Fidelity Loss in Distribution-Preserving Anonymization and Histogram Equalization

Lav R. Varshney* and Kush R. Varshney†

*Coordinated Science Laboratory, University of Illinois at Urbana-Champaign

†Mathematical Sciences and Analytics, IBM Thomas J. Watson Research Center

***Abstract*—In this paper, we show a formal equivalence between histogram equalization and distribution-preserving quantization. We use this equivalence to connect histogram equalization to quantization for preserving anonymity under the k-anonymity metric, while maintaining distributional properties for data analytics applications. Finally, we make connections to mismatched quantization. These relationships allow us to characterize the loss in mean-squared error (MSE) performance of privacy-preserving quantizers that must meet distribution-preservation constraints as compared to MSE-optimal quantizers in the high-rate regime. Thus, we obtain a formal characterization of the cost of anonymity.**

***Index Terms*—data, histogram equalization, k-anonymity, mismatched quantization, neural information processing**

I. Introduction

With the proliferation of information systems, there is growing societal interest in maintaining privacy, especially in settings such as health care [1] and education [2]. In discussions of privacy, a *transmission principle* is a constraint on the flow (distribution, dissemination, transmission) of information in a given context. The transmission principle parameter expresses terms and conditions under which such transfers ought (or ought not) to occur [3, p. 145]. As established in signal processing and information theory, there are basic trade-offs between privacy and utility inherent in transmission principles, trade-offs governed by their parameters [4]–[6].

Several formal transmission principles have been proposed [7]; we focus herein on k-anonymity, with parameter k. In releasing data tables about people under k-anonymity, the data for an individual should not be distinguishable from the data of at least $k-1$ other individuals [8], [9].

Given a data set, k-anonymity is often achieved via generalizations and suppressions, or through the addition of noise. Of particular note, k-anonymity can be achieved via quantization of data. Quantization has also been widely discussed in multimedia signal processing [10] and in theoretical explanations of neural information processing [11], [12], where physical constraints limit information rates.

In theoretical neuroscience, the principle of histogram equalization is invoked to explain the nature of neural signals that seem to maximize information carried, in the sense of maximum entropy. That is, by essentially appealing to the fact noiseless (maximum entropy) and symmetric (capacity-achieving input distribution) communication channels are best used with equiprobable inputs, this approach is meant to not waste channel capacity. Laughlin showed the response of the large monopolar cell (LMC) in the fly visual system uses this entropy maximizing scheme (the LMC responds to contrast, and the probability distribution of contrasts of natural scenes in habitats where flies live were measured) [11]. This approach to not waste channel capacity by ensuring that quantization indices can be used directly with standard channel coding schemes (which assume that input messages have equal likelihood) have also been suggested in engineering [13].

Indeed, in histogram equalization, the probability mass of the source distribution mapped to any representation index is equal to any other, yielding equiprobable (maximum entropy) codewords. Thus, as we show in the sequel, histogram equalization quantization yields a kind of k-anonymity: the source sample is divided into equal-sized groups whose number N can be chosen sufficiently large to meet any k-anonymity requirement with high probability.

Note that data-driven quantizer design algorithms such as the k-means algorithm[1] due to Lloyd and Max can be modified to take the transmission principle parameter k (number of data points in cluster) as input, rather than number of clusters N as input [1], [14], though there are numerical difficulties.

Upon reconstructing from representation indices back to the real line (and using subtractive dithering), histogram equalization also ensures that the reproduction distribution essentially matches the source distribution. This distribution-preservation property is important in many data analytics applications, such as covariate shift in regression, as we argued previously [1]. We had previously noted that distribution-preserving quantization has also been developed for perceptual reasons in audio [15], [16] and also studied using quantization theory [17].

The connections between quantization that achieves k-anonymity, distribution-preserving quantization, and histogram equalization quantization that are presented here, however, are new.

Histogram equalization achieves k-anonymity and preserves distributions, but such quantization does not optimize for fidelity criteria such as mean-squared error (MSE). For companding quantizers, the quantizer point density function $\lambda(u)$

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA.

[1]Note the "k" in k-means is distinct from the "k" in k-anonymity.

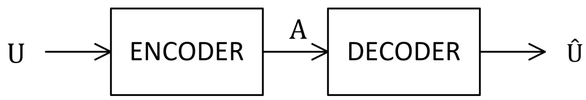

Fig. 1. Block diagram of a scalar quantizer, explicitly calling out the index variable A.

of histogram equalization is the probability density function of the source itself $f_U(u)$ rather than $\lambda(u) \propto f_U^{1/3}(u)$ as for MSE-optimality. A central contribution of this paper is to characterize loss in MSE performance due to this suboptimal choice for several typical source distributions such as Gaussian and Laplacian. Although this loss, which we refer to as the *cost of anonymity*, can often be computed precisely, we focus on asymptotic characterizations in the high-rate regime to get a sense of scaling [18]. To do so, we make connections to the theory of mismatched quantization and draw on some extant results from the theory [13], [19]–[21].

The remainder of the paper is organized as follows. Sec. II establishes notation by reviewing and reinterpreting results in quantization theory. The histogram equalization approach is also cast into the quantization theory framework. Sec. III shows that a histogram equalization compander achieves k-anonymity with high probability for sufficiently large datasets. By using results in mismatched quantization, Sec. IV computes the cost of anonymity in the high-rate regime. Sec. V concludes by discussing how other principles in theoretical neuroscience may be useful for privacy and anonymity.

II. Quantized Representations

To formalize our discussion, we review and reinterpret some basic definitions and results from quantization theory, following [22, Chapter 6.5.2] and [18].

Definition 1. *A* scalar quantizer q *is a mapping from* $\mathbb{R}$ *to a* reproduction codebook $\mathcal{C} = \{\hat{u}_i\}_{i\in\mathcal{A}}$*, where* $\mathcal{A}$ *is an arbitrary countable index set. Quantization can be decomposed into two operations,* ϕ_D *and* ϕ_E*, with* $q = \phi_D \circ \phi_E$*. The* lossy encoder $\phi_E : \mathbb{R} \mapsto \mathcal{A}$ *is specified by a partition* $\{S_i\}_{I\in\mathcal{A}}$ *of* $\mathbb{R}$ *with partition cells* $S_i = \phi_E^{-1} = \{u \in \mathbb{R} | \phi_E(u) = i\}, i \in \mathcal{A}$*. The* reproduction decoder $\phi_D : \mathcal{A} \mapsto \mathbb{R}$ *is specified by the codebook* $\mathcal{C}$*.*

The definition is depicted schematically in Fig. 1.

If each partition cell is an interval and the associated reproduction point lies within that interval, the quantizer is called *regular*. For settings we consider, there is no loss of optimality in restricting to regular quantizers [23, Sec. 6.2], which we do in the sequel.

A. Mean-Squared Error Optimality

One traditional measure of the quality of a quantizer is how well the reproduction $\hat{u}$ represents the original u, as measured using a distortion function $d(u, \hat{u})$: smaller average distortion means higher quality. A common distortion function is squared error, $d(u, \hat{u}) = |u - \hat{u}|^2$. In a data setting, the average distortion is the sample average when the quantizer is applied to a sequence of real data, but traditional theoretical development views the data as sharing a common cumulative distribution function (cdf) $F_U(u)$ and probability density function (pdf) $f_U(u)$ corresponding to a generic random variable U. The average distortion is

$$D = \mathbb{E}\left[d(U, \hat{U} = q(U))\right] = \sum_i \int_{S_i} d(u, \hat{u}_i) f(u) du.$$

When used for data compression, another traditional measure of the quality of a quantizer is the rate R. In the fixed-rate setting, $R_f = \log_2 |\mathcal{A}|$. Often there is thought to be an entropy code being applied after the quantizer, and so the entropy lower bound is used to measure rate as $R_e = H(A = \phi_E(U))$, where $H(A)$ is the entropy of the index variable A. When the index variable A is equiprobable, the entropy of the index is maximized and $R_f = R_e$.

Since precise analysis of optimal quantizers balancing rate and distortion may be difficult, Bennett's high-resolution approach with companding quantizers is often used [24]. In companding quantizers, a monotonic smooth nonlinearity G called a *compressor* is applied to the source, followed by a uniform quantizer having equally-spaced partition boundaries, and finally the inverse nonlinearity G^{-1} called the *expander* is applied when reconstructing the signal. That is, $\hat{u} = G^{-1}(q_\square(G(u)))$, where $q_\square(\cdot)$ is a uniform quantizer. Any nonuniform quantizer can be implemented as a compander.

For fixed-rate companding quantizers, the distortion is asymptotically (in the high-resolution regime) given by

$$D \cong \frac{\Delta^2}{12} \int \frac{f_U(u)}{g^2(u)} du, \tag{1}$$

where $g(u) = dG(u)/d(u)$, Δ is the width of the partition cells of the uniform quantizer $q_\square$, and the integral is over the granular range of the input. This is typically written in terms of the quantizer point density function $\lambda(u)$, whose integral over a region gives the fraction of quantizer reproduction levels in that region, as follows:

$$D \cong \frac{1}{12} \frac{1}{N^2} \int \frac{f_U(u)}{\lambda^2(u)} du, \tag{2}$$

where N satisfies $R_f = 2^N$. Using variational techniques or Hölder's inequality, Bennett's integral can be optimized to find the optimal $\lambda(u)$:

$$\lambda(u) = \frac{f_U^{1/3}(u)}{\int f_U^{1/3}(u) du}, \tag{3}$$

with corresponding mean-squared error

$$D \cong \frac{1}{12} \left(\int f_U^{1/3}(u) du \right)^3 2^{-2R_f} \tag{4}$$

This is called the Panter-Dite formula [25]. Panter and Dite's derivation demonstrated that an optimal quantizer has roughly equal contributions to total average distortion from each partition cell, a result that is sometimes called the *partial distortion theorem*.

B. Histogram Equalization

The Panter-Dite derivation suggests quantizers optimized for distortion should have each partition cell contribute roughly equally to average distortion, but the concept of histogram equalization that has been developed in both theoretical neuroscience and in multimedia signal processing requires partition cells to have probability mass that is roughly equal [11].

Let us restrict to companding quantizers and determine the point density function $\lambda(u)$. For histogram equalization, we require that the index variable be distributed equiprobably, $p_A(a_i) = 1/N$ for all $i \in \mathcal{A}$. That is to say,

$$\int_{S_i} f_U(u)du = 1/N \text{ for all } i \in \mathcal{A}. \tag{5}$$

Since $q_\Box$ is a uniform quantizer, in order to meet this condition in the high-resolution regime, we need the expanded random variable $G(U)$ to be a uniform random variable. As per the probability integral transformation, the choice $G(\cdot) = F_U(\cdot)$ yields the desired result, which further implies $\lambda(u) = f_U(u)$.

Note that since the expander reverses the compressor, i.e. $G^{-1} = F_U^{-1}$, the reproduction variable $\hat{U}$ has cdf F_U, just like U does. As a consequence, histogram equalization is distribution-preserving. This Rosenblatt transformation property of quantization is important for data analytics [1] and perceptual applications [16]. In finitary rate regimes, we have previously noted that such probability integral transformation scalings yield k-anonymity [1]. Now let us investigate k-anonymity in the high-rate regime.

III. Equivalence Theorem

By proving a concentration inequality in this section, we argue that a companding histogram equalization approach to quantization does achieve a kind of k-anonymity for specific finite-length realizations of data. This is especially important since companders are easy to implement but the quantization procedure of [14] is fraught with many numerical difficulties, only partially addressed by the numerical tricks described in the appendix therein.

Recall we are considering regular quantizers, and so partition cells are intervals. Let us denote these semi-closed intervals S_i by their boundary points, $S_i = (b_i, b_{i+1}]$. We want to show that the empirical measure of random data will give roughly the same number of points in each quantization bin as any other. This argument can be made with any concentration of measure result. The strong law of large numbers essentially shows there will be concentration, and the Glivenko-Cantelli theorem shows this concentration is uniform. We use the Dvoretzky-Kiefer-Wolfowitz inequality which quantifies the rate of convergence [26], [27].

Lemma 1 ([26], [27]). *Let $X_1, X_2, \ldots, X_n$ be real-valued independent and identically distributed random variables with cdf $F_X(\cdot)$. Let $F_X^{(n)}$ denote the associated empirical cdf defined by*

$$F_X^{(n)} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{X_i \le x}, \text{ for } x \in \mathbb{R},$$

where $\mathbb{1}$ denotes the indicator function. Then,

$$\Pr\left(\sqrt{n}\sup_{x\in\mathbb{R}} |F_X^{(n)}(x) - F_X(x)| > \lambda\right) \le 2e^{-2\lambda^2}, \tag{6}$$

for every $\lambda > 0$.

This leads to our concentration result for the k-anonymity of histogram equalization quantizers.

Theorem 1. *Let u_1^n be the actual data to be quantized, drawn i.i.d. from the common distribution $F_U(u)$. They are quantized using the companding quantizer with compressor $G = F_U$. Let the number of data points that lie in $S_i = (b_i, b_{i+1}]$ be k_i, for all $i \in \mathcal{A}$. Let $\Delta_\Box$ be a constant independent of i. Then*

$$\Pr[(k_i - \Delta_\Box) > \epsilon] \le 4e^{-\frac{n^2\epsilon^2}{2}},$$

for every $\epsilon > 0$.

Proof: Recall the probability that a random variable X lies in the semi-closed interval $(a, b]$, where $a < b$, is $\Pr[a < X \le b] = F_X(b) - F_X(a)$. Thus by the definition of the empirical measure, $k_i = n(F_U^{(n)}(b_{i+1}) - F_U^{(n)}(b_i))$. Let us subtract the quantity $n(F_U(b_{i+1}) - F_U(b_i))$ from both sides to get

$$\begin{aligned} &k_i - n(F_U(b_{i+1}) - F_U(b_i)) \\ &\quad = n(F_U^{(n)}(b_{i+1}) - F_U^{(n)}(b_i)) - n(F_U(b_{i+1}) - F_U(b_i)). \end{aligned} \tag{7}$$

From the construction of histogram equalization companding, we can note that this quantity that was subtracted, $n(F_U(b_{i+1}) - F_U(b_i))$, is a constant for all i, which we will call $\Delta_\Box$. Thus,

$$k_i - \Delta_\Box = n(F_U^{(n)}(b_{i+1}) - F_U^{(n)}(b_i)) - n(F_U(b_{i+1}) - F_U(b_i)), \tag{8}$$

which can be rearranged as

$$k_i - \Delta_\Box = n(F_U^{(n)}(b_{i+1}) - F_U(b_{i+1})) - n(F_U^{(n)}(b_i) - F_U(b_i)). \tag{9}$$

Now consider the quantity $\Pr[(k_i - \Delta_\Box) > \epsilon]$ for every $\epsilon > 0$. Letting $B_1 = (F_U^{(n)}(b_{i+1}) - F_U(b_{i+1}))$ and $B_2 = (F_U^{(n)}(b_i) - F_U(b_i))$, by the equality (9) above,

$$\begin{aligned} \Pr[(k_i - \Delta_\Box) > \epsilon] &= \Pr\left[n(B_1 - B_2) > \epsilon\right] && (10) \\ &\le \Pr[nB_1 > \tfrac{\epsilon}{2}] + \Pr[nB_2 < -\tfrac{\epsilon}{2}] && (11) \\ &\le \Pr[n|B_1| > \tfrac{\epsilon}{2}] + \Pr[n|B_2| > \tfrac{\epsilon}{2}] && (12) \\ &\le 2e^{-\frac{n^2\epsilon^2}{2}} + 2e^{-\frac{n^2\epsilon^2}{2}} \text{ for } \epsilon > 0 && (13) \\ &\le 4e^{-\frac{n^2\epsilon^2}{2}} \text{ for } \epsilon > 0, && (14) \end{aligned}$$

where the penultimate step (13) was two applications of Lemma 1, the Dvoretzky-Kiefer-Wolfowitz inequality. This is the desired result. ■

Note that this one-sided bound on the number of data points in a given quantizer representation interval can be directly extended to a two-sided bound.

Thus we see for sufficiently large numbers of data points, there is a concentration of measure such that an equal number of points lie in each cell, with exceedingly large probability. That is, k-anonymity is achieved by histogram equalization companding quantizer with exponentially high probability.

Note that this result does not require any high-rate asymptotics; it holds for any rate and number of partition cells N. The design of the companding quantizer is applicable at all rates, however, is motivated by high-rate considerations.

IV. Cost of Anonymity

Having established that quantizers that implement histogram-equalizing compander strategies achieve k-anonymity, we now concern ourselves with how much MSE distortion is incurred as compared to companding quantizers that are optimized for MSE. That is to say, we determine how much distortion is incurred by using $\lambda_{\mathrm{HE}}(u) = f_U(u)$ rather than $\lambda_{\mathrm{MSE}}(u) = f_U^{1/3}(u)$, which is optimal for MSE. We compute the average distortion as a function of N, the number of representation points in the quantizer, and term this performance as the *cost of anonymity*, C_{anon}. The analysis is in the asymptotic regime of high rate.

A. Uniform Sources

Consider a source of data, U, that is governed by a uniform distribution over the support interval $[c_1, c_2]$, i.e.

$$f_U(u) = \begin{cases} \frac{1}{c_2 - c_1}, & u \in [c_1, c_2] \\ 0, & \text{otherwise.} \end{cases}$$

Then, since $1^{1/3} = 1$, we see that $\lambda_{\mathrm{MSE}}(u) = \lambda_{\mathrm{HE}}(u) = f_U(u)$ and the two quantizers coincide. Since U is already uniform, there is no need for any companding. Thus we observe that for this source, there is no additional cost of anonymity, beyond an MSE-optimal quantizer. This is the only family of sources for which $f_U(u) = f_U^{1/3}(u)$ for all u and therefore the only class of sources that do not incur any additional cost of anonymity.

In fact there is no cost of anonymity for uniform sources in non-asymptotic regimes either. This follows directly from the facts that uniform quantizers are histogram equalizers for uniform sources and that uniform quantizers are MSE-optimal for uniform sources.

As is well-known, C_{anon} is $O(1/N^2)$ and more precisely

$$C_{\mathrm{anon}} = \frac{(c_2 - c_1)^2}{12} \frac{1}{N^2}. \tag{15}$$

B. Gaussian Sources

Although it is possible to apply Lloyd-Max type algorithms and other algorithms for finding k-anonymous quantizers for any source, as described in [1], [14], we are interested in analyzing performance in the high-rate regime, to get a closed-form asymptotic expression for C_{anon}.

Consider quantization of the zero-mean, unit-variance Gaussian source U,

$$f_U(u) = \frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{-u^2}{2} \right\}. \tag{16}$$

Then the $\lambda_{\mathrm{MSE}}(u)$ is given by:

$$\begin{aligned} \lambda_{\mathrm{MSE}}(u) &\propto \left(\frac{1}{\sqrt{2\pi}} \exp\left\{ \frac{-u^2}{2} \right\} \right)^{1/3} && (17) \\ &\propto \left(\frac{1}{\sqrt{2\pi}} \right)^{1/3} \left(\exp\left\{ \frac{-u^2}{2} \right\} \right)^{1/3} && (18) \\ &\propto \exp\left\{ \frac{-u^2}{6} \right\}. && (19) \end{aligned}$$

So we see that $\lambda_{\mathrm{MSE}}(u)$ corresponds to the pdf of a Gaussian random variable with variance 3 instead of variance 1.

We are facing a special case of Na's results on mismatched-variance quantizers [20], which has recently been studied further to determine a simplified asymptotic average distortion expression [13]. With a so-called mismatch factor of $\sqrt{3}$ in the heavy-mismatch regime, to use Na's terminology, the average distortion incurred by the histogram equalization approach, $C_{\mathrm{anon}}(N)$, is $O(1/N \ln N)$. Although the specific distortion function is fairly straightforward to write, we recall an easier-to-interpret asymptotic expression [13] that is more precise than Na's expression when particularized for scalar quantization. It is:

$$\lim_{N \to \infty} C_{\mathrm{anon}}(N) N \ln N = \frac{13}{12}. \tag{20}$$

C. Laplacian Sources

Now instead of a Gaussian source, let us consider finding the cost of anonymity for a Laplacian source. We will use the same basic technique for the Gaussian case, using Na's results for mismatched Laplacians [19], [21].

Let U be a mean-zero, unit-variance Laplacian random variable.

$$f_U(u) = \frac{1}{\sqrt{2}} \exp\left\{ -\sqrt{2}|u| \right\}.$$

Then the $\lambda_{\mathrm{MSE}}(u)$ is given by:

$$\begin{aligned} \lambda_{\mathrm{MSE}}(u) &\propto \left(\frac{1}{\sqrt{2}} \exp\left\{ -\sqrt{2}|u| \right\} \right)^{1/3} && (21) \\ &\propto \left(\frac{1}{\sqrt{2}} \right)^{1/3} \left(\exp\left\{ -\frac{\sqrt{2}|u|}{3} \right\} \right) && (22) \\ &\propto \exp\left\{ -\frac{\sqrt{2}|x|}{3} \right\} && (23) \end{aligned}$$

which is Laplacian with variance 9 or standard deviation 3. So the degree of mismatch in Na's terms is $\rho = 3$, in the heavy-mismatch regime. The main result is that N is $O(1/N)$. More precisely,

$$C_{\mathrm{anon}}(K = N/2) = \frac{c(3)}{K} - \frac{3}{8}\frac{1}{K^2} - \frac{1}{40}\frac{1}{K^4} + O\left(\frac{1}{K^2}\right), \tag{24}$$

where $c(3) = 6.8887$. Moreover, Bennett's integral gives a fairly good approximation:

$$C_{\mathrm{anon}}(N) \approx \frac{3}{4}\frac{1}{N} - \frac{3}{2}\frac{1}{N^2}. \tag{25}$$

V. Conclusion

This paper has made a novel link between privacy and theoretical neuroscience, establishing k-anonymity and distribution-preservation can be achieved simultaneously using a histogram-equalizing quantization approach. Moreover, that a compander-based approach to histogram equalization will achieve k-anonymity with exponentially large probability. Finally, the cost of anonymity has been defined as the average distortion performance of the histogram equalization quantizer and computed for several sources. For uniform sources the cost of anonymity is $O(1/N^2)$, for Gaussian sources the cost of anonymity is $O(1/N \ln N)$, and for Laplacian sources the cost of anonymity is $O(1/N)$. It remains to compute the cost of anonymity for other kinds of sources, e.g. in the generalized Gaussian family, or to determine which classes of sources suffer the greatest cost of anonymity (and are therefore hardest to release with both utility and privacy).

It is worth noting that the method of privacy preservation through histogram-equalizing quantization presented here retains more information about the data sample than simply the empirical cdf, as would happen in permutation-invariant data compression [28]. Some knowledge of the individual data points can be recovered.

The formulation and analysis herein suggests a change in the architecture for distribution-preserving anonymization as described in [1]: one should first perform companding or Rosenblatt's transformation to obtain a uniform distribution first, before quantization or clustering. This is because the equiprobable (k-member) constraint for clustering makes the optimization harder to carry out. Specifically, the procedure of [14] suffers from many numerical difficulties, that are not fully mitigated by the numerical tricks described therein. Compander-based quantization is simple to implement.

The entropy maximization principle present in histogram equalization led to further developments in neural computation, such as independent component analysis (ICA) which transforms multidimensional signals into components that are as statistically independent as possible [29]. It is of interest to explore whether ICA and similar techniques hold promise for anonymous data release in realistic multivariate settings. In fact, whitening and similar transformations have recently been proposed in the privacy literature [30], [31, Section 5.8].

Acknowledgment

Discussions with Bruno Olshausen, Ravi Kiran Raman, and Dennis Wei are appreciated.

References

[1] D. Wei, K. N. Ramamurthy, and K. R. Varshney, "Health insurance market risk assessment: Covariate shift and k-anonymity," in *Proc. 2015 SIAM Int. Conf. Data Mining*, May 2015, pp. 226–234.

[2] L. R. Varshney, A. Vempaty, and P. K. Varshney, "Assuring privacy and reliability in crowdsourcing with coding," in *Proc. 2014 Inf. Theory Appl. Workshop*, Feb. 2014.

[3] H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA: Stanford University Press, 2010.

[4] L. Sankar, S. R. Rajagopalan, and H. V. Poor, "Utility-privacy tradeoffs in databases: An information-theoretic approach," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 6, pp. 838–852, Jun. 2013.

[5] L. Sankar, W. Trappe, K. Ramchandran, H. V. Poor, and M. Debbah, "The role of signal processing in meeting privacy challenges: An overview," *IEEE Signal Process. Mag.*, vol. 30, no. 5, pp. 95–106, Sep. 2013.

[6] Q. Geng, P. Kairouz, S. Oh, and P. Viswanath, "The staircase mechanism in differential privacy," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 7, pp. 1176–1184, Oct. 2015.

[7] M. Bezzi, "An information theoretic approach for privacy metrics," *Transactions on Data Privacy*, vol. 3, no. 3, pp. 199–215, 2010.

[8] P. Samarati, "Protecting respondents' identities in microdata release," *IEEE Trans. Knowl. Data Eng.*, vol. 13, no. 6, pp. 1010–1027, Nov.-Dec. 2001.

[9] L. Sweeney, "k-anonymity: A model for protecting privacy," *Int. J. Uncertain. Fuzz.*, vol. 10, no. 5, pp. 557–570, Oct. 2002.

[10] W. Burger and M. J. Burge, *Digital Image Processing*. Springer, 2012.

[11] S. Laughlin, "A simple coding procedure enhances a neuron's information capacity," *Z. Naturforschung C*, vol. 36, no. 9–10, pp. 910–912, Sept.-Oct. 1981.

[12] J. Z. Sun, G. I. Wang, V. K. Goyal, and L. R. Varshney, "A framework for Bayesian optimality of psychophyscal laws," *J. Math. Psychol.*, vol. 56, no. 6, pp. 495–501, Dec. 2012.

[13] E. Köken, E. Tuncel, and D. Gündüz, "On the asymptotic distortion-energy tradeoff for zero-delay transmission of a Gaussian source over the AWGN channel," in *Proc. 2015 IEEE Int. Symp. Inf. Theory*, Jun. 2015, pp. 481–485.

[14] D. Rebollo-Monedero, J. Forné, E. Pallarès, and J. Parra-Arnau, "A modification of the Lloyd algorithm for k-anonymous quantization," *Inf. Sci.*, vol. 222, pp. 185–202, Feb. 2013.

[15] M. Li and W. B. Kleijn, "Quantization with constrained relative entropy and its application to audio coding," in *127th Audio Eng. Soc. Conv. Papers*, Oct. 2009, p. 7874.

[16] M. Li, J. Klejsa, and W. B. Kleijn, "Distribution preserving quantization with dithering and transformation," *IEEE Signal Process. Lett.*, vol. 17, no. 12, pp. 1014–1017, Dec. 2010.

[17] N. Saldi, T. Linder, and S. Yüksel, "Randomized quantization and source coding with constrained output distribution," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 91–106, Jan. 2015.

[18] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[19] S. Na, "Asymptotic formulas for mismatched fixed-rate minimum MSE Laplacian quantizers," *IEEE Signal Process. Lett.*, vol. 15, pp. 13–16, 2008.

[20] ——, "Asymptotic formulas for variance-mismatched fixed-rate scalar quantization of a Gaussian source," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2437–2441, May 2011.

[21] ——, "Variance-mismatched fixed-rate scalar quantization of Laplacian sources," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4561–4572, Jul. 2011.

[22] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of Signal Processing*. Cambridge: Cambridge University Press, 2014.

[23] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston: Kluwer Academic Publishers, 1992.

[24] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 446–472, Jul. 1948.

[25] P. F. Panter and W. Dite, "Quantization distortion in pulse-count modulation with nonuniform spacing of levels," *Proc. IRE*, vol. 39, no. 1, pp. 44–48, Jan. 1951.

[26] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, "Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator," *Ann. Math. Stat.*, vol. 27, no. 3, pp. 642–669, Sep. 1956.

[27] P. Massart, "The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality," *Ann. Probab.*, vol. 18, no. 3, pp. 1269–1283, Jul. 1990.

[28] L. R. Varshney and V. K. Goyal, "Toward a source coding theory for sets," in *Proc. IEEE Data Compression Conf. (DCC 2006)*, Mar. 2006, pp. 13–22.

[29] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, no. 6, pp. 1129–1159, Nov. 1995.

[30] J. Adebayo, L. Kagal, and A. Pentland, "The hidden cost of efficiency: Fairness and discrimination in predictive modeling," in *Proc. Data for Good Exchange (D4GX)*, Sep. 2015.

[31] F. du Pin Calmon, "Information-theoretic metrics for security and privacy," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, Sep. 2015.