, and this negative *AOL* indicates that it is possible to design a system that can hurt the attacker even in the worst-case scenario instead of helping.

6.2 Advanced Utility Provider

The advanced utility provider performs statistical analysis, based on his collection of (X, Y) pairs, and afterward, makes an educated guess. For such a user, the utility can be computed by considering the worst-case *average objective leakage* of the disclosed information and making sure that this worst-case *average objective leakage* is higher than some minimum utility. We have already denoted the approximated privacy mechanism of the utility provider as $Q'_{Y|X}$ (shown in Table 1), and $x_3^*(y) = \arg\max_{x \in \mathcal{X}} Q'_{X|Y}$. If based on the collection of (X, Y) pairs, the worst-utility is achieved at

Table 15. Worst-case ACB Values for Different Optimization Problems ($p_0 = 0.5$)

Optimization	$u_{min} = 0.2$	$u_{min} = 0.5$	$u_{min} = 0.8$
ASL	$q_1 = 1$	$q_1 = 0.88$	$q_1 = 1$
	$q_2 = 0.05$	$q_2 = 0$	$q_2 = 0.65$
	ACB = 0.07	ACB = 0.148	ACB = 0.72
ACB	$p_1 = 0$	$p_1 = 0.05$	$p_1 = 0.6$
	$p_2 = 0.99$	$p_2 = 0.95$	$p_2 = 0.99$
	ACB = 0.05	ACB = 0.1375	ACB = 0.72
L	$q_1 = 1$	$q_1 = 1$	$q_1 = 1$
	$q_2 = 0.15$	$q_2 = 0.15$	$q_2 = 0.65$
	ACB = 0.173	ACB = 0.173	ACB = 0.72

Table 16. Corresponding AOL Values for Different Optimization Problems ($p_0 = 0.5$)

Optimization	$u_{min} = 0.2$	$u_{min} = 0.5$	$u_{min} = 0.8$
ASL	$p_1 = 1$	$p_1 = 0.93$	$p_1 = 1$
	$p_2 = 0$	$p_2 = 0.025$	$p_2 = 0.6$
	AOL = 0	AOL = 0.063	AOL = 0.68
ACB	$p_1 = 0$	$p_1 = 0.05$	$p_1 = 0.6$
	$p_2 = 0.99$	$p_2 = 0.95$	$p_2 = 0.99$
	AOL = 0.014	AOL = 0	AOL = 0.67
L	$p_1 = 1$	$p_1 = 1$	$p_1 = 1$
	$p_2 = 0.1$	$p_2 = 0.1$	$p_2 = 0.6$
	AOL = 0.1375	AOL = 0.1375	AOL = 0.68

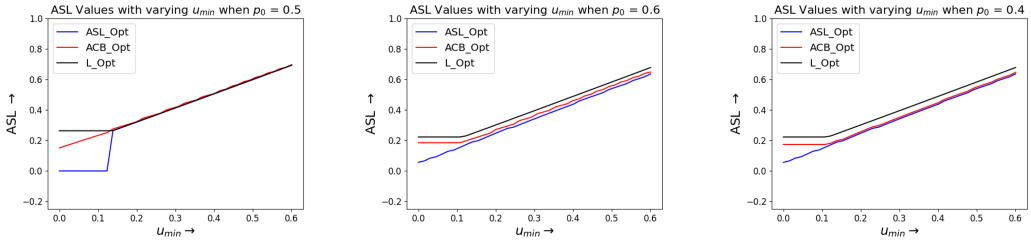
$x_3^*(y) = [0\ 1]$, then the utility constraint would be $\mathcal{U}(p_1, p_2) = H_\infty(X) + \log_2(p_0p_1 + p_2 - p_0p_2) \geq u_{min}$. Otherwise when worst-utility occurs at $x_3^*(y) = [1\ 0]$, the utility constraint becomes $\mathcal{U}(p_1, p_2) = H_\infty(X) + \log_2(1 - p_2 + p_0p_2 - p_0p_1) \geq u_{min}$.

Now, we shall explain how to compute the index $x_3^*(y)$ that corresponds to the index of worst utility. For explanation, let us consider the *Magenta* sub-region when $p_0 = 0.5$. Therefore, we have $x_1^*(y) = [1\ 0]$, and $x_3^*(y)$ can be either $[1\ 0]$ or $[0\ 1]$ (as $d_{TV}(P_{Y|X}, Q'_{Y|X}) \leq \delta$). If a specific choice of (q'_1, q'_2) results in $x_3^*(y) = [1\ 0]$, then we shall have AOL $= H_\infty(X) + \log_2(1 - p_2 + p_0p_2 - p_0p_1)$. Consequently, we have the following optimization problem:

$$\begin{gathered}\min_{p_1,p_2} 1 - p_2 + p_0p_2 - p_0p_1,\\ \text{such that } p_1p_0 < (1-p_2)(1-p_0),\\ p_2(1-p_0) < p_0(1-p_1),\\ p_0p_1 > (1-p_2)(1-p_0) - \delta,\\ \text{and } p_2(1-p_0) > p_0(1-p_1) - \delta.\end{gathered}$$

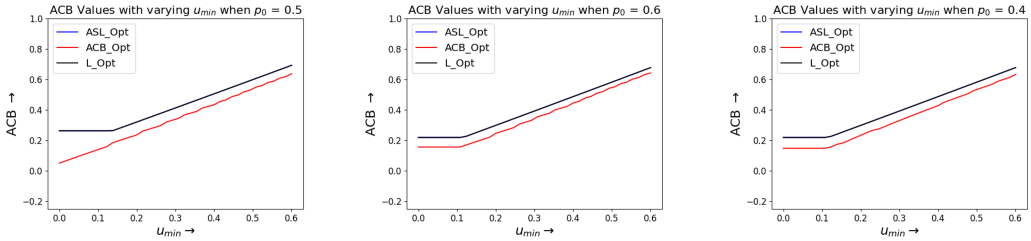
Otherwise, if the choice of (q'_1, q'_2) corresponds to $x_3^*(y) = [0\ 1]$, then we get AOL $= H_\infty(X) + \log_2(p_0p_1 + p_2 - p_0p_2)$. Accordingly, we get the following optimization problem:

$$\begin{gathered}\min_{p_1,p_2} p_0p_1 + p_2 - p_0p_2,\\ \text{such that } p_1p_0 < (1-p_2)(1-p_0),\\ p_2(1-p_0) < p_0(1-p_1),\\ p_0p_1 > (1-p_2)(1-p_0) - \delta,\\ \text{and } p_2(1-p_0) > p_0(1-p_1) - \delta.\end{gathered}$$



(a) $p_0 = 0.5$ (b) $p_0 > 0.5$ ($p_0 = 0.6$) (c) $p_0 < 0.5$ ($p_0 = 0.4$)

Fig. 8. Variation of worst-case ASL for different optimization problems when varying u_{min}, for different values of p_0 (advanced utility provider).



(a) $p_0 = 0.5$ (b) $p_0 > 0.5$ ($p_0 = 0.6$) (c) $p_0 < 0.5$ ($p_0 = 0.4$)

Fig. 9. Variation of worst-case ACB for different optimization problems when varying u_{min}, for different values of p_0 (advanced utility provider).

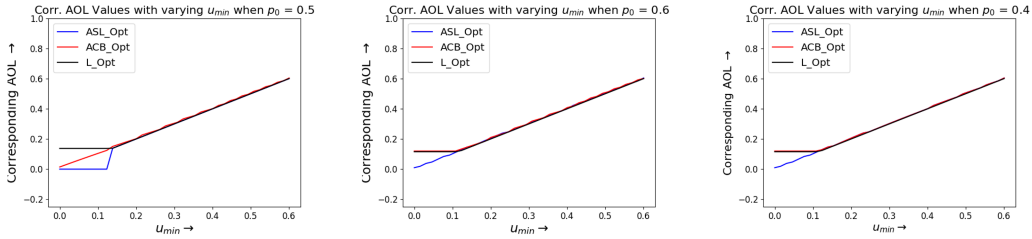
Eventually, we shall choose that value of $x_3^*(y)$ for which we get the lowest *AOL*. As it turns out for the *Magenta* sub-region, the worst-case AOL is achieved for $x_3^*(y) = [0\ 1]$ when $p_0 = 0.5$. Thus, we shall have the utility constraint as $\mathcal{U}(p_1, p_2) = H_\infty(X) + \log_2(p_0 p_1 + p_2 - p_0 p_2) \geq u_{min}$.

Similar to the generic utility provider, we have varied u_{min} and plotted the variation of each of the proposed metrics for the advanced utility provider as well. Figure 8 shows the variation of worst-case *ASL*, with varying u_{min}, for different optimization problems and different possible values of p_0. It is interesting to note that when u_{min} gets slightly higher (> 0.1) and $p_0 = 0.5$, both *ASL* and *ACB* optimization problems result in the same value of worst-case *ASL*. The reason being the optimized (p_1, p_2) tuple for *ASL* optimization problem lies in the *Cyan* sub-region. Even though the optimized (p_1, p_2) tuple for *ACB* optimization problem lies in the *Red* sub-region, the worst-case (q_1, q_2) tuple falls in the *Cyan* sub-region. As a result, the resultant worst-case values are the same. Finally, Figures 9 and 10 show the variation of worst-case *ACB* and corresponding *AOL*, respectively, with varying u_{min}, for different optimization problems and different possible values of p_0.

7 OPTIMAL PRIVACY MECHANISM FOR A GENERAL FRAMEWORK

In this section, we will provide a concise overview of the optimization problem for the general case and propose an approach to get the solution. For the general extension, the optimization can be represented as follows:

$$\begin{aligned} &\min_{P_{Y|X}} \max_{Q_{Y|X}} \mathcal{L}(X, Y), \\ &\text{such that } \mathcal{U}(X, Y) \geq u_{min}, \\ &\mathrm{d}_{TV}(P_{Y|X}, Q_{Y|X}) \leq \delta. \end{aligned} \tag{33}$$



(a) $p_0 = 0.5$ (b) $p_0 > 0.5$ ($p_0 = 0.6$) (c) $p_0 < 0.5$ ($p_0 = 0.4$)

Fig. 10. Variation of corresponding AOL for different optimization problems when varying u_{min}, for different values of p_0 (advanced utility provider).

Here, $\mathcal{L}(X, Y)$ can indicate any of the proposed metrics of interest. Let us consider $\mathcal{L}(X, Y) = ACB(P_{Y|X}, Q_{Y|X})$. Since in the general case the optimization problem may not be convex, we cannot find an exact solution as we did for the binary privacy mechanism. However, we can still obtain an approximate solution by employing a greedy approach, the details of which are given below:

— The algorithm *iteratively* finds the optimum $P_{Y|X}$ while a specific threshold condition is maintained. The details are shown in Algorithm 1.
We initialize our step size, μ, to a random positive value (Line 3). Next, we utilize the function *OPT_P* to find the optimum $P_{Y|X}$ at distance μ from the initial $P_{Y|X}$ (Line 9), and accordingly, we update our privacy mechanism to the new $P_{Y|X}$. At the same time, we keep track of the optimized worst-case leakage value. Afterward, we reduce the value of μ by half ($\mu = \frac{\mu}{2}$) (Line 11) and check if the reduced value of μ has further optimized the worst-case leakage. Such a check is done by computing the difference between the worst-case leakage values that we achieved for both μ and $\frac{\mu}{2}$. We keep repeating the process while the difference between these two leakages is higher than 0 (Lines 6–12).
— Now, we shall describe how *OPT_P* results in the optimum $P_{Y|X}$ for a fixed μ. The details are shown in Algorithm 2.
We take $P_{Y|X}$ as input and generate a list of $\hat{P}_{Y|X}$ that are μ away from $P_{Y|X}$ (Line 8). Then, we use the function *OPT_Q* to find the optimum $\hat{P}_{Y|X}$ for the next iteration (Line 9). We update our $P_{Y|X}$ to this value of $\hat{P}_{Y|X}$ and keep repeating the process while the difference between the previous leakage value and the current leakage value is higher than 0 (Lines 5–10).
— Finally, we shall discuss how *OPT_Q* finds the $\hat{P}_{Y|X}$ for the next iteration. The details of this step are shown in Algorithm 3.
Recall that we need to consider all $Q_{Y|X}$ that are at least δ away from $P_{Y|X}$. Therefore, for each $\hat{P}_{Y|X}$ in $list\_\hat{P}_{Y|X}$, we generate a list of $Q_{Y|X}$ that maintains our distance constraints (Line 3). Afterward, we compute the leakage value only for those $Q_{Y|X}$ that maintain our utility constraint and choose that $Q_{Y|X}$ that maximizes the leakage value (Lines 5–13). Once we have all the worst-case leakage values for each $\hat{P}_{Y|X}$ in $list\_\hat{P}_{Y|X}$, we choose the one that minimizes such maximization of the leakage value as our optimum $\hat{P}_{Y|X}$ (Lines 15–17).

8 RELATED WORK

The notion of information leakage encompasses a broad range of literature, from information theory to computer security. Shannon entropy and mutual information-based definitions have previously been employed to represent information leakage in References [27, 36, 44, 49]. Information-theoretic approaches to define privacy, based on such definitions of information

ALGORITHM 1: Algorithm for solving the optimization problem

Input: δ, u_{min}

Output: Optimum leakage value, optimum privacy mechanism

Initialize $P_{Y|X}$ to a transition probability matrix
$new\_P \leftarrow P_{Y|X}$
$\mu \leftarrow$ *a positive value*
$new\_leak \leftarrow$ *a large positive value*
$leak\_diff \leftarrow$ *a positive value*
while $leak\_diff > 0$ **do**
    $current\_leak \leftarrow new\_leak$
    $P_{Y|X} \leftarrow new\_P$
    $(new\_leak, new\_P) \leftarrow OPT\_P(\mu, \delta, u_{min}, P_{Y|X})$
    $leak\_diff \leftarrow current\_leak - new\_leak$
    $\mu \leftarrow \mu/2$
end while
return $current\_leak$, $P_{Y|X}$

ALGORITHM 2: Algorithm for function OPT_P

Input: μ, δ, u_{min}, $P_{Y|X}$

Output: Optimum leakage value, optimum $P_{Y|X}$ **(for a specific** μ**)**

function $(\mu, \delta, u_{min}, P_{Y|X})$
    $new\_leak \leftarrow$ *a large positive value*
    $leak\_diff \leftarrow$ *a positive value*
    $new\_P \leftarrow P_{Y|X}$
    while $leak\_diff > 0$ **do**
        $current\_leak \leftarrow new\_leak$
        $P_{Y|X} \leftarrow new\_P$
        Generate $list\_\hat{P}_{Y|X}$ *that are* μ *away from* $P_{Y|X}$
        $(new\_leak, new\_P) \leftarrow OPT\_Q\,(list\_\hat{P}_{Y|X}, \delta, u_{min})$
        $leak\_diff \leftarrow current\_leak - new\_leak$
    end while
    return $current\_leak$, $P_{Y|X}$
end function

leakage, have been explored in References [14, 57]. An information leakage games-based framework [4] has also been proposed to reduce the leakage. Issa et al. [30] provided an operational definition of leakage where an adversary tries to guess a randomized function of X, upon observing Y, either in a single or k guesses. The authors further introduced the concept of *maximal realizable leakage* in Reference [31] that represents the gain of an adversary in the worst-case scenario. The authors in References [3, 8, 51] utilized another measure of leakage, known as *min-entropy leakage*. This definition adopts *Rényi entropy* [45] and considers the difference between the initial uncertainty of guessing X (i.e., $H(X)$) and the remaining uncertainty of guessing, after some observations (i.e., $H(X|Y)$) and defines the leakage as the difference between these two measures (i.e., $H(X) - H(X|Y)$). A tunable measure for information leakage was introduced in Reference [37] to capture the various actions of an adversary based on her belief.

ALGORITHM 3: Algorithm for function OPT_Q

Input: $list\_\hat{P}_{Y|X}, \delta, u_{min}$
Output: Optimum leakage value, optimum $\hat{P}_{Y|X}$

function $(list\_\hat{P}_{Y|X}, \delta, u_{min})$
  for *each* $\hat{P}_{Y|X}$ *in the* $list\_\hat{P}_{Y|X}$ **do**
    Generate $list\_Q_{Y|X}$ *that are within* δ *distance of* $\hat{P}_{Y|X}$
    $leak\_list\_P \leftarrow [\,]$
    for *each* $Q_{Y|X}$ *in* $list\_Q_{Y|X}$ **do**
      $leak\_list\_Q \leftarrow [\,]$
      if *Utility constraint is maintained* **then**
        $leak\_Q \leftarrow \mathcal{L}(P_{Y|X}, Q_{Y|X})$
        Append $leak\_Q$ *to* $leak\_list\_Q$
      end if
    end for
    $leak\_max \leftarrow \max(leak\_list\_Q)$
    Append $leak\_max$ *to* $leak\_list\_P$
  end for
  $leak\_min \leftarrow \min(leak\_list\_P)$
  $min\_index \leftarrow leak\_list\_P.index(leak\_min)$
  $min\_P \leftarrow list\_\hat{P}_{Y|X}[min\_index]$
  return $leak\_min, min\_P$
end function

The authors of Reference [41] introduced *g-leakage*, a generalization of *min-entropy leakage*. The definition of *g-leakage* utilizes a gain function, g, to represent how much benefit the adversary has achieved by guessing the password either partially or completely. However, this definition of g-leakage utilizes the prior knowledge of the channel, which is equivalent to the privacy mechanism in our setup. Afterward, the authors provided several axioms for information leakage in Reference [5]. Cherubin et al. [12] estimated information leakage via machine learning in a black-box setup. In Reference [46], the authors estimated the g-leakage also via machine learning approaches and evaluated the performance of their approach through various experiments using k-nearest neighbors and neural network.

Another line of research for quantifying the notion of leakage is *differential privacy* [18, 19]. The authors in Reference [6] derived the bound of min-entropy leakage of an ϵ differentially private mechanism and showed the absence of domain-sized independent bounds for such a mechanism. Moreover, optimizing the tradeoff between utility and privacy, for any differential private mechanism, has been studied extensively [9, 25, 28, 32, 35, 60]. In Reference [55], the authors analyzed *f-information* as a measure of privacy for a database containing public and private entries, while the χ^2*-information* was adopted as a measure of privacy and utility in Reference [56] for a similar database. Note that both *f-information* and χ^2*-information* assume the correct joint distribution is known. The authors identified that the privacy metric would be different under exact and approximated distributions and provided the bound on the difference between the exact and approximated privacy measures.

Searchable Encryption is a cryptographic technique that allows users to search over encrypted data without revealing the significant contents of the data to the server or any other third party [52]. This cryptographic scheme ensures the leaking of only harmless information that is termed as *leakage profile* [15]. Significant developments were made in this direction of research, such as

the introduction of dynamic searchable encryption [10] and dynamic local searchable symmetric encryption [39]. Numerous leakage-abuse attacks have also been explored [7, 21, 33], where the adversary is interested in reconstructing the underlying plaintext database once they have access to the query leakages. Consequently, several works study the quantification of privacy in searchable encryption [34] and the security of the data from a system-wide viewpoint [26].

While the concept of searchable encryption may initially appear relevant, there exists a significant distinction between it and the setup proposed in our article. Encryption, in general, involves the conversion of plaintext into ciphertext using a key that can only be deciphered by those with knowledge of the corresponding decryption key. Its primary purpose is to prevent unauthorized access to data, with only authorized parties possessing information about the sensitive data. This setup differs from what is presented in our article. In our work, we employ a privacy mechanism to privatize the data before its disclosure. Subsequently, the disclosed information becomes accessible to an adversary who attempts to infer the private data, similar to the actions of an adversary in the encryption process. However, the crucial difference lies in the role of the utility provider. In encryption, the utility provider possesses the corresponding decryption key and uses it to access the private information. In our setup, the utility provider does not hold any information that would enable them to "*decrypt*" the disclosed information. Instead, the design of the privacy mechanism ensures that the utility provider achieves their desired utility while performing analysis using the disclosed information. This paradigm comes in handy, especially in the overwhelming number of situations in which the attacker and the utility provider are one and the same physical entity.

The authors of References [22, 23] considered the *Fisher information* as a measure of information leakage and derived the optimal privacy-preserving policy. In Reference [17], the authors utilized the minimax techniques for local privacy and derived bounds on both mutual information quantities and KL-divergence under such a privacy model. Acharya et al. [1] performed identity and closeness testing of discrete distributions in the differential privacy settings. The authors in Reference [53] derived an estimator of an unknown discrete distribution, based on entropy, number of distinct elements, and distance metrics, and later modified the method in Reference [54] to obtain the optimal estimator. Clarkson et al. [13] discussed how an attacker's beliefs change by observing the execution of a program. The authors introduced a metric to capture an adversary who observes the execution of a program and updates her belief accordingly. Authors in Reference [29] unified the notion of belief and leakage for an adversary. They introduced metrics to represent the belief of the attacker when they had different (and potentially wrong) initial beliefs regarding the distribution of the secret and presented several properties to measure the accuracy and belief of the adversary. Finally, the leakage metrics for incomplete statistics were introduced in Reference [47], and later the authors extended these metrics to a more general Bayesian framework in Reference [48].

9 CONCLUSION AND FUTURE WORK

Usually an adversary lacks complete knowledge of the privacy mechanism and tries to approximate the privacy mechanism by analyzing several input-output pairs. This article identified the lack of traditional information leakage measures to adequately capture the leakage in such scenarios and introduced diverse novel leakage metrics: *average subjective leakage*, *average confidence boost*, and *average objective leakage*. We formalized the definitions of these metrics and developed a minimax optimization problem, the solution of which results in the optimum binary privacy mechanism. Additionally, we formulated the optimization problems analytically and numerically computed the probabilities that achieve the optimium privacy mechanism. Interestingly, we found out that the *AOL*, computed using the optimal privacy mechanism, can be negative, meaning that the disclosed information would hurt the attacker rather than help her. We have observed that

optimizing for *min-entropy leakage (L)* leads to increased values of ASL, ACB, and the corresponding AOL, in comparison to optimizing for either ASL or ACB alone. A higher value of corresponding AOL indicates a higher amount of information leakage in the system. Furthermore, higher values of ASL and ACB suggest that the adversary will have greater confidence in their ability to infer information, potentially leading to severe consequences if they act based on this confidence. Therefore, it is essential to minimize the worst-case values of the leakage metrics, which can arise due to incomplete statistics, when designing privacy mechanisms. This ensures that even in the worst-case scenarios, the actual information leakage of the system is minimized.

One important direction for future research is to extend the proposed measures to non-Bayesian frameworks, allowing for their applicability to large datasets. By doing so, we can ensure that the metrics become effective in a broader range of scenarios. Additionally, an interesting direction to explore would be to investigate how to design an optimal privacy mechanism specifically tailored to these metrics when applied in a non-Bayesian framework. This research can provide valuable insights into the development of privacy-preserving techniques that are effective and efficient in handling large-scale datasets outside of the Bayesian context.

ACKNOWLEDGMENTS

The findings achieved herein are solely the responsibility of the authors. We appreciate the anonymous reviewers for their valuable suggestions and comments.

REFERENCES

[1] Jayadev Acharya, Ziteng Sun, and Huanyu Zhang. 2018. Differentially private testing of identity and closeness of discrete distributions. In *Conference on Advances in Neural Information Processing Systems*. 6878–6891.

[2] Shaukat Ali, Naveed Islam, Azhar Rauf, Ikram Ud Din, Mohsen Guizani, and Joel J. P. C. Rodrigues. 2018. Privacy and security issues in online social networks. *Fut. Internet* 10, 12 (2018), 114.

[3] Mário S. Alvim, Miguel E. Andrés, Konstantinos Chatzikokolakis, Pierpaolo Degano, and Catuscia Palamidessi. 2011. Differential privacy: On the trade-off between utility and information leakage. In *International Workshop on Formal Aspects in Security and Trust*. Springer, 39–54.

[4] Mário S. Alvim, Konstantinos Chatzikokolakis, Yusuke Kawamoto, and Catuscia Palamidessi. 2018. A game-theoretic approach to information-flow control via protocol composition. *Entropy* 20, 5 (2018), 382.

[5] Mário S. Alvim, Konstantinos Chatzikokolakis, Annabelle McIver, Carroll Morgan, Catuscia Palamidessi, and Geoffrey Smith. 2016. Axioms for information leakage. In *EEE 29th Computer Security Foundations Symposium (CSF'16)*. IEEE, 77–92.

[6] Gilles Barthe and Boris Kopf. 2011. Information-theoretic bounds for differentially private mechanisms. In *IEEE 24th Computer Security Foundations Symposium*. IEEE, 191–204.

[7] Laura Blackstone, Seny Kamara, and Tarik Moataz. 2019. Revisiting leakage abuse attacks. *Cryptology ePrint Archive* (2019).

[8] Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2009. Quantitative notions of leakage for one-try attacks. *Electronic Notes in Theoretical Computer Science* 249 (2009), 75–91.

[9] Hai Brenner and Kobbi Nissim. 2010. Impossibility of differentially private universally optimal mechanisms. In *IEEE 51st Annual Symposium on Foundations of Computer Science*. IEEE, 71–80.

[10] David Cash, Joseph Jaeger, Stanislaw Jarecki, Charanjit Jutla, Hugo Krawczyk, Marcel-Cătălin Roşu, and Michael Steiner. 2014. Dynamic searchable encryption in very-large databases: Data structures and implementation. *Cryptology ePrint Archive* (2014).

[11] Konstantinos Chatzikokolakis, Tom Chothia, and Apratim Guha. 2010. Statistical measurement of information leakage. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 390–404.

[12] Giovanni Cherubin, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. 2019. F-BLEAU: Fast black-box leakage estimation. In *IEEE Symposium on Security and Privacy (SP'19)*. IEEE, 835–852.

[13] Michael R. Clarkson, Andrew C. Myers, and Fred B. Schneider. 2009. Quantifying information flow with beliefs. *J. Comput. Secur.* 17, 5 (2009), 655–701.

[14] Paul Cuff and Lanqing Yu. 2016. Differential privacy as a mutual information constraint. In *ACM SIGSAC Conference on Computer and Communications Security*. 43–54.

[15] Reza Curtmola, Juan Garay, Seny Kamara, and Rafail Ostrovsky. 2006. Searchable symmetric encryption: Improved definitions and efficient constructions. In *13th ACM Conference on Computer and Communications Security*. 79–88.
[16] Damien Desfontaines and Balázs Pejó. 2019. SoK: Differential privacies. *arXiv preprint arXiv:1906.01337* (2019).
[17] John C. Duchi, Michael I. Jordan, and Martin J. Wainwright. 2013. Local privacy, data processing inequalities, and statistical minimax rates. *arXiv preprint arXiv:1302.3203* (2013).
[18] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference*. Springer, 265–284.
[19] Cynthia Dwork, and Aaron Roth. 2014. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science* 9, 3–4 (2014), 211–407.
[20] Barbara Espinoza and Geoffrey Smith. 2013. Min-entropy as a resource. *Inf. Comput.* 226 (2013), 57–75.
[21] Francesca Falzon, Evangelia Anna Markatou, Akshima, David Cash, Adam Rivkin, Jesse Stern, and Roberto Tamassia. 2020. Full database reconstruction in two dimensions. In *ACM SIGSAC Conference on Computer and Communications Security*. 443–460.
[22] Farhad Farokhi and Henrik Sandberg. 2019. Ensuring privacy with constrained additive noise by minimizing Fisher information. *Automatica* 99 (2019), 275–288.
[23] Farhad Farokhi and Henrik Sandberg. 2020. Fisher information privacy with application to smart meter privacy using HVAC units. In *Privacy in Dynamical Systems*. Springer, 3–17.
[24] Terri D. Fisher. 2013. Gender roles and pressure to be truthful: The bogus pipeline modifies gender differences in sexual but not non-sexual behavior. *Sex Roles* 68, 7 (2013), 401–414.
[25] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. 2012. Universally utility-maximizing privacy mechanisms. *SIAM J. Comput.* 41, 6 (2012), 1673–1693.
[26] Zichen Gui, Kenneth G. Paterson, and Sikhar Patranabis. 2023. Rethinking searchable symmetric encryption. In *2023 IEEE Symposium on Security and Privacy (SP)*. IEEE, 1401–1418.
[27] Deniz Gunduz, Elza Erkip, and H. Vincent Poor. 2008. Lossless compression with security constraints. In *IEEE International Symposium on Information Theory*. IEEE, 111–115.
[28] Mangesh Gupte and Mukund Sundararajan. 2010. Universally optimal privacy mechanisms for minimax agents. In *29th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*. 135–146.
[29] Sardaouna Hamadou, Catuscia Palamidessi, and Vladimiro Sassone. 2017. Quantifying leakage in the presence of unreliable sources of information. *J. Comput. Syst. Sci.* 88 (2017), 27–52.
[30] Ibrahim Issa, Sudeep Kamath, and Aaron B. Wagner. 2016. An operational measure of information leakage. In *Annual Conference on Information Science and Systems (CISS'16)*. IEEE, 234–239.
[31] Ibrahim Issa and Aaron B. Wagner. 2017. Operational definitions for some common information leakage metrics. In *IEEE International Symposium on Information Theory (ISIT'17)*. IEEE, 769–773.
[32] Kousha Kalantari, Lalitha Sankar, and Anand D. Sarwate. 2018. Robust privacy-utility tradeoffs under differential privacy and Hamming distortion. *IEEE Trans. Inf. Forens. Secur.* 13, 11 (2018), 2816–2830.
[33] Seny Kamara, Abdelkarim Kati, Tarik Moataz, Thomas Schneider, Amos Treiber, and Michael Yonli. 2022. SoK: Cryptanalysis of encrypted search with LEAKER–A framework for LEakage AttacK evaluation on real-world data. In *IEEE 7th European Symposium on Security and Privacy (EuroS&P'22)*. IEEE, 90–108.
[34] Evgenios M. Kornaropoulos, Nathaniel Moyer, Charalampos Papamanthou, and Alexandros Psomas. 2022. Leakage inversion: Towards quantifying privacy in searchable encryption. In *ACM SIGSAC Conference on Computer and Communications Security*. 1829–1842.
[35] Chao Li, Michael Hay, Vibhor Rastogi, Gerome Miklau, and Andrew McGregor. 2010. Optimizing linear counting queries under differential privacy. In *29th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*. 123–134.
[36] Simon Li, Ashish Khisti, and Aditya Mahajan. 2018. Information-theoretic privacy for smart metering systems with a rechargeable battery. *IEEE Trans. Inf. Theor* 64, 5 (2018), 3679–3695.
[37] Jiachun Liao, Oliver Kosut, Lalitha Sankar, and Flavio P. Calmon. 2018. A tunable measure for information leakage. In *IEEE International Symposium on Information Theory (ISIT'18)*. IEEE, 701–705.
[38] Linyuan Lü, Matúš Medo, Chi Ho Yeung, Yi-Cheng Zhang, Zi-Ke Zhang, and Tao Zhou. 2012. Recommender systems. *Phys. Rep.* 519, 1 (2012), 1–49.
[39] Brice Minaud and Michael Reichle. 2022. Dynamic local searchable symmetric encryption. In *42nd Annual International Cryptology Conference: Advances in Cryptology (CRYPTO'22)*. Springer, 91–120.
[40] Maryam Mohsin. 2021. 10 Facebook Statistics Every Marketer Should Know in 2021 [Infographic]. Retrieved from https://www.oberlo.com/blog/facebook-statistics
[41] S. Alvim M'rio, Kostas Chatzikokolakis, Catuscia Palamidessi, and Geoffrey Smith. 2012. Measuring information leakage using generalized gain functions. In *IEEE 25th Computer Security Foundations Symposium*. IEEE, 265–279.

[42] Jingmei Qiu. 2012. *Iterative Solutions of Linear Systems*. Retrieved from https://www.math.uh.edu/~jingqiu/math4364/iterative_linear_system.pdf
[43] Borzoo Rassouli and Deniz Gündüz. 2019. Optimal utility-privacy trade-off with total variation distance as a privacy measure. *IEEE Trans. Inf. Forens. Secur.* 15 (2019), 594–603.
[44] David Rebollo-Monedero, Jordi Forne, and Josep Domingo-Ferrer. 2009. From t-closeness-like privacy to postrandomization via information theory. *IEEE Trans. Knowl. Data Eng.* 22, 11 (2009), 1623–1636.
[45] Alfréd Rényi et al. 1961. On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California.
[46] Marco Romanelli, Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Pablo Piantanida. 2020. Estimating g-leakage via machine learning. In *ACM SIGSAC Conference on Computer and Communications Security*. 697–716.
[47] Shahnewaz Karim Sakib, George T. Amariucai, and Yong Guan. 2021. Information leakage metrics for adversaries with incomplete information: Binary privacy mechanism. In *IEEE International Conference on Communications (ICC'21)*. IEEE, 1–7.
[48] Shahnewaz Karim Sakib, George T. Amariucai, and Yong Guan. 2022. Variations and extensions of information leakage metrics with applications to privacy problems with imperfect statistical information. In *IEEE 36th Computer Security Foundations Symposium (CSF'23)*. IEEE Computer Society, 95–110.
[49] Lalitha Sankar, S. Raj Rajagopalan, and H. Vincent Poor. 2013. Utility-privacy tradeoffs in databases: An information-theoretic approach. *IEEE Trans. Inf. Forens. Secur.* 8, 6 (2013), 838–852.
[50] Shahnewaz. 2023. ACM_TOPS_Codes. Retrieved from https://github.com/Shahanewaz/ACM_TOPS_Codes.git
[51] Geoffrey Smith. 2009. On the foundations of quantitative information flow. In *International Conference on Foundations of Software Science and Computational Structures*. Springer, 288–302.
[52] Dawn Xiaoding Song, David Wagner, and Adrian Perrig. 2000. Practical techniques for searches on encrypted data. In *IEEE Symposium on Security and Privacy (S&P'00)*. IEEE, 44–55.
[53] Gregory Valiant and Paul Valiant. 2011. Estimating the unseen: An n/log (n)-sample estimator for entropy and support size, shown optimal via new CLTs. In *43rd Annual ACM Symposium on Theory of Computing*. 685–694.
[54] Paul Valiant and Gregory Valiant. 2013. Estimating the unseen: Improved estimators for entropy and other properties. In *Conference on Advances in Neural Information Processing Systems*. 2157–2165.
[55] Hao Wang, Mario Diaz, Flavio P. Calmon, and Lalitha Sankar. 2018. The utility cost of robust privacy guarantees. In *IEEE International Symposium on Information Theory (ISIT'18)*. IEEE, 706–710.
[56] Hao Wang, Lisa Vo, Flavio P. Calmon, Muriel Médard, Ken R. Duffy, and Mayank Varia. 2019. Privacy with estimation guarantees. *IEEE Trans. Inf. Theor.* 65, 12 (2019), 8025–8042.
[57] Weina Wang, Lei Ying, and Junshan Zhang. 2016. On the relation between identifiability, differential privacy, and mutual-information privacy. *IEEE Trans. Inf. Theor.* 62, 9 (2016), 5018–5029.
[58] Anita Whiting and David Williams. 2013. Why people use social media: a uses and gratifications approach. *Qualitative Market Research: an International Journal* 16, 4 (2013), 362–369.
[59] Robert L. Winkler and Leroy A. Franklin. 1979. Warner's randomized response model: A Bayesian approach. *J. Am. Stat. Assoc.* 74, 365 (1979), 207–214.
[60] Tianrui Xiao and Ashish Khisti. 2019. Maximal information leakage based privacy preserving data disclosure mechanisms. In *16th Canadian Workshop on Information Theory (CWIT'19)*. IEEE, 1–6.

Received 31 December 2022; revised 15 June 2023; accepted 12 September 2023