Rate-Distortion in Molecular Signal Sensing with Ligand Receptors

Arash Einolghozati, Faramarz Fekri
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, GA 30332
Email:{einolghozati, fekri}@ece.gatech.edu

***Abstract*—Molecular communication between biological entities is a new paradigm in which biological nodes sense the environment, communicate and cooperate with each other. Ligand receptors are among the most common mechanisms used by biological entities such as bacteria to sense the molecular signals in their surroundings. In such a mechanism, molecules (i.e., ligands) bind to certain proteins (i.e., receptors) and activate a signaling cascade inside the cell. In this paper, we study the distortion in sensing and estimation of the concentration of molecular signals by ligand receptors in biological agents. The sensing distortion is caused by both the randomness in the ligand reception and the quantization of the final receiver output. The random measurement of the signal by the binding activity differentiates this case from classical quantization problems. We propose an optimal random quantization technique that minimizes the overall distortion described above. The performance of this optimal technique is compared with a uniform quantizer design and the regions where the optimal quantizer can offer a considerable advantage are identified. Furthermore, we analyze the effect of the number of the output levels (i.e., the output rate) on the overall distortion compared with the theoretical limit given by Shannon. Following this, the best practical number of levels beyond which no significant improvement can be made is presented.**

I. Introduction and Background

Molecular communication is a paradigm in which molecules are used in some forms to encode, transmit and decode information. Recent advances in synthetic biology have encouraged the applications of primitive bio agents (e.g., the engineered bacteria) as the basic components of communication networks [1], [2]. As such, primitive bio agents are designed to behave collaboratively for performing the tasks that would be impossible otherwise. This paradigm is inspired by existing molecular communication in nature in various forms. Quorum sensing, for instance, is used by bacteria to sense the concentration of specific types of molecules in order to estimate bacteria population density, communicate their estimate to each other, and synchronize their collective actions [3].

Among the most promising envisioned applications of molecular communication is the design of biological communication networks in which a node acts as a sensor who measures and reports the existence and/or the intensity of specific types of molecules in the environment [4]. Ligand reception is among the most common mechanism through which the biological entities sense the molecules at their vicinity. In this context, the receptors are trans-membrane proteins on a receiving cell that are bound by their corresponding ligand to form a complex molecule, and hence, stimulating the cell to activate specific genes. These genes, in turn, trigger further reactions inside the cell indicating the reception of the molecule. The simplest reaction between a ligand and its receptor can be written as a two-way process [5]:

$$\text{Ligand} + \text{Receptor} \leftrightarrow \text{Complex}$$

In this paper, we focus on the equilibrium behavior of the ligand reception in which the rate of removal of the ligand by the cell is equal to rate at which new ligands are supplied by diffusion [5].

There are two main sources that lead to distortion in sensing the concentration of molecules by a node via ligand receptors. The first is due to random discretization of the continuous concentration of molecules through the inherent discrete nature of ligand reception described above. As the concentration of molecules at the vicinity of receptors increases, the probability of ligand binding increases as well. Hence, the number of activated receptors is a random discrete indicator for inferring the concentration of signal molecules. As we will see, we model this source of distortion with a binomial random variable which represents the mapping from the continuous binding probability (or equivalently the corresponding concentration of molecules) onto a discrete level of activated receptors. The second source of the distortion is due to the limited number of output levels at the sensor node. As a primitive bio sensor, the node needs to map multiple levels of activated receptors to a single output representation value. We view this distortion as quantization of the discrete binomial random variable from the previous stage.

Optimal Quantization techniques have been discussed extensively in the classical literature. Max in [6] described the optimal choice of quantizing intervals and the reconstruction points that result in minimum average quantization error for an arbitrary signal distribution. Uniform quantization (i.e., having uniform quantization levels) is the most common and practical quantization method [7]. Authors in [8] showed that the uniform quantization is asymptotically optimal with the number of quantization levels. Universal quantizers have been introduced for the cases where the input distribution is

This material is based upon work supported by the National Science Foundation under Grant No. CNS-111094

unknown beforehand [7]. A universal quantizer in the minimax sense (i.e., minimizing the distortion for the distribution that yields the maximum error) is also introduced in [9].

The analysis of the ligand reception process in molecular communication has been discussed to some extent in the literature. The capacity of ligand reception has been obtained in the steady state [10] and in the transient mode [11]. In [12], a noise model was also proposed for the ligand-binding reception. The capacity of a molecular communication system with ligand receptors, under various assumptions, has been analyzed as well [13]–[15].

In this paper, we study the fundamental problem of imposed distortion in sensing a molecular signal through the ligand reception process described earlier. An optimal quantization technique is proposed that minimizes the distortion caused by both the reception and output reconstruction processes. We compare the performance of this optimal technique with an uniform quantizer and identify the regions where the optimal quantizer has superior performance. We also compare the distortion resulted by using different number of output levels with the theoretical limit given by Shannon and obtain the optimal number of outputs.

The rest of the paper is organized as follows. In Section. II, the ligand reception model that we use is introduced and the optimization problem formalized. In Section. III, we solve the aforementioned optimization problem and discuss the main results. Finally, Section. IV concludes the paper.

II. Optimal Random Quantization

We consider a molecular node consisting of N receptors acting as a sensor node. The node senses the presence of a type of molecules in the environment through the ligand receptors at its surface and maps the continuous concentration of signal molecules to one of the discrete output levels. There is an abundance of literature describing the chemical reactions concerning the the ligand receptor activation mechanism (e.g., [16]). The equation for a single receptor is described by

$$\frac{dp(t)}{dt} = k_+ A(t)(1 - p(t)) - k_- p(t), \tag{1}$$

where A(t) is the concentration of signal molecules at the vicinity of the receptor, p(t) is the binding probability, and k_+ and k_- are two constants corresponding to binding and unbinding of molecules. Here, we focus on the steady-state behavior of the node. Hence, the steady-state probability of receptor activation p^* is obtained as:

$$p^* = \frac{k_+ A^*}{k_+ A^* + k_-}, \tag{2}$$

where A^* is the steady-state concentration of molecules. In this paper, we consider a node that comprises N ligand receptors that act independently to measure the steady-state concentration of molecules. We wish to measure the distortion imposed on the system by both the ligand reception process as well as the output representation with finite number of levels and ignore the distortion due to the latter stages of the output production. Hence, the measured distortion would be a lower bound for any node who uses ligand reception to sense the signal molecules. We will also introduce an optimal quantization scheme that minimizes the aforementioned distortion.

Upon being stimulated by the concentration of molecules A^* in the steady state (or equivalently, the activation probability p^*), a random S number of the receptors will be activated where $S \sim \text{Binomial}(N, p^*)$. From now on, we drop the superscript $*$ when referring to the steady-state values. The quantized value of p (i.e., the corresponding value of A in (2)) is inferred by observing the quantized value of S as the final output. Hence, the distortion on p has two sources: 1. Random and discrete nature of S in measuring p, and 2. Quantization error in representing S.

We formalize the problem as follows: Given the random variable corresponding to the number of activated receptors $S \sim \text{Binomial}(N, p)$, where p is an instantiation of P with the density function $f_P(p)$, and k quantization levels at the output, the objective is to find an optimal (the notion of optimality will follow) mapping π from $N+1$ levels of S to k levels of the node output. In other words, we need to find the optimal mutually exclusive sets S_i, $i \in \{1, \ldots, k\}$ and $S_i \subset \{0, 1, \ldots, N\}$, where $\bigcup_{i=1}^{k} S_i = \{0, 1, \ldots, N\}$. This would correspond to finding the optimal quantization intervals in a classical quantization problem [7]. Further, we need to find the optimal reconstruction points p_i, $i \in \{1, \ldots, k\}$, corresponding to each S_i. note that each A_i can then be computed from (2). Here, the optimality is defined as minimizing the modified distortion rate [7] for a noisy source. In other words, we minimize the conditional expectation $\mathbf{E}[d(P, \hat{p})|S = s]$ where $d()$ is a distance measure, which is assumed to be the squared error in this paper.

Following the footsteps of [17], we solve the above optimization problem, which can be shown to be convex, by simultaneously solving the following two problems: 1. For a given set of reconstruction points $p_1, \ldots, p_k$, find the best quantization sets $S_1, \ldots, S_k$, and 2. For a given set of S_i's, find the best corresponding reconstruction points p_i's. If we order the reconstruction points p_i in increasing order, Fine has shown that optimal S_i sets can be described as follows [17]:

$$S_i = \{s : \frac{p_i + p_{i-1}}{2} \leq \mathbf{E}[P|S = s] \leq \frac{p_i + p_{i+1}}{2}\}, \tag{3}$$

where $i \in \{1, \ldots, k\}$ and we define p_0 and p_{k+1} to be negative and positive infinity, respectively. Note that (3) resembles the first optimality criterion in a classical quantization problem [6] which declares each interval end point must be halfway between the reconstruction points immediately before and after that.

In the second step of the algorithm, we need to find the optimal reconstruction points for each S_i set. In other words, we need to minimize $\mathbf{E}[(P - p_i)^2|S \in S_i]$. It can be easily shown that the conditional mean of P minimizes this error. Hence, we have:

$$p_i = \mathbf{E}[P|S \in S_i], \quad i \in \{1, \ldots, k\}. \tag{4}$$

Rewriting (4) for our problem using the Bayes formula results in:

$$p_i = \frac{\int_0^1 p\, \mathbf{P}(S \in S_i|p) f_P(p) dp}{\mathbf{P}(S \in S_i)} = \frac{\sum_{s \in S_i} \binom{N}{s} \int_0^1 p^{s+1}(1-p)^{N-s} f_P(p) dp}{\sum_{s \in S_i} \binom{N}{s} \int_0^1 p^{s}(1-p)^{N-s} f_P(p) dp}, \quad (5)$$

where we have exchanged the order of summation and integration at the end. In the next section, we solve (3) and (5) simultaneously and compare the resulting distortion with both a simple uniform quantizer and the rate-distortion function.

Note that we have focused on the distortion in estimating p, instead of concentration of molecules A. The main reason is the simplicity it offers for the presentation of equations but in fact, (2) can be viewed as the compressor in a common compander quantizing system where the signal is mapped to the unit interval [0,1] through a nonlinear function $G()$ and then is quantized. The output reconstruction $\hat{A}$ can then be obtained by $\hat{A} = G^{-1}(q(G(A)))$ where q is the quantization scheme used for the normalized signal P [18]. The distortion $D(q)$ on the original signal, A, can be obtained by the "Bennett's integral" [7] as follows:

$$D(q) \cong D_0 \int \frac{f_A(a)}{g^2(a)} da, \quad (6)$$

where D_0 is the distortion in detecting the normalized signal, $f_A(a)$ is the input probability distribution and $g(a) = \frac{dG}{da}$. The optimal compressor has been described in the quantization literature [7]. The performance analysis of $G()$ in (2) as a compressor is outside of the scope of the current study.

III. Main Results

We define the distortion $D(N,k)$ as the expected value of squared error over the input probability distribution. Hence, we have:

$$D(N,k) = \mathbf{E}_P[(P-\hat{p})^2] = \mathbf{E}_S[\mathbf{E}_{P|S}[(P-\hat{p})^2|S]] = \sum_{s=0}^{N} \int_0^1 (p - p_{\pi(s)})^2 \mathbf{P}(S=s|P=p) f_P(p) dp = \sum_{s=0}^{N} \binom{N}{s} \int_0^1 (p - p_{\pi(s)})^2 p^s (1-p)^{N-s} f_P(p) dp, \quad (7)$$

where $\pi(s)$ is the optimal mapping from [0,...,N] to [1,...,k] introduced in the previous section. Note that we have used the Law of total expectation and the Bayes rule to obtain (7).

In Fig. 1, we have shown the average distortion versus the number of quantization levels k for different values of N using the optimal technique described above. We have chosen $f_P(p)$ to be uniform in $[0,1]$ but the following observations can be easily extended for any arbitrary distribution. As we observe in the plot, the effect of using larger quantization levels in reducing the overall distortion diminishes after a certain point. On the other hand, for small values of k, the distortion is approximately the same regardless of N. In particular, at $k = 1$ which corresponds to representing the input distribution with its expected value, the distortion is equal to the distribution variance, here $\frac{1}{12}$. These behavior can be explained by the observation that we made about the sources of the overall distortion in the previous section. The second source which stems from the size of the quantization levels and is independent of N, is the dominant term for small values of k but its effect diminishes in the form of $\frac{1}{k^2}$. The first source of the overall distortion which is due to the randomness in the observation of S, is enduring and the same for all the values of k. Hence, it becomes approximately the sole source of distortion for high values of k.

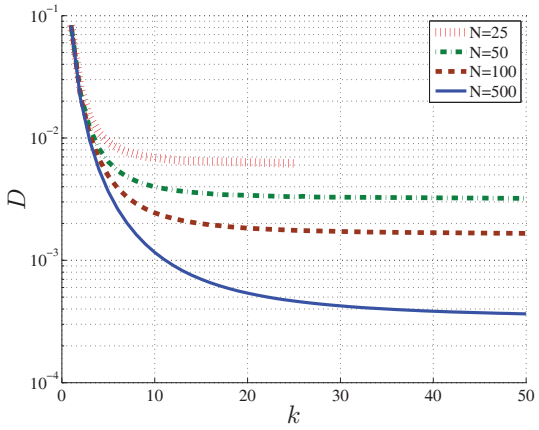

Fig. 1. Distortion versus number of quantization levels for different values of N

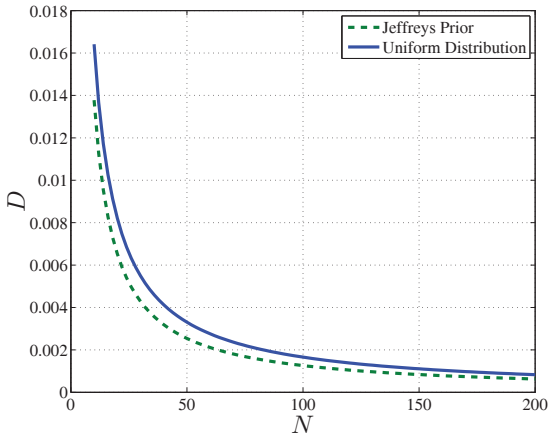

Fig. 2. Asymptotic distortion of optimal random quantizer versus N

Here, we take a closer look at the first source of the overall distortion due to randomness in the observation of S. Given N and observing the value of $S = s$, from classical statistics, $\frac{s}{N}$ is the sufficient statistics for estimating the input p. For large values of k, this statistic can be represented with a near perfect precision. Hence the average asymptotic distortion D_∞ in inferring the value of p becomes the expected value of the variance of $\frac{S}{N}$. In other words:

$$D_\infty \cong \int_0^1 \frac{p(1-p)}{N} f_P(p) dp, \quad (8)$$

where we have used the fact that $\mathbf{Var}[\frac{S}{N}] = \frac{p(1-p)}{N}$. Here, we consider two input distributions $f_P(p)$. First is the uniform distribution discussed above which results in $D_\infty \cong \frac{1}{6N}$. The second distribution is the Jeffreys prior for a Binomial observation given by:

$$f_P(p) = \frac{1}{\pi\sqrt{p(1-p)}}, \quad 0 < p < 1. \tag{9}$$

In [10], it has been shown that the Jeffreys prior is the capacity-achieving distribution for the receiver model in which a continuous concentration of molecules is mapped to a Binomial observation. Calculating (8) for the distribution in (9) results in the asymptotic distortion $D_\infty \cong \frac{1}{8N}$. Moreover, based on (8), the worst case asymptotic distortion is resulted by the probability distribution having all its probability mass at $p = \frac{1}{2}$ which would result in $D_\infty \cong \frac{1}{4N}$. In Fig. 2, we have shown the asymptotic distortion of the optimal random quantizer for both the uniform distribution and the Jeffreys prior. We can observe perfect match of the two curves with the asymptotic distortion $D_\infty \cong \frac{1}{6N}$ and $D_\infty \cong \frac{1}{8N}$, respectively, as obtained above.

In order to show the advantage of using the optimal random quantizer, we compare its distortion performance with the performance of a uniform random quantizer where the binomial observations are uniformly assigned to the quantization levels and also, the reconstruction levels are uniformly picked in the interval $[0, 1]$. We have plotted the normalized difference between the distortion of the optimal and uniform random quantizers versus the number of quantization levels k for the uniform distribution in Fig. 3. As we can see in Fig. 3, the two techniques result in approximately the same distortion for small values of k where the quantization error is dominant. The relative advantage of the optimal random quantizer becomes more apparent in this case for larger values of k but falls sharply for $k = N + 1$. This latter case corresponds to the point where a one-to-one mapping from the receptor outputs and the reconstruction levels becomes possible and the optimal quantizer loses its advantage in picking the optimal mapping. We can also observe from the plot that the overall advantage of using the optimal quantizer diminishes for large values of N and k. This reinforces the observation in [8] that uniform quantization is asymptotically optimal.

In Fig. 4, we have made the same comparison for the case of the capacity-achieving distribution in (9). As we can observe in the plot, except for $k = 1$ which corresponds to representing the distribution with its expected value, the optimal quantizer outperforms its uniform counterpart considerably for small values of k. Moreover, unlike the uniform distribution case, the optimal quantizer advantage is even more accentuated for larger values of N. This phenomenon can be explained by the observation that for small k and large N, the optimal distribution can fully take advantage of the asymmetric nature of the distribution in (9). Moreover, similar to the previous case, the uniform quantizer approaches its optimal counterpart for large values of N and k. Note that in both Fig. 3 and Fig. 4 , the curves have been smoothed to offset the illegibility due

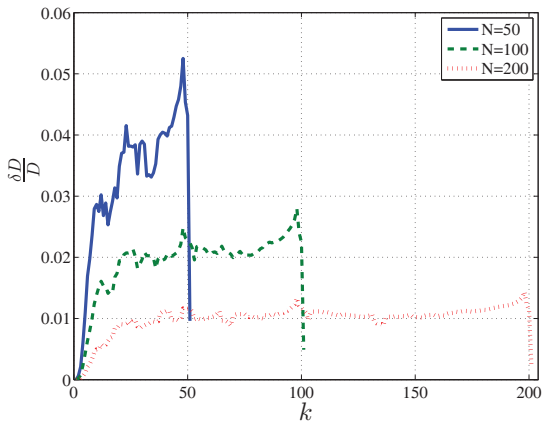

Fig. 3. Comparison of the optimal versus uniform random quantization for the uniform distribution

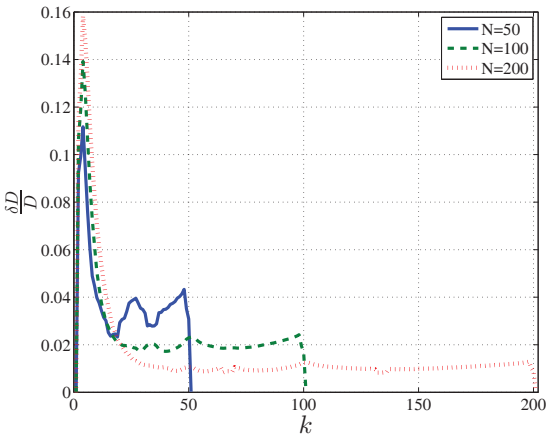

Fig. 4. Comparison of the optimal versus uniform random quantization for the Jeffreys prior

to their discrete nature.

We also compare the performance of the optimal random quantizer with the theoretical limit given by the distortion-rate function as defined by Shannon in [19]. Based on Shannon, $D(R)$ is defined as:

$$D(R) = \min_{Q(\hat{X}|X): I(Q(\hat{X}|X)) \leq R} d(Q(\hat{X}|X)), \tag{10}$$

where $Q(\hat{X}|X)$ is the conditional probability distribution over the reproduction alphabets given the source. Moreover, $d(Q(\hat{X}|X))$ and $I(Q(\hat{X}|X))$ are the average distortion and the average Shannon mutual information associated with $Q(\hat{X}|X)$, respectively, and are given by:

$$d(Q(\hat{X}|X)) = \sum_x \sum_{\hat{x}} p(x)Q(\hat{x}|x)d(x, \hat{x}),$$

where $p(x)$ is the input distribution and $d()$ is a distance measure (here the squared error) and

$$I(Q(\hat{X}|X)) = \sum_x \sum_{\hat{x}} p(x)Q(\hat{x}|x) \log \frac{Q(\hat{x}|x)}{\sum_x p(x)Q(\hat{x}|x)}$$

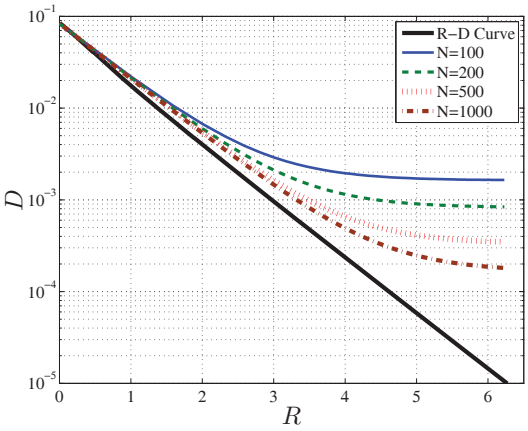

Fig. 5. Comparison of the optimal random quantizer with distortion rate curve

Note that in the ligand receiver model, $Q(\hat{x}|x)$ has been fixed to be a Binomial distribution. Based on the definition above, $D(R)$ gives the lower bound on the distortion for any quantizing methods with the rate limited to R. In order to solve the constrained optimization in (10), we use the Blahut algorithm described in [20] which obtains the points on the rate-distortion curve (or equivalently, distortion-rate curve) iteratively. In Fig. 5, we have shown the distortion of the optimal random quantizer versus $R = \log(k)$ for different values of N and compared it with the Shannon distortion-rate curve. As we can see in the plot, the optimal random quantizer curves follow the theoretical limit closely for small R but fall short afterwards and approach their asymptotic distortion. Also, it is possible to get closer to the distortion-rate curve by using larger values of N but the curves will eventually diverge where the random observation error dwarfs the quantization error. This will also correspond to the best practical choice for the number of levels $k^*(N)$, i.e., the minimum k that satisfies $\frac{k^2}{N} > a$ where the left side is the relative variance of the two errors and a is a constant. Hence, $R^*(N) = \log k^*(N) = \frac{\log N}{2} + \frac{\log a}{2}$. Choosing $a = 10$, we observe a good match between $R^*(N)$ and the points of divergence in Fig. 5.

IV. Conclusion

In this paper, we studied the distortion in molecular signal sensing via ligand receptors. We identified two sources of distortion, namely the random ligand reception and the receiver output quantization. We modeled the problem as finding the optimal quantization of a binomial random variable in order to infer its parameter with minimum distortion. We proposed an optimal random quantization technique that minimizes the overall distortion through optimal mapping of the binomial output and finding the best corresponding reconstruction levels. We showed that the first source of distortion decreases inversely with the number of receptors while the second decreases as inverse squared of the number of quantization levels. As such, increasing the number of quantization levels (i.e., increasing the receiver output rate) after a certain point has negligible effect on improving the overall distortion where the random observation of the input becomes the dominant source of the distortion. We compared the performance of the optimal random quantizer with a uniform quantizer and showed that even though the optimal technique offers a considerable advantage in certain regions, the uniform quantizer approachers the optimal one asymptotically. We also compared the performance of the optimal quantizer with the theoretical limit given by the distortion-rate curve proposed by Shannon. We observed that it is possible to approach the distortion-rate curve in low rates by using large number of receptors but the curves always finally diverge as the optimal quantizer's distortion approaches its asymptotic value which is independent of the rate.

References

[1] N. Nandagopal and M. B. Elowitz, "Synthetic biology: integrated gene circuits," *science*, vol. 333, no. 6047, pp. 1244–1248, 2011.

[2] A. Einolghozati, M. Sardari, A. Beirami, and F. Fekri, "Data gathering in networks of bacteria colonies: Collective sensing and relaying using molecular communication," in *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on*, 2012, pp. 256–261.

[3] B. L. Bassler, "How bacteria talk to each other: regulation of gene expression by quorum sensing." *Curr Opin Microbiol*, vol. 2, no. 6, pp. 582–587, Dec 1999.

[4] T. Nakano, M. Moore, F. Wei, A. Vasilakos, and J. Shuai, "Molecular communication and networking: Opportunities and challenges," *NanoBioscience, IEEE Transactions on*, vol. 11, no. 2, pp. 135 –148, June 2012.

[5] M. A. Model and G. M. Omann, "Ligand-receptor interaction rates in the presence of convective mass transport," *Biophysical Journal*, vol. 69, no. 5, pp. 1712–1720, 1995.

[6] J. Max, "Quantizing for minimum distortion," *Information Theory, IRE Transactions on*, vol. 6, no. 1, pp. 7–12, 1960.

[7] R. M. Gray and D. L. Neuhoff, "Quantization," *Information Theory, IEEE Transactions on*, vol. 44, no. 6, pp. 2325–2383, 1998.

[8] H. Gish and J. Pierce, "Asymptotically efficient quantizing," *Information Theory, IEEE Transactions on*, vol. 14, no. 5, pp. 676–683, 1968.

[9] J. Morris and V. Vandelinde, "Robust quantization of discrete-time signals with independent samples," *Communications, IEEE Transactions on*, vol. 22, no. 12, pp. 1897–1902, 1974.

[10] A. Einolghozati, M. Sardari, and F. Fekri, "Capacity of diffusion-based molecular communication with ligand receptors," in *Information Theory Workshop (ITW), 2011 IEEE*. IEEE, 2011, pp. 85–89.

[11] A. W. Eckford and P. J. Thomas, "Capacity of a simple intercellular signal transduction channel," *arXiv preprint arXiv:1305.2245*, 2013.

[12] M. Pierobon and I. F. Akyildiz, "Noise analysis in ligand-binding reception for molecular communication in nanonetworks," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4168–4182, 2011.

[13] Q. Liu and K. Yang, "Channel capacity analysis of a diffusion-based molecular communication system with ligand receptors," *International Journal of Communication Systems*, 2014.

[14] A. Noel, K. C. Cheung, and R. Schober, "Optimal receiver design for diffusive molecular communication with flow and additive noise," *arXiv preprint arXiv:1308.0109*, 2013.

[15] A. Einolghozati, M. Sardari, and F. Fekri, "Design and analysis of wireless communication systems using diffusion-based molecular communication among bacteria," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 12, pp. 6096–6105, 2013.

[16] D. Higham, "Modeling and simulating chemical reactions," *SIAM Review*, vol. 50, no. 2, pp. 347–368, 2008.

[17] T. Fine, "Optimum mean-square quantization of a noisy input (corresp.)," *Information Theory, IEEE Transactions on*, vol. 11, no. 2, pp. 293–294, 1965.

[18] W. R. Bennett, "Spectra of quantized signals," *Bell System Technical Journal*, vol. 27, no. 3, pp. 446–472, 1948.

[19] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat. Conv. Rec*, vol. 4, no. 142-163, p. 1, 1959.

[20] R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *Information Theory, IEEE Transactions on*, vol. 18, no. 4, pp. 460–473, 1972.