Almost Perfect Privacy for Additive Gaussian Privacy Filters

Shahab Asoodeh*, Fady Alajaji, and Tamás Linder

Department of Mathematics and Statistics, Queen's University
Jeffery Hall, 48 University Ave., Kingston, ON, Canada
{asoodehshahab,fady,linder}@mast.queensu.ca

Abstract. We study the maximal mutual information about a random variable Y (representing non-private information) displayed through an additive Gaussian channel when guaranteeing that only ε bits of information is leaked about a random variable X (representing private information) that is correlated with Y. Denoting this quantity by $g_\varepsilon(X,Y)$, we show that for perfect privacy, i.e., $\varepsilon = 0$, one has $g_0(X,Y) = 0$ for any pair of absolutely continuous random variables (X,Y) and then derive a second-order approximation for $g_\varepsilon(X,Y)$ for small ε. This approximation is shown to be related to the strong data processing inequality for mutual information under suitable conditions on the joint distribution P_{XY}. Next, motivated by an operational interpretation of data privacy, we formulate the privacy-utility tradeoff in the same setup using estimation-theoretic quantities and obtain explicit bounds for this tradeoff when ε is sufficiently small using the approximation formula derived for $g_\varepsilon(X,Y)$.

Keywords: Data privacy, rate-privacy function, estimation noise-to-signal ratio, MMSE, additive Gaussian channel, mutual information, maximal correlation.

1 Introduction

The ever increasing growth of social networks has brought major challenges in terms of data privacy. This paper focuses on a privacy problem which is relevant for users or designers of social networks: the trade-off between data privacy and customized services performance. On the one hand, users want their private data to remain secret, and on the other hand, they also desire to benefit from customized services that require personal information in order to function properly. In this context, it is reasonable to assume that the user has two kinds of data: private data such as passport numbers, credit cards numbers, etc; and non-private data such as gender, age, etc. In general, private and non-private data are correlated. Thus, it is possible that enough non-private data discloses a non-negligible amount of private data. Therefore, it is necessary to develop techniques to provide/store personal data (user's point of view/designer's point of view) that

* This work was supported in part by NSERC of Canada.

yield the best customized services performance without compromising privacy. The goal of these techniques is to provide displayed data that will be used by customized services which contains as much non-private data as possible while revealing as little private data as possible. Also, for security reasons, the displayed data has to be produced using only non-private data. In general, this implies that the displayed data should be a randomized version of the non-private data.

To formulate this problem, we need to specify a privacy function and a utility function that respectively measure the amount of private and non-private data *leaked* into the displayed data. The authors of this paper recently suggested in [1] to use mutual information as the measure of both utility and privacy. Let X and Y denote the private and non-private data, respectively. The *rate-privacy function* $g_\varepsilon^{\mathsf{dis}}(X,Y)$ for discrete random variables X and Y having finite alphabets $\mathcal{X}$ and $\mathcal{Y}$, respectively is defined for any $\varepsilon \geq 0$ as the privacy-utility tradeoff

$$g_\varepsilon^{\mathsf{dis}}(X,Y) := \max_{\substack{P_{Z|Y}: X \multimap Y \multimap Z, \\ I(X;Z)\leq\varepsilon}} I(Y;Z), \tag{1}$$

where the auxiliary random variable Z is the privacy-constrained displayed data and $X \multimap Y \multimap Z$ denotes that X, Y, and Z form a Markov chain in this order. The channel $P_{Z|Y}$ is called the *privacy filter*. It is shown in [2] that $g_\varepsilon^{\mathsf{dis}}(X,Y)$ is in fact a corner point of an outer bound on the achievable region of the "dependence dilution" coding problem which provides an information-theoretic operational interpretation. It is also shown that if the channel from Y to X displays certain symmetry properties, then $g_\varepsilon^{\mathsf{dis}}(X,Y)$ can be calculated in closed form. For instance, if $P_{X|Y}$ is a binary symmetric channel (BSC) and $Y \sim \mathsf{Bernoulli}(0.5)$, then $g_\varepsilon^{\mathsf{dis}}(X,Y) = \frac{\varepsilon}{I(X;Y)}$.

As a more practical and operational notion of privacy, estimation-theoretic formulations of privacy are introduced in [3] and [4]. In particular, Calmon et al. [3] studied the case where $X = Y$ and defined the utility by $\Pr(\hat{Y}(Z) = Y)$ where $\hat{Y} : \mathcal{Z} \to \mathcal{Y}$ is the Bayes decoding map satisfying $I(Y;Z) \leq \varepsilon$ for discrete Y. Motivated by [5], which suggested the use of maximal correlation $\rho_m^2(X,Z)$ to measure the privacy level between X and Z, the authors in [4] recently generalized this model to arbitrary discrete X and Y, with the same utility function except that Z is required to satisfy $\rho_m^2(X,Z) \leq \varepsilon$. It was shown independently in [1] and [6] that if *perfect privacy* is required, i.e., Z must be statistically independent of X, then Z is also independent of Y unless the probability vectors $\{P_{Y|X}(\cdot|x) : x \in \mathcal{X}\}$ are linearly dependent (in which case Y is called *weakly independent* of X, see [7, Appendix II]). Hence, if Y is not weakly independent of X, then $g_0^{\mathsf{dis}}(X,Y) = 0$. Other formulations for privacy have appeared in [8,9,10,11,12,13].

The setting where (X,Y) is a pair of absolutely continuous random variables with $\mathcal{X} = \mathcal{Y} = \mathbb{R}$ is studied in [2] with both utility and privacy being measured by mutual information, and in [4], where both utility and privacy are measured in terms of the minimum mean-squared error (MMSE). In both cases, it is assumed that the privacy filter is an additive Gaussian channel with signal-to-noise ratio

(SNR) $\gamma \geq 0$, i.e.,

$$Z = Z_\gamma := \sqrt{\gamma} Y + N_{\mathsf{G}}, \tag{2}$$

where $N_{\mathsf{G}} \sim \mathcal{N}(0,1)$ is independent of (X,Y). In particular, the rate-privacy function [2] is defined as

$$g_\varepsilon(X,Y) := \max_{\substack{\gamma \geq 0, \\ I(X;Z_\gamma) \leq \varepsilon}} I(Y;Z_\gamma). \tag{3}$$

Letting $\mathsf{mmse}(U|V)$ denote the MMSE of estimating U by observing V and letting var denote the variance, the estimation-theoretic privacy-utility tradeoff is defined in [4] by the *estimation noise-to-signal ratio* (ENSR):

$$\mathsf{sENSR}_\varepsilon(X,Y) := \min \frac{\mathsf{mmse}(Y|Z_\gamma)}{\mathsf{var}(Y)}, \tag{4}$$

where the minimum is taken over all $\gamma \geq 0$ such that $\mathsf{mmse}(f(X)|Z_\gamma) \geq (1-\varepsilon)\mathsf{var}(f(X))$ for any non-constant measurable function $f : \mathcal{X} \to \mathbb{R}$. Unlike $g_\varepsilon(X,Y)$, $\mathsf{sENSR}_\varepsilon(X,Y)$ has a clear operational interpretation; it is the smallest MMSE associated with estimating Y given Z from which no non-degenerate function f of X can be estimated efficiently. This notion is related to *semantic security* [14] in cryptography. An encryption mechanism is said to be semantically secure if the adversary's advantage for correctly guessing any function of the private data given an observation of the mechanism's output (i.e., the ciphertext) is required to be negligible. As opposed to the discrete case, perfect privacy is achieved if and only if $\gamma = 0$, which gives rise to $g_0(X,Y) = 0$ (or equivalently $\mathsf{sENSR}_0(X,Y) = 1$) for any absolutely continuous (X,Y).

1.1 Contributions

In this work, we investigate the "almost" perfect privacy regime, that is, when $\varepsilon > 0$ is close to zero and derive a second-order approximation for $g_\varepsilon(X,Y)$ (Corollary 2). We also obtain the first and second derivatives of the mapping $\varepsilon \mapsto g_\varepsilon(X,Y)$ for $\varepsilon \in [0, I(X;Y))$ (Theorem 1). For a pair of Gaussian random variables (X,Y), an expression for $g_\varepsilon(X,Y)$ is derived (Example 1) and it is shown that the optimal filter has SNR equal to $\frac{2^{2\varepsilon}-1}{1-2^{-2(I(X;Y)-\varepsilon)}}$ for all $\varepsilon < I(X;Y)$ and the SNR is infinity if $\varepsilon \geq I(X;Y)$. Functional properties of the map $\varepsilon \mapsto g_\varepsilon(X,Y)$ are obtained (Proposition 1); in particular, it is shown than although the map $\varepsilon \mapsto g_\varepsilon^{\mathsf{dis}}(X,Y)$ is concave [2], the map $\varepsilon \mapsto g_\varepsilon(X,Y)$ is neither convex nor concave, and is infinitely differentiable (Corollary 1). Using a recent result on the strong data processing inequality by Anantharam et al. [15], a lower bound is obtained for $g_\varepsilon(X,Y)$. Assuming $P_{Y|X}$ is a convolution with a Gaussian distribution, i.e., $Y = aX + M_{\mathsf{G}}$, where $a \neq 0$ and $M_{\mathsf{G}} \sim \mathcal{N}(0,\sigma_M^2)$ is independent of X, we obtain an inequality relating $\mathsf{mmse}(Y|Z_\gamma, X)$ to $\mathsf{mmse}(Y|Z_\gamma)$ from which a stronger version of Anantharam's data processing inequality is derived for our setup (Theorem 2).

One main result of this paper is to connect $g_\varepsilon(X,Y)$ with $\mathsf{sENSR}_\varepsilon(X,Y)$ in the almost perfect privacy regime when X is Gaussian (Theorem 4). This connection allows us to translate the approximation obtained for $g_\varepsilon(X,Y)$ to a lower bound for $\mathsf{sENSR}_\varepsilon(X,Y)$.

1.2 Preliminaries

For a given pair of absolutely continuous random variables (U,V), we interchangeably use P_{UV} to denote the joint probability distribution and also the joint probability density function (pdf). The MMSE of estimating U given V is given by

$$\mathsf{mmse}(U|V) := \mathbb{E}[(U - \mathbb{E}[U|V])^2] = \mathbb{E}[\mathsf{var}(U|V)],$$

where $\mathsf{var}(U|V) = \mathbb{E}[(U - \mathbb{E}[U|V])^2|V]$. Guo et al. [16] proved the following so-called I-MMSE formula relating the input-output mutual information of the additive Gaussian channel $Z_\gamma = \sqrt{\gamma}Y + N_\mathsf{G}$, where $N_\mathsf{G} \sim \mathcal{N}(0,1)$ is independent of X, with the MMSE of the input given the output:

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} I(Y;Z_\gamma) = \frac{1}{2}\mathsf{mmse}(Y|Z_\gamma). \tag{5}$$

Since X, Y and Z_γ form the Markov chain $X \multimap Y \multimap Z_\gamma$, it follows that $I(X;Z_\gamma) = I(Y;Z_\gamma) - I(Y;Z_\gamma|X)$ and hence two applications of (5) yields [16, Theorem 10]

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} I(X;Z_\gamma) = \frac{1}{2}\left[\mathsf{mmse}(Y|Z_\gamma) - \mathsf{mmse}(Y|Z_\gamma, X)\right]. \tag{6}$$

The second derivative of $I(Y;Z_\gamma)$ and $I(X;Z_\gamma)$ are also known via the formula [17]

$$\frac{\mathrm{d}}{\mathrm{d}\gamma}\mathsf{mmse}(Y|Z_\gamma, X) = -\mathbb{E}[\mathsf{var}^2(Y|Z_\gamma, X)]. \tag{7}$$

Rényi [18] defined the *one-sided maximal correlation between U and V* (see also [13, Definition 7.4]) as

$$\eta_V^2(U) := \sup_g \rho^2(U, g(V)) = \frac{\mathsf{var}(\mathbb{E}[U|V])}{\mathsf{var}(U)}, \tag{8}$$

where $\rho(\cdot,\cdot)$ is the (Pearson) correlation coefficient, the supremum is taken over all measurable functions g, and the equality follows from the Cauchy-Schwarz inequality. The law of total variance implies that

$$\mathsf{mmse}(U|V) = \mathsf{var}(U)(1 - \eta_V^2(U)). \tag{9}$$

In an attempt of symmetrizing $\eta_V^2(U)$, Rényi [18] (see also [19] and [20]) defined the *maximal correlation* as

$$\rho_m^2(U,V) = \sup_{f,g} \rho^2(f(U), g(V)). \tag{10}$$

Comparing (8) with (10) reveals that

$$\rho^2(X,Y) \leq \eta_X^2(Y) \leq \rho_m^2(X,Y). \tag{11}$$

Clearly, unlike maximal correlation, $\eta_X(Y)$ is asymmetric, i.e., in general $\eta_X(Y) \neq \eta_Y(X)$, and hence according to Rényi's postulates [18], it is not a "proper" measure of dependence. However, it turns out to be an appropriate measure of separability between private and non-private information in the almost perfect privacy regime (see Corollary 2). On the other hand, maximal correlation satisfies all the Rényi's postulates [18]. In particular, it is symmetric and for jointly Gaussian random variables U and V with correlation coefficient ρ, we have $\rho_m^2(U,V) = \rho^2$.

2 Rate-Privacy Function for Additive Privacy Filters

Consider a pair of absolutely continuous random variables (X,Y) distributed according to P_{XY}. Let X and Y represent the *private data* and the *non-private data*, respectively. We think of X as having fixed distribution P_X and Y being generated by the channel $P_{Y|X}$, predefined by nature. Now consider the setting where Alice observes Y and wishes to describe it as accurately as possible to Bob in order to get a utility from him. Due to the correlation between Y and the private data X, Alice needs to provide Bob a noisy version Z of Y, such that Z cannot reveal more than ε bits of information about X. In fact, we assume that Z is obtained via the privacy filter, $Z = Z_\gamma$ defined in (2). The aim is to pick $\gamma \geq 0$ such that Z_γ preserves the maximum amount of the information about Y while satisfying the privacy constraint. The rate-privacy function $g_\varepsilon(X,Y)$, defined in (3), quantifies the tradeoff between these conflicting goals [2]. Note that since $I(Y;Z_\gamma) = I(Y;Y+\frac{1}{\sqrt{\gamma}}N_{\mathsf{G}})$, we can interpret $\frac{1}{\gamma}$ as the noise variance. Due to the data processing inequality, one can restrict ε to the interval $[0, I(X;Y))$ in the definition of $g_\varepsilon(X,Y)$ and consequently for any $\varepsilon \geq I(X;Y)$ the optimal noise variance must be zero and hence $g_\varepsilon(X,Y) = \infty$. The case where the displayed data is required to carry no information at all about X, i.e., where $\varepsilon = 0$, is often called *perfect privacy*.

The maps $\gamma \mapsto I(Y;Z_\gamma)$ and $\gamma \mapsto I(X;Z_\gamma)$ are strictly increasing over $[0,\infty)$ [2, Lemmas 16, 17] and hence there exists a unique $\gamma_\varepsilon \in [0,\infty)$ such that $I(X;Z_{\gamma_\varepsilon}) = \varepsilon$ and $g_\varepsilon(X,Y) = I(Y;Z_{\gamma_\varepsilon})$. This observation yields the following proposition.

Proposition 1. *For absolutely continuous random variables (X,Y), we have*

1. *The map $\varepsilon \mapsto \gamma_\varepsilon$ is strictly increasing and continuous, and it satisfies $\gamma_0 = 0$ and $\gamma_{I(X;Y)} = \infty$.*
2. *The map $\varepsilon \mapsto g_\varepsilon(X,Y)$ is non-negative, increasing and, continuous on $[0, I(X;Y))$, and it satisfies $g_0(X,Y) = 0$ and $g_{I(X;Y)}(X,Y) = \infty$.*
3. *Let $D(Y)$ denote the "non-Gaussianness" of Y, defined as $D(Y) := D(P_Y||P_{Y_{\mathsf{G}}})$ (here $D(\cdot||\cdot)$ is the Kullback-Leibler divergence) with Y_{G} being a Gaussian*

random variable having the same mean and variance as Y. Then we have

$$\frac{1}{2}\log\left(1+\gamma_\varepsilon 2^{-2D(Y)}\mathsf{var}(Y)\right) \le g_\varepsilon(X,Y) \le \frac{1}{2}\log(1+\gamma_\varepsilon \mathsf{var}(Y)).$$

Proof. Parts 1 and 2 can be proved directly from continuity and strict monotonicity of the maps $\gamma \mapsto I(Y;Z_\gamma)$ and $\gamma \mapsto I(X;Z_\gamma)$. The upper bound in part 3 is a direct consequence of the fact that a Gaussian input maximizes the mutual information between input and output of an additive Gaussian channel. The lower bound follows from the entropy power inequality [21, Theorem 17.7.3] which states that $2^{2h(Z_\gamma)} \ge \gamma 2^{2h(Y)} + 2\pi e$ and hence

$$g_\varepsilon(X,Y) = I(Y;Z_{\gamma_\varepsilon}) \le \frac{1}{2}\log\left(\gamma_\varepsilon 2^{2h(Y)} + 2\pi e\right) - \frac{1}{2}\log(2\pi e),$$

from which and the fact that $D(Y) = h(Y_\mathsf{G}) - h(Y)$, the lower bound immediately follows. □

In light of Proposition 1, it is clear that, unless X and Y are independent, Z_γ is independent of X if and only if $\gamma = 0$, which implies $g_0(X,Y) = 0$. As mentioned in the introduction, this is in contrast with the discrete rate-privacy function (1), where $g_0^{\mathsf{dis}}(X,Y)$ may be positive (for example, when Y is an erased version of X, see [2, Lemma 12]).

Example 1. Let $(X_\mathsf{G}, Y_\mathsf{G})$ be a pair of Gaussian random variables with zero mean and correlation coefficient ρ. Then Z_γ is also a Gaussian random variable with variance $\gamma\mathsf{var}(Y_\mathsf{G}) + 1$. Without loss of generality assume that Y_G has unit variance. Then

$$I(X_\mathsf{G};Z_\gamma) = \frac{1}{2}\log\left(\frac{\gamma+1}{\gamma-\gamma\rho^2+1}\right),$$

and hence for any $\varepsilon \in [0, I(X_\mathsf{G};Y_\mathsf{G}))$ the equation $I(X_\mathsf{G};Z_\gamma) = \varepsilon$ has the unique solution

$$\gamma_\varepsilon = \frac{1-2^{-2\varepsilon}}{2^{-2\varepsilon}+\rho^2-1}.$$

Thus, we obtain

$$\begin{aligned} g_\varepsilon(X_\mathsf{G},Y_\mathsf{G}) &= \frac{1}{2}\log(1+\gamma_\varepsilon) = \frac{1}{2}\log\left(\frac{\rho^2}{2^{-2\varepsilon}+\rho^2-1}\right) \\ &= \frac{1}{2}\log\left(1+\frac{2^{2\varepsilon}-1}{1-2^{-2(I(X_\mathsf{G};Y_\mathsf{G})-\varepsilon)}}\right). \end{aligned} \tag{12}$$

The graph of $g_\varepsilon(X_\mathsf{G},Y_\mathsf{G})$ is depicted in Fig. 1 for $\rho = 0.5$ and $\rho = 0.8$. It is worth noting that $g_\varepsilon(X_\mathsf{G},Y_\mathsf{G})$ is related to the Gaussian rate-distortion function $R_\mathsf{G}(D)$ [21]. In fact, $g_\varepsilon(X_\mathsf{G},Y_\mathsf{G}) = R_\mathsf{G}(D_\varepsilon)$ for $\varepsilon \le I(X_\mathsf{G};Y_\mathsf{G})$ where

$$D_\varepsilon = \frac{2^{-2\varepsilon} - 2^{-2I(X_\mathsf{G};Y_\mathsf{G})}}{\rho^2},$$

is the mean squared distortion incurred in reconstructing Y given the displayed data Z_γ.

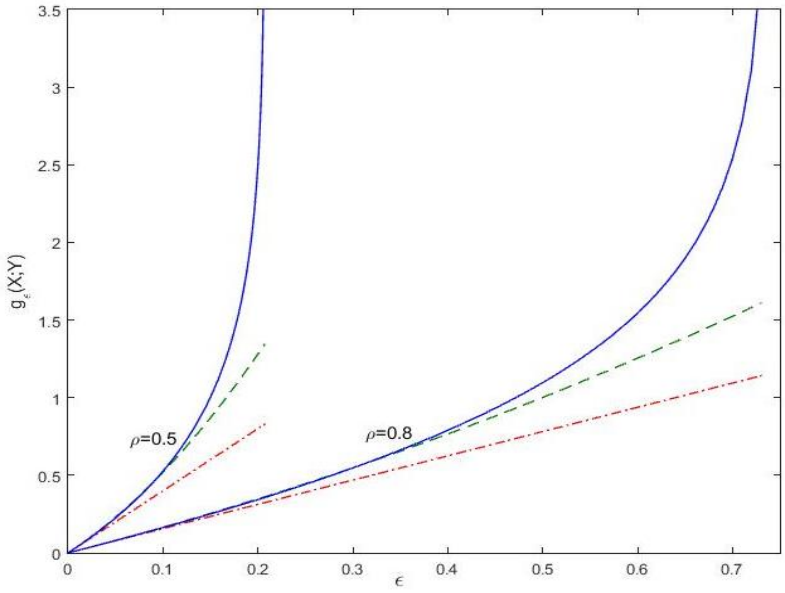

Fig. 1. The rate-privacy function for a pair of Gaussian $(X_{\mathsf{G}}, Y_{\mathsf{G}})$, given by (12), for $\rho = 0.5$ and $\rho = 0.8$. The first and second-order approximations are also shown in red and green, respectively.

The next result provides the first derivative $g'_\varepsilon(X,Y)$ of the function $\varepsilon \mapsto g_\varepsilon(X,Y)$ at any $\varepsilon < I(X;Y)$.

Theorem 1. *For any absolutely continuous random variables (X,Y), we have*

$$g'_\varepsilon(X,Y) = \frac{\mathsf{mmse}(Y|Z_{\gamma_\varepsilon})}{\mathsf{mmse}(Y|Z_{\gamma_\varepsilon}) - \mathsf{mmse}(Y|Z_{\gamma_\varepsilon}, X)}.$$

Proof. Since $g_\varepsilon(X,Y) = I(Y;Z_{\gamma_\varepsilon})$, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon} g_\varepsilon(X,Y) &= \left[\frac{\mathrm{d}}{\mathrm{d}\gamma} I(Y;Z_\gamma)\right]_{\gamma=\gamma_\varepsilon} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\gamma_\varepsilon \\ &\overset{(a)}{=} \frac{1}{2}\mathsf{mmse}(Y|Z_{\gamma_\varepsilon})\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\gamma_\varepsilon, \end{aligned} \tag{13}$$

where (a) follows from (5). In order to calculate $\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\gamma_\varepsilon$, notice that $\varepsilon = I(X;Z_{\gamma_\varepsilon})$ and hence taking the derivative of both sides of this equation with respect to ε yields

$$1 = \left[\frac{\mathrm{d}}{\mathrm{d}\gamma} I(X;Z_\gamma)\right]_{\gamma=\gamma_\varepsilon} \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\gamma_\varepsilon,$$

and hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\gamma_\varepsilon &= \frac{1}{\left[\frac{\mathrm{d}}{\mathrm{d}\gamma}I(X;Z_\gamma)\right]_{\gamma=\gamma_\varepsilon}} \\ &\overset{(a)}{=} \frac{2}{\mathsf{mmse}(Y|Z_{\gamma_\varepsilon}) - \mathsf{mmse}(Y|Z_{\gamma_\varepsilon},X)},\end{aligned} \tag{14}$$

where (a) follows from (6). The result then follows by plugging (14) into (13). □

As a simple illustration of Theorem 1, consider jointly Gaussian X_G and Y_G whose rate-privacy function is computed in Example 1. In particular, (12) gives

$$g'_\varepsilon(X_\mathsf{G},Y_\mathsf{G}) = \frac{2^{-2\varepsilon}}{2^{-2\varepsilon}+\rho^2-1}. \tag{15}$$

On the other hand, since $X_\mathsf{G} = \sqrt{\alpha}Y_\mathsf{G} + N_1$ where $\alpha = \rho^2\mathsf{var}(X)$, $N_1 \sim \mathcal{N}(0,\sigma_N^2)$ is independent of Y_G, and $\sigma_N^2 = (1-\rho^2)\mathsf{var}(X)$, one can conclude from [16, Proposition 3] that

$$\mathsf{mmse}(Y_\mathsf{G}|Z_\gamma,X_\mathsf{G}) = \mathsf{mmse}\left(Y_\mathsf{G}|Z_\gamma,\frac{1}{\sigma_N^2}X_\mathsf{G}\right) = \mathsf{mmse}(Y_\mathsf{G}|Z_{\gamma+a}),$$

where $a = \frac{\rho^2}{1-\rho^2}$. Recalling that $\mathsf{mmse}(Y_\mathsf{G}|Z_\gamma) = \frac{1}{1+\gamma}$, we obtain

$$\begin{aligned}\frac{\mathsf{mmse}(Y_\mathsf{G}|Z_\gamma)}{\mathsf{mmse}(Y_\mathsf{G}|Z_\gamma) - \mathsf{mmse}(Y_\mathsf{G}|Z_{\gamma+a})} &= \frac{1+(1-\rho^2)\gamma_\varepsilon}{\rho^2} \\ &= \frac{2^{-2\varepsilon}}{2^{-2\varepsilon}+\rho^2-1},\end{aligned}$$

which equals (15).

In light of Theorem 1, we can now show that the map $\varepsilon \mapsto g_\varepsilon(X,Y)$ is in fact infinitely differentiable over $(0, I(X;Y))$.

Corollary 1. *For a pair of absolutely continuous (X,Y), the map $\varepsilon \mapsto g_\varepsilon(X,Y)$ is infinitely differentiable at any $\varepsilon \in (0,I(X;Y))$. Moreover, if all the moments of Y is finite, then $\varepsilon \mapsto g_\varepsilon(X,Y)$ is infinitely right differentiable at $\varepsilon = 0$.*

Proof. It is shown in [17, Proposition 7] that $\gamma \mapsto \mathsf{mmse}(Y|Z_\gamma)$ is infinitely differentiable at any $\gamma > 0$ and infinitely right differentiable at $\gamma = 0$ if all the moments of Y are finite. Thus the corollary follows from Theorem 1 noting that since $\mathbb{E}[Y^k] < \infty$ for all k, we also have $\mathbb{E}[Y^k|X=x] < \infty$ for almost all x (except for x in a set of zero P_X-measure). It therefore follows that $\gamma \mapsto \mathsf{mmse}(Y|Z_\gamma,X)$ is also infinitely right differentiable at $\gamma = 0$. □

We remark that using (7) and Theorem 1, one can easily calculate the second derivative as

$$\begin{aligned}g''_\varepsilon(X,Y) &= \frac{\mathrm{d}^2}{\mathrm{d}\varepsilon^2}g_\varepsilon(X,Y) \\ &= \frac{2\left(\mathsf{mmse}(Y|Z_{\gamma_\varepsilon},X)\mathbb{E}[\mathsf{var}^2(Y|Z_{\gamma_\varepsilon})] - \mathsf{mmse}(Y|Z_{\gamma_\varepsilon})\mathbb{E}[\mathsf{var}^2(Y|Z_{\gamma_\varepsilon},X)]\right)}{[\mathsf{mmse}(Y|Z_{\gamma_\varepsilon}) - \mathsf{mmse}(Y|Z_{\gamma_\varepsilon},X)]^3}.\end{aligned} \tag{16}$$

The following corollary, which is an immediate consequence of Theorem 1, provides a second-order approximation for $g_\varepsilon(X,Y)$ as $\varepsilon \downarrow 0$ and thus an approximation to the the rate-privacy function in the almost perfect privacy regime.

Corollary 2. *For a given pair of absolutely continuous random variables (X,Y), we have as $\varepsilon \downarrow 0$,*

$$g_\varepsilon(X,Y) = \frac{\varepsilon}{\eta_X^2(Y)} + \Delta(X,Y)\varepsilon^2 + o(\varepsilon^2),$$

where

$$\Delta(X,Y) = \frac{1}{\eta_X^4(Y)}\left(\frac{\mathsf{var}^2(Y) - \mathbb{E}[\mathsf{var}^2(Y|X)]}{\mathsf{var}^2(Y)\eta_X^2(Y)} - 1\right), \tag{17}$$

and $\eta_X^2(Y)$ is the one-sided maximal correlation between X and Y defined in (8).

Proof. According to Corollary 1, we can use the second-order Taylor expansion to approximate $g_\varepsilon(X,Y)$ around $\varepsilon = 0$, resulting in

$$g_\varepsilon(X,Y) = \varepsilon g_0'(X,Y) + \frac{\varepsilon^2}{2} g_0''(X,Y) + o(\varepsilon^2).$$

From Theorem 1 and (16) we have $g_0'(X,Y) = \frac{1}{\eta_X^2(Y)}$ and $g_0''(X,Y) = 2\Delta(X,Y)$, respectively, from which the corollary follows. □

Since $\rho_m^2(X_\mathsf{G}, Y_\mathsf{G}) = \rho^2$ for jointly Gaussian X_G and Y_G with correlation coefficient ρ, (11) implies that $\eta_{X_\mathsf{G}}^2(Y_\mathsf{G}) = \rho^2$ and $\Delta(X_\mathsf{G}, Y_\mathsf{G}) = \frac{1-\rho^2}{\rho^4}$, and therefore Corollary 2 implies that for small $\varepsilon > 0$,

$$g_\varepsilon(X_\mathsf{G}, Y_\mathsf{G}) = \frac{1}{\rho^2}\varepsilon + \frac{1-\rho^2}{\rho^4}\varepsilon^2 + o(\varepsilon^2).$$

This second-order approximation as well as the first-order approximation are illustrated in Fig. 1 for $\rho = 0.5$ and $\rho = 0.8$.

Polyanskiy and Wu [22] have recently generalized the strong data processing inequality of Anantharam et al. [15] for the case of continuous random variables X and Y with joint distribution P_{XY}. Their result states that

$$\sup_{\substack{X \multimap Y \multimap U,\\ 0<I(U;Y)<\infty}} \frac{I(X;U)}{I(Y;U)} = S^*(Y,X), \tag{18}$$

where

$$S^*(Y,X) := \sup_{\substack{Q_Y,\\ 0<D(Q_Y||P_Y)<\infty}} \frac{D(Q_X||P_X)}{D(Q_Y||P_Y)},$$

where P_X and P_Y are the marginals of P_{XY} and $Q_X(\cdot) = \int P_{X|Y}(\cdot|y)Q_Y(\mathrm{d}y)$. In addition, it is shown in [22] that the supremum in (18) is achieved by a binary U. Replacing U with Z_γ, we can conclude from (18) that

$$\frac{I(X;Z_\gamma)}{I(Y;Z_\gamma)} \le S^*(Y,X),$$

for any $\gamma \geq 0$. Letting $\gamma = \gamma_\varepsilon$, the above yields that

$$g_\varepsilon(X,Y) \geq \frac{\varepsilon}{S^*(Y,X)}. \tag{19}$$

Clearly, this bound may be expected to be tight only for small $\varepsilon > 0$ since $g_\varepsilon(X,Y) \to \infty$ as $\varepsilon \to I(X;Y)$, as shown in Proposition 1. Note that Theorem 1 implies $\lim_{\varepsilon \downarrow 0} \frac{g_\varepsilon(X,Y)}{\varepsilon} = \frac{1}{\eta_X^2(Y)}$. On the other hand, it can be easily shown that $\eta_X^2(Y) \leq S^*(Y,X)$, with equality when X and Y are jointly Gaussian and hence the inequality (19) becomes tight for small ε and jointly Gaussian X and Y.

The bound in (19) would be significantly improved if we could show that $g_\varepsilon(X,Y) \geq g_\varepsilon(X_\mathsf{G}, Y_\mathsf{G})$, where X_G and Y_G are jointly Gaussian having the same means, variances, and correlation coefficient as (X,Y). This is because in that case we could write

$$g_\varepsilon(X,Y) \geq g_\varepsilon(X_\mathsf{G}, Y_\mathsf{G}) \overset{(a)}{\geq} \frac{\varepsilon}{\eta_{X_\mathsf{G}}^2(Y_\mathsf{G})} = \frac{\varepsilon}{\rho^2(X_\mathsf{G}, Y_\mathsf{G})} = \frac{\varepsilon}{\rho^2(X,Y)} \overset{(b)}{\geq} \frac{\varepsilon}{\eta_X^2(Y)}, \tag{20}$$

where (a) and (b) follow from (12) and (11), respectively. However, as shown in Appendix A, the inequality $g_\varepsilon(X,Y) \geq g_\varepsilon(X_\mathsf{G}, Y_\mathsf{G})$ does not in general hold[1]. It is therefore possible to have $g_\varepsilon(X,Y) < \frac{\varepsilon}{\eta_X^2(Y)}$ for some $0 < \varepsilon < I(X;Y)$. To construct an example, it suffices to construct P_{XY} for which $\varepsilon \mapsto g_\varepsilon(X,Y)$ is locally concave at zero (i.e., $g_0''(X,Y) < 0$) and hence its graph lies below the tangent line $\frac{\varepsilon}{\eta_X^2(Y)}$ for some $\varepsilon > 0$. Let $Y \sim \mathcal{N}(0,1)$ and $X = Y \cdot 1_{\{Y \in [-1,1]\}}$. Then it can be readily shown that $\mathbb{E}[\mathsf{var}(Y|X)] < \mathbb{E}[\mathsf{var}^2(Y|X)]$, which implies that $\Delta(X,Y) < 0$. Hence, since $g_0''(X,Y) = 2\Delta(X,Y)$, we have that $g''(X,Y) < 0$. This observation is illustrated in Fig. 2.

As remarked earlier, the map $\varepsilon \mapsto g_\varepsilon(X,Y)$ is in general not convex and thus one cannot conclude that $g_\varepsilon'(X,Y) \geq g_0'(X,Y) = \frac{1}{\eta_X^2(Y)}$. However, it can be shown that this implication holds if P_{XY} has more structure. In the next theorem, we assume that Y is a noisy version of X through an additive Gaussian channel.

Theorem 2. *For a given $X \sim P_X$ with variance σ_X^2, and $Y = aX + M_\mathsf{G}$ with $M_\mathsf{G} \sim \mathcal{N}(0, \sigma_M^2)$ independent of X, we have:*

1. *If $a^2\sigma_X^2 \geq \sigma_M^2$, then $\varepsilon \mapsto g_\varepsilon(X,Y)$ is convex.*
2. *For any $a > 0$ and $\varepsilon \in [0, I(X;Y))$, we have*

$$g_\varepsilon(X,Y) \geq \frac{\varepsilon}{\eta_X^2(Y)}. \tag{21}$$

Furthermore, we have

$$\inf_{\gamma \geq 0} \frac{\mathsf{mmse}(Y|Z_\gamma, X)}{\mathsf{mmse}(Y|Z_\gamma)} = 1 - \eta_X^2(Y), \tag{22}$$

[1] We will see in the next section that this holds in the estimation-theoretic formulation of privacy, i.e., the Gaussian case is the *worst* case when the privacy filter is an additive Gaussian channel and the utility and privacy are measured as $\mathsf{mmse}(Y|Z_\gamma)$ and $\mathsf{mmse}(X|Z_\gamma)$, respectively.

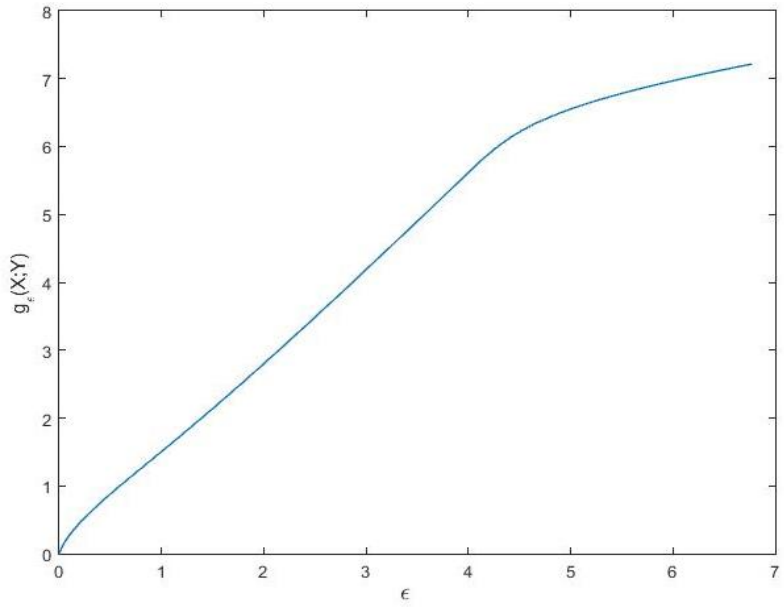

Fig. 2. The rate-privacy function for $Y \sim \mathcal{N}(0,1)$ and $X = Y \cdot 1_{\{Y\in[-1,1]\}}$. The map $\varepsilon \mapsto g_\varepsilon(X,Y)$ is clearly locally concave at zero. Note that here $I(X;Y) = \infty$ and hence ε is unbounded.

and

$$\sup_{\gamma>0} \frac{I(X;Z_\gamma)}{I(Y;Z_\gamma)} = \eta_X^2(Y). \tag{23}$$

Proof. The first part follows from a straightforward computation showing that if $a^2\mathsf{var}(X) \geq \sigma_M^2$, then $\Delta(X,Y) \geq 0$.

To prove the second part, note that for any $\gamma \geq 0$ we have

$$\begin{aligned}
\mathsf{mmse}(Y|Z_\gamma) &= \mathsf{mmse}(aX + M_\mathsf{G}|a\sqrt{\gamma}X + \sqrt{\gamma}M_\mathsf{G} + N_\mathsf{G}) \\
&\overset{(a)}{=} \frac{1}{\gamma}\mathsf{mmse}\left(N_\mathsf{G}|a\sqrt{\gamma}X + \sqrt{\gamma}M_\mathsf{G} + N_\mathsf{G}\right) \\
&\overset{(b)}{\leq} \frac{a^2\mathsf{var}(X) + \sigma_M^2}{1 + \gamma(a^2\mathsf{var}(X) + \sigma_M^2)} < \frac{a^2\mathsf{var}(X) + \sigma_M^2}{1 + \gamma\sigma_M^2} \\
&\overset{(c)}{=} \frac{1}{\gamma}\left(\frac{a^2\mathsf{var}(X) + \sigma_M^2}{\sigma_M^2}\right)\mathsf{mmse}\left(N_\mathsf{G}|\sqrt{\gamma}M_\mathsf{G} + N_\mathsf{G}\right) \\
&\overset{(d)}{=} \left(\frac{a^2\mathsf{var}(X) + \sigma_M^2}{\sigma_M^2}\right)\mathsf{mmse}(Y|Z_\gamma, X),
\end{aligned} \tag{24}$$

where (a) follows from the fact that $\mathsf{mmse}(U|\alpha U + V) = \frac{1}{\alpha^2}\mathsf{mmse}(V|\alpha U + V)$ for $\alpha \neq 0$, (b) and (c) hold by [23, Theorem 12] which states that $\mathsf{mmse}(U|U + V_\mathsf{G}) \leq$

$\mathsf{mmse}(U_\mathsf{G}|U_\mathsf{G}+V_\mathsf{G}) = \frac{\mathsf{var}(U)\mathsf{var}(V)}{\mathsf{var}(U)+\mathsf{var}(V)}$. Finally, (d) follows from the following chain of equalities

$$\begin{aligned}\mathsf{mmse}(Y|Z_\gamma, X) &= \mathsf{mmse}(aX + M_\mathsf{G}|a\sqrt{\gamma}X + \sqrt{\gamma}M_\mathsf{G} + N_\mathsf{G}, X)\\ &= \mathsf{mmse}(M_\mathsf{G}|\sqrt{\gamma}M_\mathsf{G} + N_\mathsf{G}, X)\\ &\overset{(e)}{=} \mathsf{mmse}(M_\mathsf{G}|\sqrt{\gamma}M_\mathsf{G} + N_\mathsf{G})\\ &= \frac{1}{\gamma}\mathsf{mmse}(N_\mathsf{G}|\sqrt{\gamma}M_\mathsf{G} + N_\mathsf{G})\end{aligned}$$

where (e) holds since X and M_G are independent.

We can therefore write

$$\begin{aligned}g'_\varepsilon(X,Y) &= \frac{\mathsf{mmse}(Y|Z_{\gamma_\varepsilon})}{\mathsf{mmse}(Y|Z_{\gamma_\varepsilon}) - \mathsf{mmse}(Y|Z_{\gamma_\varepsilon}, X)}\\ &\overset{(a)}{\geq} \frac{a^2\mathsf{var}(X)+\sigma_M^2}{a^2\mathsf{var}(X)} \overset{(b)}{=} \frac{1}{\eta_X^2(Y)} = g'_0(X,Y),\end{aligned} \tag{25}$$

where (a) is due to (24) and (b) holds since $\mathsf{var}(Y) = a^2\mathsf{var}(X) + \sigma_M^2$ and $\mathsf{var}(\mathbb{E}[Y|X]) = a^2\mathsf{var}(X)$. The identity $g_\varepsilon(X,Y) = \int_0^\varepsilon g'_t(X,Y)\mathrm{d}t$, and inequality (25) together imply that $g_\varepsilon(X,Y) \geq \frac{\varepsilon}{\eta_X^2(Y)}$ for $\varepsilon \leq I(X;Y)$.

Furthermore, according to Theorem 1, the inequality (25) yields (22). Using the integral representation of mutual information in (5) and (6), we can write for any $\gamma \geq 0$

$$\begin{aligned}I(X;Z_\gamma) &= \frac{1}{2}\int_0^\gamma [\mathsf{mmse}(Y|Z_t) - \mathsf{mmse}(Y|Z_t, X)]\,\mathrm{d}t\\ &\leq \frac{\eta_X^2(Y)}{2}\int_0^\gamma \mathsf{mmse}(Y|Z_t)\mathrm{d}t = \eta_X^2(Y)I(Y;Z_\gamma),\end{aligned} \tag{26}$$

where the inequality is due to (22). The equality (23) then follows from (26). □

It should be noted that both MMSE and mutual information satisfy the data processing inequality, see, [23] and [15], that is, $\mathsf{mmse}(U|V) \leq \mathsf{mmse}(U|W)$, and $I(U;W) \leq I(U;V)$ for $U \multimap V \multimap W$. Therefore, (22) can be thought of as a strong version of the data processing inequality for MMSE for the trivial Markov chain $Y \multimap (Z_\gamma, X) \multimap Z_\gamma$. Also, (23) can be viewed as a strong data processing inequality for the mutual information for the Markov chain $X \multimap Y \multimap Z_\gamma$ which is slightly stronger than (18) in the special case of an additive Gaussian channel as $\eta_X^2(Y) \leq S^*(Y,X)$.

3 Estimation-Theoretic Formulation

Consider the same scenario as in the previous section: Alice observes Y, which is correlated with the private data X according to a given joint distribution P_{XY}, and wishes to transmit a random variable Z to Bob to receive a utility from

him. An *operational* measure of privacy is proposed in [4] where Alice generates the displayed data Z via a privacy filter $P_{Z|Y}$ such that Bob cannot efficiently estimate any non-trivial function of X given Z. As before, her goal is to maximize the utility (or equivalently minimize the cost) between Y and the displayed data Z. The next definition formalizes this privacy guarantee. We call a function f of random variable X *non-degenerate* if $f(X)$ is not almost everywhere constant with respect to the probability measure P_X. Also, we assume throughout this section that X and Y have finite second moments.

Definition 1. *Given a pair of jointly absolutely continuous random variables (X,Y) with joint distribution P_{XY} and $0 \leq \varepsilon \leq 1$, we say Z satisfies* ε-strong estimation privacy, *if there exists a channel $P_{Z|Y}$ that induces a joint distribution $P_X \times P_{Z|X}$, via the Markov condition $X \multimap Y \multimap Z$, satisfying*

$$\mathsf{mmse}(f(X)|Z) \geq (1-\varepsilon)\mathsf{var}(f(X)), \quad \textit{or equivalently,} \quad \eta_Z^2(f(X)) \leq \varepsilon, \tag{27}$$

for any non-degenerate Borel function f. Similarly, Z is said to satisfy ε-weak estimation privacy, *if* (27) *is satisfied only for the identity function $f(x) = x$.*

It is shown in [4] that ε-strong estimation privacy is equivalently characterized by the requirement $\rho_m^2(X,Z) \leq \varepsilon$. In other words, $\mathsf{mmse}(f(X)|Z) \geq (1-\varepsilon)\mathsf{var}(f(X))$ for any non-degenerate Borel function f if and only if $\rho_m^2(X,Z) \leq \varepsilon$. Let the utility that Alice receives from Bob be measured by $\frac{\mathsf{var}(Y)}{\mathsf{mmse}(Y|Z)}$, which she aims to maximize. For mathematical convenience, we define the *cost* that Alice suffers by describing Z in lieu of Y as the estimation noise-to-signal ratio (ENSR), $\frac{\mathsf{mmse}(Y|Z)}{\mathsf{var}(Y)}$, and hence Alice equivalently aims to minimize the ENSR. Focusing on additive Gaussian privacy filter $Z = Z_\gamma$, we can formalize the privacy-utility tradeoff as

$$\mathsf{sENSR}_\varepsilon(X,Y) := \inf_{\gamma \in \mathcal{C}_\varepsilon(P_{XY})} \frac{\mathsf{mmse}(Y|Z_\gamma)}{\mathsf{var}(Y)} = 1 - \sup_{\gamma \in \mathcal{C}_\varepsilon(P_{XY})} \eta_{Z_\gamma}^2(Y),$$

where $\mathcal{C}_\varepsilon(P_{XY})$ is the set of parameters γ corresponding to ε-strong privacy, i.e.,

$$\mathcal{C}_\varepsilon(P_{XY}) := \{\gamma \geq 0 : \rho_m^2(X, Z_\gamma) \leq \varepsilon\}.$$

Similarly,

$$\mathsf{wENSR}_\varepsilon(X,Y) := 1 - \sup_{\gamma \in \partial\mathcal{C}_\varepsilon(P_{XY})} \eta_{Z_\gamma}^2(Y),$$

where

$$\partial\mathcal{C}_\varepsilon(P_{XY}) := \{\gamma \geq 0 : \eta_{Z_\gamma}^2(X) \leq \varepsilon\}.$$

Note that both the maximal correlation and the one-sided maximal correlation satisfy the data processing inequality, that is, $\rho_m^2(X, Z_\gamma) \leq \rho_m^2(Y, Z_\gamma)$ and $\eta_{Z_\gamma}^2(X) \leq \eta_Y(X)$. Therefore, in the definition of $\mathsf{sENSR}_\varepsilon(X,Y)$ and $\mathsf{wENSR}_\varepsilon(X,Y)$, we can restrict ε as $0 \leq \varepsilon \leq \rho_m^2(X,Y)$ and $0 \leq \varepsilon \leq \eta_Y^2(X)$, respectively.

Example 2. Let X_{G} and Y_{G} be jointly Gaussian with correlation coefficient ρ. Without loss of generality assume that $\mathbb{E}[X_{\mathsf{G}}] = \mathbb{E}[Y_{\mathsf{G}}] = 0$. Since $\rho_m^2(X_{\mathsf{G}}, Z_\gamma) = \rho^2(X_{\mathsf{G}}, Z_\gamma)$, we have

$$\rho_m^2(X_{\mathsf{G}}, Z_\gamma) = \rho^2 \frac{\gamma \mathsf{var}(Y_{\mathsf{G}})}{1 + \gamma \mathsf{var}(Y_{\mathsf{G}})},$$

which implies that the mapping $\gamma \mapsto \rho_m^2(X_{\mathsf{G}}, Z_\gamma)$ is strictly increasing. Also, the equation $\rho_m^2(X_{\mathsf{G}}, Z_\gamma) = \varepsilon$ for $0 \le \varepsilon \le \rho_m^2(X_{\mathsf{G}}, Y_{\mathsf{G}}) = \rho^2$ has a unique solution

$$\gamma_\varepsilon := \frac{\varepsilon}{\mathsf{var}(Y_{\mathsf{G}})(\rho^2 - \varepsilon)},$$

and $\rho_m^2(X, Z_\gamma) \le \varepsilon$ for any $\gamma \le \gamma_\varepsilon$. On the other hand, $\mathsf{mmse}(Y_{\mathsf{G}}|Z_\gamma) = \frac{\mathsf{var}(Y_{\mathsf{G}})}{1+\gamma \mathsf{var}(Y_{\mathsf{G}})}$, which shows that the map $\gamma \mapsto \mathsf{mmse}(Y_{\mathsf{G}}|Z_\gamma)$ is strictly decreasing. Hence,

$$\mathsf{sENSR}_\varepsilon(X_{\mathsf{G}}, Y_{\mathsf{G}}) = \frac{\mathsf{mmse}(Y_{\mathsf{G}}|Z_{\gamma_\varepsilon})}{\mathsf{var}(Y_{\mathsf{G}})} = 1 - \frac{\varepsilon}{\rho^2}. \tag{28}$$

Clearly for jointly Gaussian X_{G} and Y_{G} we have $\eta_{Z_\gamma}^2(X_{\mathsf{G}}) = \rho_m^2(X_{\mathsf{G}}, Z_\gamma) = \varepsilon$, for any $\gamma \ge 0$ and consequently $\mathcal{C}_\varepsilon(P_{X_{\mathsf{G}}Y_{\mathsf{G}}}) = \partial\mathcal{C}_\varepsilon(P_{X_{\mathsf{G}}Y_{\mathsf{G}}})$, that is, for $0 \le \varepsilon \le \rho^2$,

$$\mathsf{sENSR}_\varepsilon(X_{\mathsf{G}}, Y_{\mathsf{G}}) = \mathsf{wENSR}_\varepsilon(X_{\mathsf{G}}, Y_{\mathsf{G}}) = 1 - \frac{\varepsilon}{\rho^2}. \tag{29}$$

Unlike $g_\varepsilon(X, Y)$, the quantity $\mathsf{sENSR}_\varepsilon(X, Y)$ is maximized among all pairs of random variables (X, Y) with identical means, variances and correlation coefficient when X and Y are jointly Gaussian. Thus, Example 2 yields a sharp upper-bound for $\mathsf{sENSR}_\varepsilon(X, Y)$. This is stated in the following theorem.

Theorem 3 ([4]). *For any given jointly absolutely continuous (X, Y), we have for $0 \le \varepsilon \le \rho_m^2(X, Y)$,*

$$\mathsf{wENSR}_\varepsilon(X, Y) \le \mathsf{sENSR}_\varepsilon(X, Y) \le \mathsf{sENSR}_\varepsilon(X_{\mathsf{G}}, Y_{\mathsf{G}}) = 1 - \frac{\varepsilon}{\rho_m^2(X, Y)},$$

where $(X_{\mathsf{G}}, Y_{\mathsf{G}})$ is a pair of Gaussian random variables with the same means, variances, and correlation coefficient as (X, Y).

Next, we turn our attention to the approximation of $\mathsf{sENSR}_\varepsilon(X, Y)$ in the almost perfect privacy regime. Unfortunately, there is no known approximation for $\rho_m^2(X, Z_\gamma)$ and $\mathsf{mmse}(X|Z_\gamma)$ around $\gamma = 0$. Nevertheless, we can use the first-order approximation of $g_\varepsilon(X, Y)$ to derive an approximation for $\mathsf{sENSR}_\varepsilon(X, Y)$ around $\varepsilon = 0$. The next theorem shows this approximation for the special case where $P_{Y|X}$ is an additive noise channel.

Theorem 4. *If $X \sim \mathcal{N}(b, \sigma_X^2)$ and $Y = aX + M$, where $a, b \in \mathbb{R}^+$, and M is a noise random variable having a density, then for sufficiently small ε*

$$\mathsf{sENSR}_\varepsilon(X_{\mathsf{G}}, Y) \ge 2^{-D(Y)} 2^{-2g_{\varepsilon + o(\varepsilon)}(X_{\mathsf{G}}, Y)}. \tag{30}$$

Proof. We start by deriving an inequality relating $\mathsf{mmse}(Y|Z_\gamma)$ and $I(Y;Z_\gamma)$ which originates from the Shannon lower bound for the rate-distortion function. Since the Gaussian distribution maximizes the differential entropy [21, Theorem 8.6.5], we have $h(Y|Z=z) \leq \frac{1}{2}\log(2\pi e \mathsf{var}(Y|Z=z))$ for any random variable Z. It immediately follows from Jensen's inequality that

$$h(Y|Z_\gamma) \leq \frac{1}{2}\log(2\pi e \mathsf{mmse}(Y|Z_\gamma)),$$

and hence

$$\mathsf{mmse}(Y|Z_\gamma) \geq \frac{1}{2\pi e} 2^{2h(Y|Z_\gamma)} = \mathsf{var}(Y) 2^{2(h(Y)-h(Y_\mathsf{G}))} 2^{-2I(Y;Z_\gamma)}, \tag{31}$$

from which we obtain

$$\inf_{\substack{\gamma \geq 0, \\ I(X;Z_\gamma)\leq \varepsilon}} \frac{\mathsf{mmse}(Y|Z_\gamma)}{\mathsf{var}(Y)} \geq 2^{-D(Y)} 2^{-2g_\varepsilon(X,Y)}, \tag{32}$$

where $D(Y)$ is the non-Gaussianness of Y defined in Proposition 1. We note that a similar inequality is proved in [2, Lemma 13] for arbitrary noise distribution provided that Y is Gaussian. Although, inequality (32) provides an operational interpretation of $g_\varepsilon(X,Y)$, it does not relate $g_\varepsilon(X,Y)$ to $\mathsf{sENSR}_\varepsilon(X,Y)$. Such a relationship would follow if $\rho_m^2(X,Z_\gamma) \leq \varepsilon$ implied $I(X;Z_\gamma) \leq \varepsilon$ for a given (X,Y), because then according to (32), one could conclude that $\mathsf{sENSR}_\varepsilon \geq 2^{-D(Y)}2^{-2g_\varepsilon(X,Y)}$. However, this implication does not hold in general. Nevertheless, we show in the sequel that this implication holds for Gaussian X in the almost perfect privacy regime when $P_{Y|X}$ is an additive noise channel. First we notice that for jointly Gaussian X_G and Y_G, we have $I(X_\mathsf{G}; \sqrt{\gamma}Y_\mathsf{G} + N_\mathsf{G}) = -\frac{1}{2}\log(1-\rho^2(X_\mathsf{G}, \sqrt{\gamma}Y_\mathsf{G}+N_\mathsf{G}))$. Hence, since $\rho_m^2(X_\mathsf{G}, \sqrt{\gamma}Y_\mathsf{G}+N_\mathsf{G}) = \rho^2(X_\mathsf{G}, \sqrt{\gamma}Y_\mathsf{G}+N_\mathsf{G})$, the above implication clearly holds, i.e., $\rho_m^2(X_\mathsf{G}, \sqrt{\gamma}Y_\mathsf{G}+N_\mathsf{G}) \leq \varepsilon$ implies $I(X_\mathsf{G}; \sqrt{\gamma}Y_\mathsf{G}+N_\mathsf{G}) \leq \varepsilon$. On the other hand, specializing the decomposition (37) proved in Appendix A for $U = X_\mathsf{G}$ and $V = Z_\gamma$, we can write

$$I(X_\mathsf{G};Z_\gamma) = I(X_\mathsf{G}; \sqrt{\gamma}Y_\mathsf{G}+N_\mathsf{G}) + D(Z_\gamma|X_\mathsf{G}) - D(Z_\gamma), \tag{33}$$

where $D(V|U)$ for a pair of absolutely continuous random variables (U,V) is defined as

$$D(V|U) := D(P_{V|U}||P_{V_\mathsf{G}|U_\mathsf{G}}|P_U) = \mathbb{E}_{UV}\left[\log \frac{P_{V|U}}{P_{V_\mathsf{G}|U_\mathsf{G}}}\right], \tag{34}$$

where $(U_\mathsf{G}, V_\mathsf{G})$ is a pair of Gaussian random variables having the same means, variances and correlation coefficient as (U,V), and $P_{V_\mathsf{G}|U_\mathsf{G}}(\cdot|u)$ and $P_{V|U}(\cdot|u)$ are the conditional densities of V_G and V given $U_\mathsf{G} = u$ and $U = u$, respectively. As shown in [16, Appendix II] if $\mathsf{var}(Y) < \infty$, then as $\gamma \to 0$

$$D(Z_\gamma) = o(\gamma). \tag{35}$$

Lemma 1 in Appendix B shows that $D(Z_\gamma|X_\mathsf{G})$ also behaves like $o(\gamma)$ if $\mathsf{mmse}(Y|X_\mathsf{G}) = \mathsf{mmse}(Y_\mathsf{G}|X_\mathsf{G})$. In light of this lemma, (33), and (35), we can conclude that

$$I(X_\mathsf{G}; Z_\gamma) \leq I(X_\mathsf{G}; \sqrt{\gamma}Y_\mathsf{G} + N_\mathsf{G}) + \frac{\gamma}{2}\left[\mathsf{mmse}(Y_\mathsf{G}|X_\mathsf{G}) - \mathsf{mmse}(Y|X_\mathsf{G})\right] + o(\gamma).$$

Thus if P_{XY} satisfies $\mathsf{mmse}(Y|X_\mathsf{G}) = \mathsf{mmse}(Y_\mathsf{G}|X_\mathsf{G})$, or equivalently $\mathbb{E}[\mathsf{var}(Y|X_\mathsf{G})] = 1 - \rho^2(X,Y)$, we have

$$I(X_\mathsf{G}; Z_\gamma) \leq I(X_\mathsf{G}; \sqrt{\gamma}Y_\mathsf{G} + N_\mathsf{G}) + o(\gamma). \tag{36}$$

Since $\rho_m^2(X_\mathsf{G}, Z_\gamma) \geq \rho_m^2(X_\mathsf{G}, \sqrt{\gamma}Y_\mathsf{G} + N_\mathsf{G})$, we can conclude from (36) that, $\rho_m^2(X_\mathsf{G}, Z_\gamma) \leq \varepsilon$ implies $I(X_\mathsf{G}; Z_\gamma) \leq \varepsilon + o(\gamma)$ for sufficiently small γ (or equivalently ε). Note that it is straightforward to show that $\rho_m^2(X_\mathsf{G}, Z_\gamma) \leq \varepsilon$ implies $\gamma \leq \frac{\varepsilon}{\rho^2(X_\mathsf{G},Y)-\varepsilon}$ (see Example 2). Hence, in the almost perfect privacy regime, $\rho_m^2(X_\mathsf{G}, Z_\gamma) \leq \varepsilon$ is satisfied with γ which is at most linear in ε. Therefore, (36) allows us to conclude that $\rho_m^2(X_\mathsf{G}, Z_\gamma) \leq \varepsilon$ implies that $I(X_\mathsf{G}; Z_\gamma) \leq \varepsilon + o(\varepsilon)$.

The condition $\mathbb{E}[\mathsf{var}(Y|X_\mathsf{G})] = 1 - \rho^2(X,Y)$ is satisfied if the channel from X_G to Y is additive, that is, $Y = aX_\mathsf{G} + M$, where $a \in \mathbb{R}^+$ and M is a noise random variable with a density having variance $1 - \rho^2(X_\mathsf{G}, Y)$. However, since $\mathbb{E}[\mathsf{var}(Y|X_\mathsf{G})] = \mathbb{E}[\mathsf{var}(Y|rX_\mathsf{G})]$ for any $r \neq 0$, the variance condition can be removed. □

The lower-bound (30) can be further simplified by invoking Corollary 2, which results in

$$\mathsf{sENSR}_\varepsilon(X_\mathsf{G}, Y) \geq 2^{-D(Y)}\left(1 - \frac{2\varepsilon}{\eta_{X_\mathsf{G}}^2(Y)}\right) + o(\varepsilon).$$

One the other hand, as proved in [4], when Y is Gaussian, Y_G, then

$$1 - \frac{\varepsilon}{\rho^2(X, Y_\mathsf{G})} \leq \mathsf{sENSR}_\varepsilon(X, Y_\mathsf{G}) \leq 1 - \frac{\varepsilon}{\rho_m^2(X, Y_\mathsf{G})},$$

for any $\varepsilon \leq \rho_m^2(X,Y)$. We have therefore tight lower bounds for $\mathsf{sENSR}_\varepsilon(X,Y)$ when either X or Y is Gaussian.

4 Conclusion

In this paper, we studied the problem of approximating the maximal amount of information one can transmit about a random variable Y over an additive Gaussian channel without revealing more than a certain (small) amount of information about another random variable X that represents sensitive or private data. Specifically, letting $g_\varepsilon(X,Y)$ denote the maximum of $I(Y;Z_\gamma)$ over $\gamma \geq 0$, where $Z_\gamma := \sqrt{\gamma}Y + N_\mathsf{G}$ and $N_\mathsf{G} \sim \mathcal{N}(0,1)$ is independent of (X,Y), subject to $I(X;Z_\gamma) \leq \varepsilon$, we showed that $g_\varepsilon(X,Y) = \frac{\varepsilon}{\eta_X^2(Y)} + \Delta(X,Y)\varepsilon^2 + o(\varepsilon)$ where $\eta_X^2(Y)$ and $\Delta(X,Y)$ are two asymmetric measures of correlation between X and Y. For the special case of jointly Gaussian X and Y, the approximation was

compared with the exact value of $g_\varepsilon(X,Y)$. As a side result, we also showed that this approximation leads to a slightly improved version of the strong data processing inequality under some suitable conditions on $P_{Y|X}$.

We also studied an estimation-theoretic formulation of the privacy-utility tradeoff for the same setup. Let $\mathsf{sENSR}_\varepsilon(X,Y)$ be the smallest achievable MMSE in estimating Y given Z_γ such that MMSE in estimating any function f of X given Z_γ is lower bound by $(1-\varepsilon)\mathsf{var}(f(X))$. We then showed that when X is Gaussian and Y is the output of an additive noise channel then $\mathsf{sENSR}_\varepsilon(X,Y) \geq 2^{-D(Y)}2^{-2g_\varepsilon(X,Y)}$ for sufficiently small ε, where $D(Y)$ is the non-Gaussianness of Y. The significance of this bound is that it gives an operational interpretation for $g_\varepsilon(X,Y)$ in terms of MMSE. Using the approximation obtained for $g_\varepsilon(X,Y)$, we derived a lower bound for $\mathsf{sENSR}_\varepsilon(X,Y)$ for small ε which is linear in ε.

References

1. S. Asoodeh, F. Alajaji, and T. Linder, "Notes on information-theoretic privacy," in *Proc. 52nd Annual Allerton Conference on Communication, Control, and Computing*, Sept. 2014, pp. 1272–1278.
2. S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Information extraction under privacy constraints," *Information*, vol. 7, no. 1, 2016. [Online]. Available: http://www.mdpi.com/2078-2489/7/1/15
3. F. P. Calmon, M. Varia, M. Médard, M. M. Christiansen, K. R. Duffy, and S. Tessaro, "Bounds on inference," in *Proc. 51st Annual Allerton Conference on Communication, Control, and Computing*, Oct 2013, pp. 567–574.
4. S. Asoodeh, F. Alajaji, and T. Linder, "Privacy-aware MMSE estimation," in *Proc. IEEE Int. Symp. Inf. (ISIT)*, July 2016. [Online]. Available: arXiv:1511.02381v3
5. A. Makhdoumi and N. Fawaz, "Privacy-utility tradeoff under statistical uncertainty," in *Proc. 51st Allerton Conference on Communication, Control, and Computing*, Oct 2013, pp. 1627–1634.
6. F. P. Calmon, A. Makhdoumi, and M. Médard, "Fundamental limits of perfect privacy," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2015, pp. 1796–1800.
7. T. Berger and R. Yeung, "Multiterminal source encoding with encoder breakdown," *IEEE Trans. Inf. Theory*, vol. 35, no. 2, pp. 237–244, March 1989.
8. I. S. Reed, "Information theory and privacy in data banks," in *Proc. of the National Computer Conference and Exposition, ser. AFIPS'73.* New York, NY, USA: ACM, 1973, pp. 581–587.
9. H. Yamamoto, "A source coding problem for sources with additional outputs to keep secret from the receiver or wiretappers," *IEEE Trans. Inf. Theory*, vol. 29, no. 6, pp. 918–923, Nov. 1983.
10. L. Sankar, S. Rajagopalan, and H. Poor, "Utility-privacy tradeoffs in databases: An information-theoretic approach," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 6, pp. 838–852, 2013.
11. S. Asoodeh, F. Alajaji, and T. Linder, "On maximal correlation, mutual information and data privacy," in *Proc. IEEE 14th Canadian Workshop on Inf. Theory (CWIT)*, June 2015, pp. 27–31.
12. D. Rebollo-Monedero, J. Forne, and J. Domingo-Ferrer, "From t-closeness-like privacy to postrandomization via information theory," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 11, pp. 1623–1636, Nov 2010.

13. F. P. Calmon, "Information-theoretic metrics for security and privacy," Ph.D. dissertation, MIT, Sep. 2015.
14. S. Goldwasser and S. Micali, "Probabilistic encryption," *Journal of Computer and System Sciences*, vol. 28, no. 2, pp. 270 – 299, 1984.
15. V. Anantharam, A. Gohari, S. Kamath, and C. Nair, "On maximal correlation, hypercontractivity, and the data processing inequality studied by Erkip and Cover," *Preprint, arXiv:1304.6133v1*, 2014.
16. D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inf. Theory,*, vol. 51, no. 4, pp. 1261–1282, April 2005.
17. D. Guo, Y. Wu, S. Shamai, and S. Verdú, "Estimation in Gaussian noise: properties of the minimum mean-square error," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2371–2385, April 2011.
18. A. Rényi, "On measures of dependence," *Acta Mathematica Academiae Scientiarum Hungarica*, vol. 10, no. 3, pp. 441–451, 1959.
19. H. Gebelein, "Das statistische problem der korrelation als variations- und eigenwertproblem und sein zusammenhang mit der ausgleichungsrechnung," *Zeitschrift f ur angew. Math. und Mech.*, no. 21, pp. 364–379, 1941.
20. O. Sarmanov, "The maximum correlation coefficient (nonsymmetric case)," *Dokl. Akad. Nauk SSSR*, vol. 120, no. 4, pp. 715–718, 1958.
21. T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.
22. Y. Polyanskiy and Y. Wu, "Dissipation of information in channels with input constraints," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 35–55, Jan 2016.
23. Y. Wu and S. Verdú, "Functional properties of minimum mean-square error and mutual information," *IEEE Trans. Inf. Theory,*, vol. 58, no. 3, pp. 1289–1301, March 2012.
24. V. V. Prelov, "Capacity of communication channels with almost Gaussian noise," *Teor. Veroyatnost. i Primenen.*, vol. 33, no. 3, pp. 433–452, 1988.

A Connection Between Mutual Information and Non-Gaussianness

For any pair of random variables (U, V) with $I(U;V) < \infty$, let $P_{V|U}(\cdot|u)$ be the conditional density of V given $U = u$. Then, we have

$$\begin{aligned} I(U;V) &= \mathbb{E}_{UV}\left[\log \frac{P_{V|U}(V|U)}{P_V(V)}\right] \\ &= \mathbb{E}_{UV}\left[\log \frac{P_{V|U}(V|U)}{P_{V_{\mathsf{G}}|U_{\mathsf{G}}}(V|U)}\right] + \mathbb{E}_{UV}\left[\log \frac{P_{V_{\mathsf{G}}|U_{\mathsf{G}}}(V|U)}{P_{V_{\mathsf{G}}}(V)}\right] - \mathbb{E}_{UV}\left[\log \frac{P_V(V)}{P_{V_{\mathsf{G}}}(V)}\right] \\ &= I(U_{\mathsf{G}};V_{\mathsf{G}}) + D(V|U) - D(V), \end{aligned} \tag{37}$$

where $(U_{\mathsf{G}}, V_{\mathsf{G}})$ is a pair of Gaussian random variable having the same means, variances and correlation coefficient as (U, V), and $P_{V_{\mathsf{G}}|U_{\mathsf{G}}}(\cdot|u)$ is the conditional density of V_{G} given $U_{\mathsf{G}} = u$, and the quantity $D(V|U)$ is defined in (34). Replacing U and V with X and Z_γ, respectively, the decomposition (37) allows us to conclude that

$$I(X;Z_\gamma) = I(X_{\mathsf{G}}; \sqrt{\gamma}Y_{\mathsf{G}} + N_{\mathsf{G}}) + D(Z_\gamma|X) - D(Z_\gamma),$$

and therefore, if $Y = Y_{\mathsf{G}}$ is Gaussian, we have

$$I(X;Z_\gamma) = I(X_{\mathsf{G}};Z_\gamma) + D(Z_\gamma|X) \geq I(X_{\mathsf{G}};Z_\gamma),$$

from which we conclude that when Y is Gaussian then $I(X;Z_\gamma) \leq \varepsilon$ implies that $I(X_{\mathsf{G}};Z_\gamma) \leq \varepsilon$ and hence $g_\varepsilon(X,Y_{\mathsf{G}}) \leq g_\varepsilon(X_{\mathsf{G}},Y_{\mathsf{G}})$.

B Completion of the Proof of Theorem 4

Lemma 1. *For Gaussian X_{G} and absolutely continuous Y with unit variance, we have*

$$D(Z_\gamma|X_{\mathsf{G}}) \leq \frac{\gamma}{2}\left[\mathsf{mmse}(Y_{\mathsf{G}}|X_{\mathsf{G}}) - \mathsf{mmse}(Y|X_{\mathsf{G}})\right] + o(\gamma).$$

Proof. Let E be an auxiliary random variable defined as

$$E = \begin{cases} 1, & |Y| \leq L \\ 0, & \text{otherwise,} \end{cases}$$

for some real number $M > 0$. Note that

$$\begin{aligned} D(Z_\gamma|X_{\mathsf{G}}=x) &= h(\sqrt{\gamma}Y_{\mathsf{G}} + N_{\mathsf{G}}|X_{\mathsf{G}}=x) - h(Z_\gamma|X_{\mathsf{G}}=x) \\ &\leq h(\sqrt{\gamma}Y_{\mathsf{G}} + N_{\mathsf{G}}|X_{\mathsf{G}}=x) - h(Z_\gamma|X_{\mathsf{G}}=x,E) \\ &= \frac{1}{2}\log(2\pi e(1+\gamma\mathsf{var}(Y_{\mathsf{G}}|X_{\mathsf{G}}=x))) \\ &\quad - \Pr(E=1)h(Z_\gamma|X_{\mathsf{G}}=x,E=1) - \Pr(E=0)h(Z_\gamma|X_{\mathsf{G}}=x,E=0) \\ &\overset{(a)}{\leq} \frac{1}{2}\log(2\pi e(1+\gamma\mathsf{var}(Y_{\mathsf{G}}|X_{\mathsf{G}}=x)) - \Pr(E=0)h(N_{\mathsf{G}}) \\ &\quad - \Pr(E=1)h(Z_\gamma|X_{\mathsf{G}}=x,E=1) \end{aligned} \tag{38}$$

where (a) follows from the fact that $h(Z_\gamma|X_{\mathsf{G}}=x,E=0) \geq h(N_{\mathsf{G}})$.

Prelov [24] showed that for any random variable Y such that

$$\mathbb{E}[|Y|^{2+\alpha}] \leq K < \infty, \tag{39}$$

for some $\alpha > 0$, then

$$h(\sqrt{\gamma}Y + N_{\mathsf{G}}) = \frac{1}{2}\log(2\pi e) + \frac{\mathsf{var}(Y)}{2}(\gamma + o(\gamma)), \tag{40}$$

where $o(\gamma)$ term depends only on K. Since $Y|\{E=1\}$ satisfies (39), we can use (40) to evaluate $h(Z_\gamma|X_{\mathsf{G}}=x,E=1)$ in (38) which yields

$$\begin{aligned} D(Z_\gamma|X_{\mathsf{G}}=x) &\leq \frac{1}{2}\log(2\pi e(1+\gamma\mathsf{var}(Y_{\mathsf{G}}|X_{\mathsf{G}}=x)) - \Pr(E=0)\frac{1}{2}\log(2\pi e) \\ &\quad - \Pr(E=1)\left[\frac{1}{2}\log(2\pi e) + \frac{\mathsf{var}(Y|X_{\mathsf{G}}=x,E=1)}{2}(\gamma+o(\gamma))\right] \\ &= \frac{1}{2}\log(1+\gamma\mathsf{var}(Y_{\mathsf{G}}|X_{\mathsf{G}}=x)) \\ &\quad - \frac{\mathsf{var}(Y|X_{\mathsf{G}}=x,E=1)}{2}(\gamma+o(\gamma))\Pr(E=1). \end{aligned} \tag{41}$$

Note that since $\mathsf{var}(Y) < \infty$ and X_{G} has a positive density, $\mathsf{var}(Y|X_{\mathsf{G}} = x) < \infty$ for almost all x (except for x in a set of zero Lebesgue measure). Hence, we can choose L sufficiently large such that for any given $\delta > 0$,

$$\Pr(E = 1) \geq 1 - \delta,$$

and

$$\mathsf{var}(Y|X_{\mathsf{G}} = x, E = 1) \geq \mathsf{var}(Y|X_{\mathsf{G}} = x) - \delta.$$

Therefore, invoking the inequality $\log(1 + u) \leq u$ for $u > 0$, we can write

$$D(Z_\gamma|X_{\mathsf{G}} = x) \leq \frac{\gamma}{2} \left[\mathsf{var}(Y_{\mathsf{G}}|X_{\mathsf{G}} = x) - (\mathsf{var}(Y|X_{\mathsf{G}} = x) - \delta)(1 - \delta)\right] + o(\gamma),$$

from which and the fact the δ is arbitrarily small the result follows. □