

# Media-based Modulation: Converting Static Rayleigh Fading to AWGN

Amir K. Khandani

E&CE Department, University of Waterloo, Waterloo, ON, Canada

**Abstract**—It is shown in [1] that embedding information in the (intentional) variation of the transmission media (end-to-end channel) can offer significant gains vs. traditional SISO, SIMO and MIMO systems. In particular, it is shown that using a single transmit antenna and  $K$  receive antennas; significant savings in energy vs. a  $K \times K$  MIMO can be achieved [1]. This article proves that, a  $1 \times K$  media-based modulation over a static multi-path channel asymptotically achieves the capacity of  $K$  parallel AWGN channels, where for each unit of energy over the single transmit antenna, the effective energy for each of the  $K$  AWGN channels is the statistical average of channel fading. The rate of convergence is computed. Significant gains can be realized even in a SISO media-based setup. An example for the practical construction of the system and its realistic RF simulation are presented. Issues of equalization and selection gain are discussed.

## I. INTRODUCTION

In [1], the author established the advantages of perturbing the end-to-end channel, according to the input data, in a wireless system with multi-path fading. These variations are detected at the receiver end, resulting in a modulation scheme with Additive White Gaussian Noise (AWGN). This method was coined in [1] as Media-Based Modulation (MBM). Assuming rich scattering, a small perturbation in the propagation environment will be augmented by reflections, resulting in an independent channel. As a result, in rich scattering, the MBM constellation size will be  $2^\kappa$  where  $\kappa$  is the number of channel perturbations.

In [3] [4], data is embedded in two orthogonal antenna beam-patterns, which transmit a binary signal set. Although use of orthogonal basis is common in communications, it usually does not bring any benefits on its own, it simplifies detection by keeping the noise uncorrelated. This means there are no clear advantages in designing the RF system to support orthogonal patterns as used in [3] [4]. The motivation in [3] [4] is to reduce the number of transmit chains and no other benefits are discussed. Bains [5] discusses using parasitic elements for data modulation, and shows limited energy saving, which again is indirectly due to the effect of classical RF beam-forming. Spatial Modulation (SM) [6] uses multiple transmit antennas with a single RF chain, where a single transmit antenna is selected according to the input data, with the rest of the data modulating the signal transmitted through the selected antenna. SM is in essence

a diagonal space-time code, where the trade-off between diversity and multiplexing gain has been in favour of the latter. A shortcoming of SM is that the rate due to the spatial portion increases with  $\log_2$  of the number of antennas, while in MBM, it increases exponentially. As an example, in a particular practical realization of MBM, a central antenna is surrounded by 14 RF mirrors (same height as the main antenna), resulting in  $2^{14}$  options, while for SM, this would be  $\log_2(14)$ . In SM, antennas should be sufficiently separated to have independent fading, while in MBM, the RF mirrors are placed side by side. In general, in MBM, the limit on the physical separation to have independent fading is non-existent. The switches used in SM are high power, which means expensive and slow, or each antenna needs a separate Power Amplifier (PA) with switches placed before PAs. The switches used for RF mirrors in MBM are cheap, low power and fast. The use of tuneable parasitic elements external to antenna(s) for beam-forming is an old topic. In terms of “varying the carrier after leaving the antenna”, this class of beam-forming schemes share similarities with MBM. However, the objective in parasitic antenna beam-forming is “to focus” the energy, which realizes energy gain, but does not realize the advantages of MBM, namely: (1) Additive information over multiple receive antennas, (2) Mixing good and bad channel conditions over a single constellation, which essentially removes the bottleneck of deep fades in static fading channels without compromising the rate.

Unveiling benefits of MBM (e.g., additivity of information over multiple receive antennas, inherent diversity over a signal constellation, exponential growth of data carrying options, etc.) and methods to realize them are the contributions of the author in [1] continued in the current article.

## II. SYSTEM MODEL

Figure 1 shows a  $1 \times K = Q/2$  SIMO-MBM. For simplicity, the concept is explained by focusing only on the MBM part of the transmission, and the combination with traditional Source-Based Modulation (SBM) is straightforward.

There are  $M$  messages,  $m = 1, \dots, M$ , selecting channel realizations  $\vec{h}(m)$  with components  $h_d(m)$ ,  $d = 1, \dots, Q$ , and  $m = 1, \dots, M$ .  $E|h_d(m)|^2 = 1$  where  $E$  denotes statistical averaging. AWGN  $\vec{z}$  has independent identically

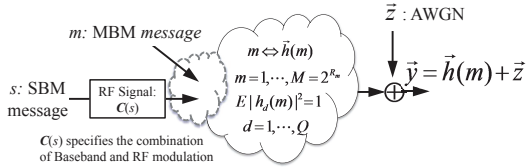


Fig. 1: System block diagram.

distributed (i.i.d) components  $z_d$ ,  $d = 1, \dots, K$ ,  $E|z_d|^2 = \sigma_z^2$ . Although transmitter selects the channel realization  $\vec{h}(m)$ , it is oblivious to its details. The receiver knows  $\vec{h}(m)$  and can perform reliable detection, if the mutual information is sufficient. Receiver training is achieved by sending a set of pilots for different channel realizations to measure  $h_d(m)$ ,  $\forall d, \forall m$ . For a Rayleigh fading channel,  $h_d(m)$  are i.i.d complex Gaussian. As the transmitter is oblivious to the details of  $\vec{h}(m)$ , an outage may occur. We have,

$$I(\vec{y}; m) \equiv I(\vec{y}; \vec{h}(m)) = h(\vec{y}) - h(\vec{z}) = h(\vec{y}) - K \log_2(\pi e \sigma_z^2).$$

For low Signal to Noise Ratio (*SNR*), we have (see [2]):

$$\lim_{\epsilon \rightarrow 0} \frac{I(\epsilon)}{\epsilon} \simeq \frac{Q}{2} (M_2 - M_1^2),$$

where  $I(\epsilon)$  is the mutual information per real dimension as a function of a small increase in energy,  $\epsilon$ , starting from zero energy,  $M_1$ ,  $M_2$  are the first and second sample moments of the  $M$ -points constellation, respectively. The SBM part of the modulation in its simplest form adds one bit of information by negating each vector of constellation points (180 degree phase shift applied to the carrier), resulting in  $M_1 = 0$ . In this case,

$$E \left[ \lim_{\epsilon \rightarrow 0} \frac{I(\epsilon)}{\epsilon} \right] = \frac{Q^2}{2}.$$

Although  $M_2$  is a random variable, its variance vanishes as  $1/M$ . However, in a  $K \times K$  MIMO-SBM the scaling of rate vs. *SNR* at low *SNR* is at best (i.e., assuming feedback and water filling) limited to the largest eigenvalue of a  $K \times K$  random Wishart matrix, the expected value of which is limited to 4 (approaches 4 as  $K \rightarrow \infty$ ) [7].

### III. CONVERGENCE OF RAYLEIGH CHANNEL TO AWGN

For a channel with a random static gain, MBM reduces the variance of the mutual information. It will be shown that this effect asymptotically converts a static Rayleigh fading channel into an AWGN channel where the *SNR* is determined by the statistical average of the channel gain. Assuming Rayleigh fading, the components of the constellation points are i.i.d. Gaussian. This is in agreement with random coding over an AWGN channel. However, the constellation structure in MBM remains the same over subsequent transmissions instead of having independent realizations as is required in random coding. This causes some loss in the achievable

rate as compared to the capacity of the underlying AWGN channel, i.e., AWGN with an energy gain equal to the statistical average of fading, which is normalized to one. It is difficult to compute the mutual information. For a given constellation, although the components are i.i.d. Gaussian, as the structure is static, the channel output will not be Gaussian. A method is presented to compute, on the average, the loss in the capacity vs. an AWGN channel with a Gaussian random code-book. Averaging is performed with respect to all possible random code-books, each corresponding to a realization of the constellation. It is shown this loss tends to zero as  $(1/M)^{1/K}$  for  $M \rightarrow \infty$ . Let us consider two ensembles of random codes:

*Ensemble I:* Given a realization of the  $M$ -points constellation as the underlying alphabet, construct an i.i.d. random code-book of cardinality  $c$ . Let us form a *Composite Ensemble I* by concatenating all such possible code-books.

*Ensemble II:* An ensemble of cardinality  $c$  with i.i.d Gaussian components of variance one. For Ensemble II, the composite version will be a Gaussian code-book with the same block length as *Composite Ensemble I*.

Note that both of the Composite Ensembles are Ergodic. The statistical average of rate over different constellations is equivalent to time average over Composite Ensemble I. We select  $c$  to guarantee reliable communication for Composite Ensemble I.

Capital vectors in Fig. 2(a) correspond to composite code-words, with the corresponding per code-book components shown in Fig. 2(b). Let us select a code-book from each of these two Composite Ensembles: For each code-word  $\vec{X}$  from Composite Ensemble I, find the code-word  $\vec{G}$  from the Composite Ensemble II that is at the minimum Euclidean distance to  $\vec{X}$ . Let us denote the corresponding vector of quantization error by  $\vec{N}_q$ , i.e.  $\vec{N}_q = \vec{G} - \vec{X}$ , and its per code-book components by  $\vec{n}_q$ , see Fig. 2. This setup allows to tackle the bottleneck in providing a closed form solution for the statistical average of mutual information, as  $\vec{G}$ , and consequently  $\vec{Y}$ , in Fig. 2 have i.i.d. Gaussian components. Two points should be mentioned: (1) Quantization noise,  $\vec{N}_q$ , depends on  $\vec{X}$ . (2) Mapping from  $\vec{X} \rightarrow \vec{G}$  is not necessarily one-to-one. As a result, elements of  $\vec{Y}$  may occur with non-equal probabilities. We have reliable transmission for Composite Ensemble I if the mutual information from  $\vec{X}$  to  $\vec{Y}$  is sufficient. Under these circumstances, as the receiver is able to reliably detect  $\vec{X}$ , it means that the probability that multiple  $\vec{X}$  are mapped to the same  $\vec{G}$  has asymptotically vanished. Note that if  $\vec{G}$ s occur with slightly (asymptotically vanishing) non-equal probabilities, this does not affect the hardening effect in the Gaussian channel and  $\vec{Y}$  will still have a uniform distribution over a sphere which is the basis for reliable transmission.

The following assumptions, which upper-bound the entropy of  $\vec{n}_q + \vec{z}$ , simplify the computations: (1) Dependency

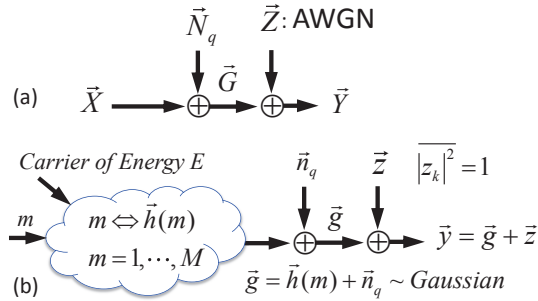


Fig. 2: (a) Block diagram for the composite Ensemble of code-books, and (b) their per code-book components.

among components of  $\vec{n}_q$  is ignored. (2)  $\vec{n}_q + \vec{z}$  is replaced by i.i.d Gaussian, with a variance conditioned on the constellation point. As  $\vec{G}$  and  $\vec{Y}$  will be i.i.d. Gaussian regardless of these assumptions, entropy of the channel output in Fig. 2(a) will not be affected. Overall, this results in a lower bound on the mutual information. Appendix A, through establishing an achievable rate for Composite Ensemble I, shows that the statistical average of mutual information in Fig. 2(b), and consequently  $\log_2 c$ , can approach the capacity of  $K$  parallel AWGN channels with  $K$  times energy saving with a gap behaving as:

$$\left(\frac{1}{M}\right)^{1/K} \rightarrow 0, \text{ as } M \rightarrow \infty. \quad (1)$$

*Interpretation of the mapping between two code-books to induce Gaussian distribution:* Let us assume transmitter picks a code-word, and a third party maps it to a code-word from the second code-book. Under proper conditions on mutual information and cardinality of the sets, this mapping can be based on minimum distance quantization with a vanishing probability of ties. Transmitter and receiver do not know the mapping rule. Such a mapping adds to the confusion, but if the mutual information is sufficient, its effect can be absorbed in the channel coding and the transmitted code-word can be recovered. Up to this point, computational complexity still remains intractable. However, it turns out that by a sequence of boundings, the computation reduces to finding the conditional variance of the quantization noise. This computation, in the asymptotic case of large constellation size, is tractable.

#### IV. ADDITIONAL REMARKS

**Equalization:** SBM exploits the LTI property and rely on equalization to combat Inter Symbol Interference (ISI). Due to ISI, the energy of a single symbol is spread over  $L$  dimensions, where  $L$  is the length of the channel impulse

response. In SBM, this  $L$ -D vector spans a single dimension. In contrast, in MBM, the  $L$ -D vector spans an  $L$  dimensions. As a result, a transmission in MBM can be followed by  $L$  zeroes to clear channel memory without sacrificing the dimensionality. An alternative would be to apply sequence detection in time. MBM equalization is simpler in multi-user setups, and in particular in networks using Orthogonal Frequency Division Multiple Access (OFDMA). In OFDMA, different nodes use different tones. In this case, the channel at each transmitter can be perturbed from OFDM symbol to OFDM symbol. As the channel remains static over each OFDM symbol, receivers can rely on a simple single tap equalizer. Such a setup can be used if transmitters are separated in space, e.g., in the uplink (transmitters use different tones to send to a common receiver), and in interfering links (each link uses an OFDM tone).

**Selection Gain:** As the constellation points have equal probability, selecting a subset can improve mutual information. The optimum solution will be too complex. The slope of the mutual information vs. energy at low  $SNR$  is determined by the sample second moment of the selected subset. This motivates selecting the highest energy points. Figure 3 shows an example for the achievable gain. As very low  $SNR$ , this slope scales with the maximum norm of the constellation points. Reference [2] contains mathematical derivations regarding this gain, which is similar to the scheduling gain:

$$E \left[ \max \left( \|\vec{h}(1)\|^2, \dots, \|\vec{h}(M)\|^2 \right) \right] \simeq A(Q) + \ln(M+1),$$

$$A(Q) = 2 \left[ (Q/2 - 1) \ln(Q/2 - 1) - \ln \Gamma(Q/2) \right].$$

Table I shows examples for values of  $A(Q)$ . We have  $A(Q) \simeq K = 2Q$  for large  $K$ .

$K = 2Q$	2	4	8	16	32	64
$A(Q)$	0	1.5	5	12.7	28.4	60

TABLE I: Example for the values of  $A(Q)$ .

**Pros and Cons:** *Pros:* Legacy schemes suffer from imperfections such as I/Q imbalance, PA non-linearity, PA efficiency, LNA non-linearity, etc. Most of these problems disappear in MBM. *Cons:* In MBM, one needs training with a complexity that grows with the constellation size. Two bits of information can be embedded in the carrier sign change and exchanging the role of I and Q, reducing the required number of training symbols. Gradual movement of constellation points over time can be partially tracked and readjusted before next training phase.

**Practical Implementation:** A novel RF mirror using a periodic switched structure [2] is used to perturb the channel. There are 14 mirrors enabling  $2^{14}$  combinations. The structure is very compact [2]. The system, including antenna and RF mirrors, is designed both as a on a PCB

as well as a three-dimensional structure. Figure 4 shows examples of MBM constellations. Figure 5 shows the corresponding propagation environments, for indoor (a typical house) and outdoor (a typical down-town). Simulations are performed using HFSS to extract antenna patterns, which are imported to Remcom Wireless Insite for ray tracing. Switches are simple PIN diodes, and  $S_{11}$  (antenna efficiency) is acceptable (below -10dB over a band of 100MHz around 2.4GHz).

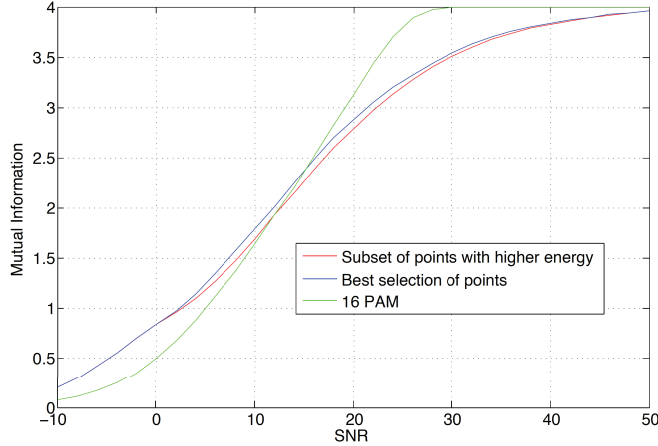


Fig. 3: Selection gain in a 256 points constellation.

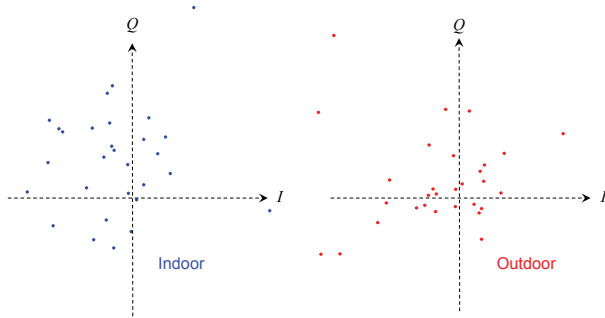


Fig. 4: Examples of constellation points corresponding to propagation environments in Fig. 5.

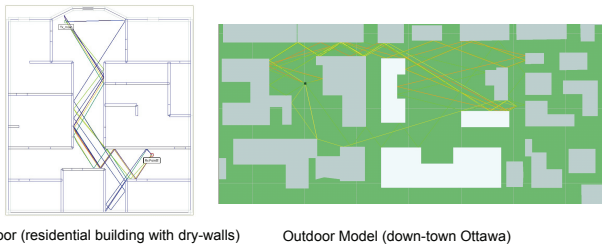


Fig. 5: Indoor/Outdoor Propagation environments corresponding to the MBM constellation points shown in Fig. 4

**Acknowledgment:** Author is thankful to E. Seifi for many useful inputs and computer simulations, to M. Salehi for RF related simulations, and to K. Moshksar for comments on this article.

## REFERENCES

- [1] A. K. Khandani, Media-based modulation: A new approach to wireless transmission, IEEE International Symposium on Information Theory (ISIT 2013), July 2013, pp. 3050—3054
- [2] A. K. Khandani, Media-based Modulation, Technical Report, Univ. Waterloo (cst.uwaterloo.ca/reports)
- [3] O. N. Alrabadi, A. Kalis, C. B. Papadias, R. Prasad, “Aerial modulation for high order PSK transmission schemes,” *1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, VITAE 2009*, pp. 823—826 .
- [4] O. N. Alrabadi, A. Kalis, C. B. Papadias, R. Prasad, “A universal encoding scheme for MIMO transmission using a single active element for PSK modulation schemes”, *IEEE Transactions on Wireless Communications*, Vol. 8 , No. 10, pp. 5133—5142.
- [5] R. Bains, “On the Usage of Parasitic Antenna Elements in Wireless Communication Systems”, PhD Thesis, Department of Electronics and Telecommunications, Norwegian University of Science and Technology, May 2008.
- [6] Raed Y. Mesleh, Harald Haas, Sinan Sinanovic, Chang Wook Ahn, and Sangboh Yun, , ”Spatial Modulation:, *Trans. on Vehicular Technology*, Vol. 57, No. 4, July 2008, pp. 2228—2241
- [7] Edelman, Rao. Random matrix theory. Acta # 14:233297, 2005.

## APPENDIX A

From Fig. 2(b), we have,

$$\begin{aligned} I(\vec{h}; \vec{y}) &= h(\vec{y}) - h(\vec{y} | \vec{h}) = h(\vec{y}) - h(\vec{n}_q + \vec{z} | \vec{h}) \\ &= K \log(\pi e \sigma_y^2) - h(\vec{n}_q + \vec{z} | \vec{h}), \end{aligned} \quad (2)$$

where  $\sigma_y^2$  is the variance of elements of  $\vec{y}$ . To obtain a lower bound, the entropy of  $\vec{n}_q + \vec{z}$  is upper-bounded by an i.i.d. Gaussian noise of the same conditional variance. This results in:

$$E_{\vec{h}} [I(\vec{h}; \vec{y}) | \vec{h}] \geq K \log(\pi e \sigma_y^2) - K E_{\vec{h}} [\log(\pi e \sigma_z^2 + \pi e \sigma_{n_q}^2) | \vec{h}] \quad (3)$$

where  $\sigma_y^2 = 1 + \sigma_z^2$  (signal power is normalized to one) and  $\sigma_{n_q}^2$  is the variance of elements of  $\vec{n}_q$ . Inequality (3) holds since entropy of  $\vec{n}_q + \vec{z}$  is bounded by the entropy of an i.i.d. Gaussian vector of the same variance. Due to concavity of log function, we have

$$E_{\vec{h}} [\log(\pi e \sigma_z^2 + \pi e \sigma_{n_q}^2) | \vec{h}] \leq \log(\pi e \sigma_z^2 + \pi e \sigma_{n_q}^2 | \vec{h}), \text{ where}$$

$$\sigma_{n_q}^2 | \vec{h} = \frac{1}{K} E [\|\vec{n}_q\|^2 | \vec{h}].$$

For given  $\vec{h}$ , let  $\vec{g}$  denote to the Gaussian vector at the minimum distance to  $\vec{h}$ , i.e.,  $\vec{h}$  is quantized to  $\vec{g}$ . Let us denote the  $Q \times 1$  vector of real and imaginary parts of all components of  $\vec{h}$  and  $\vec{g}$  by  $\vec{h}$  and  $\vec{g}$ , respectively.

$$\sigma_{n_q}^2 | \vec{h} = \frac{1}{Q} \int_{\vec{g} \in \mathbb{R}^Q} \|\vec{h} - \vec{g}\|^2 \frac{f_{\vec{g}, \vec{h}}(\vec{g}, \vec{h})}{f_{\vec{h}}(\vec{h})} d\vec{g}. \quad (4)$$

where  $f_{\vec{h}}(\vec{\cdot})$  and  $f_{\vec{g}, \vec{h}}(\vec{\cdot}, \vec{\cdot})$  are the marginal and joint distribution of  $\vec{h}$  and  $\vec{g}$ , respectively. Note that power of quantization noise will be averaged over all realizations of the constellation, simplifying the computations. Let us explicitly include the index of the selected constellation point as a random variable,

$$\begin{aligned} \sigma_{n_q|\vec{h}}^2 &= \frac{1}{Q} \int_{\vec{g} \in \mathbb{R}^Q} \|\vec{h} - \vec{g}\|^2 \frac{\sum_{i=1}^M f_{\vec{g}, \vec{h}, I}(\vec{g}, \vec{h}, i)}{f_{\vec{h}}(\vec{h})} d\vec{g} \\ &= \frac{1}{Q} \int_{\vec{g} \in \mathbb{R}^Q} \|\vec{h} - \vec{g}\|^2 \frac{M f_{\vec{g}, \vec{h}, I}(\vec{g}, \vec{h}, \check{i})}{f_{\vec{h}}(\vec{h})} d\vec{g} \end{aligned} \quad (5)$$

where index  $\check{i} \in \{1, \dots, M\}$  is used to reflect that the corresponding terms are all equal to each other,  $f_{\vec{g}, \vec{h}, I}(\vec{g}, \vec{h}, i)$  is the joint distribution of the event:  $i$ th message is mapped to the constellation point  $\vec{h}$ , which is quantized to  $\vec{g}$ . Note that the event  $\{\vec{G} = \vec{g}, \vec{H} = \vec{h}, I = \check{i}\}$  is equivalent to

$$\{\vec{G} = \vec{g}, \vec{h}(\check{i}) = \vec{h}, \vec{h}(j) \notin \mathcal{S}^Q(\vec{g}, \|\vec{h} - \vec{g}\|), \forall j \neq \check{i}\}, \text{ where} \quad (6)$$

$$\mathcal{S}^Q(\vec{o}, r) = \{\vec{x} \in \mathbb{R}^Q : \|\vec{x} - \vec{o}\| \leq r\} \quad (7)$$

is a sphere of radius  $r$  centered at  $\vec{o}$ . Note that (6) captures the operation of quantization. We can rewrite (5) as

$$\frac{M}{Q} \int_{\vec{g} \in \mathbb{R}^Q} \|\vec{h} - \vec{g}\|^2 \frac{\mathbf{G}(\vec{h})\mathbf{G}(\vec{g}) (1 - P(\vec{g}, \vec{h}))^{M-1}}{\mathbf{G}(\vec{h})} d\vec{g}, \quad (8)$$

where  $f_{\vec{h}}(\vec{h})$  and  $f_{\vec{g}}(\vec{g})$  are replaced by  $\mathbf{G}(\vec{h})$  and  $\mathbf{G}(\vec{g})$  to emphasize Gaussian distribution.  $P(\vec{g}, \vec{h})$  is the probability of event  $\{\vec{h} \in \mathcal{S}^Q(\vec{g}, \|\vec{h} - \vec{g}\|)\}$ , i.e., a Gaussian point falls in a  $Q$ -dimensional hyper-sphere,  $\mathcal{S}^Q$ , centered at  $\vec{g}$  with radius  $r = \|\vec{h} - \vec{g}\|$ . This means,

$$P(\vec{g}, \vec{h}) = \int_{\vec{x} \in \mathcal{S}^Q(\vec{g}, \|\vec{h} - \vec{g}\|)} \mathbf{G}(\vec{x}) d\vec{x} \quad (9)$$

Applying  $\ln(A) \leq A - 1$ ,  $\forall A > 0$  to (8), results in

$$\begin{aligned} \frac{M}{Q} \int_{\vec{g} \in \mathbb{R}^Q} \|\vec{h} - \vec{g}\|^2 \mathbf{G}(\vec{g}) (1 - P(\vec{g}, \vec{h}))^{M-1} d\vec{g} &\leq \\ \frac{M}{Q} \int_{\vec{g} \in \mathbb{R}^Q} \|\vec{h} - \vec{g}\|^2 \mathbf{G}(\vec{g}) \exp[-(M-1)P(\vec{g}, \vec{h})] d\vec{g} \end{aligned} \quad (10)$$

The leading asymptotic behaviour of (10) is obtained using the Laplace method. For integrals of form:

$$S(M) = \int_a^b \psi(x) \exp[M\phi(x)] dx \quad (11)$$

if the real continuous function  $\phi(x)$  has its maximum in the interval  $a \leq x \leq b$  at an intermediate point  $x = c$ , then it is only the immediate neighbour of  $x = c$  that contributes to asymptotic expansion of  $S(M)$ . In our case,

$$S(M) = \frac{M}{Q} \int_{\vec{g} \in \mathbb{R}^Q} \mathbf{G}(\vec{g}) \|\vec{h} - \vec{g}\|^2 \exp[-(M-1)P(\vec{g}, \vec{h})] d\vec{g}. \quad (12)$$

$\phi(\cdot) = -P(\vec{g}, \vec{h})$  is maximized at  $\vec{g} = \vec{h}$ . For large values of  $M$ , the leading asymptotic behavior of the integral is governed by  $\vec{g}$  within a small sphere, say of radius  $r_0$ , around  $\vec{h}$ . To find the first order term, we approximate (12) as

$$S(M) \simeq \frac{M}{Q} \int_{\vec{g} \in \mathcal{S}^Q(\vec{h}, r_0)} \mathbf{G}(\vec{g}) \|\vec{h} - \vec{g}\|^2 \exp[-(M-1)P(\vec{g}, \vec{h})] d\vec{g} \quad (13)$$

where  $r_0$  is small enough such that  $P(\vec{g}, \vec{h})$ , i.e., the integration of Gaussian distribution over hyper-sphere  $\mathcal{S}^Q(\vec{h}, r_0)$ , can be approximated with  $\mathbf{G}(\vec{h})$  times the corresponding volume. Due to spherical symmetry of Gaussian,  $\sigma_{n_q|\vec{h}}^2$  only depends on norm  $\|\vec{h}\|$ , hence it suffices to compute the conditional variance for  $\vec{h} = \frac{\|\vec{h}\|}{\sqrt{Q}}(1, \dots, 1)_Q$ , where  $(1, \dots, 1)_Q$  is the all-one vector of size  $Q$ . Therefore,

$$\begin{aligned} S(M) &\simeq \frac{M}{Q} \int_{\vec{g} \in \mathbb{R}^Q} \mathbf{G}(\vec{h}) \left[ \left( g_1 - \frac{\|\vec{h}\|}{\sqrt{Q}} \right)^2 + \dots + \left( g_Q - \frac{\|\vec{h}\|}{\sqrt{Q}} \right)^2 \right] \\ &\exp \left\{ \frac{-\pi^{Q/2} (M-1) \mathbf{G}(\vec{h})}{\Gamma(Q/2+1)} \left[ \left( g_1 - \frac{\|\vec{h}\|}{\sqrt{Q}} \right)^2 + \dots \right]^{Q/2} \right\} dg_1 \dots dg_Q \end{aligned}$$

Applying change of variables, we obtain

$$\begin{aligned} \frac{M}{Q} \int_{r \in \mathbb{R}} \mathbf{G}(\vec{h}) r^2 \exp \left[ -(M-1) \mathbf{G}(\vec{h}) \frac{\pi^{Q/2}}{\Gamma(Q/2+1)} r^Q \right] \frac{\pi^{Q/2}}{\Gamma(Q/2+1)} dr^Q \\ = \frac{M}{Q} \int_{s \in \mathbb{R}} \mathbf{G}(\vec{h}) \left[ \frac{s \Gamma(Q/2+1)}{\pi^{Q/2}} \right]^{2/Q} \exp[-(M-1) \mathbf{G}(\vec{h}) s] ds \\ \sigma_{n_q|\vec{h}}^2 \simeq \frac{2\Gamma(2/Q+1)}{Q} \left[ \frac{\Gamma(Q/2+1)}{M} \right]^{2/Q} \exp \left( \frac{\|\vec{h}\|^2}{Q} \right), \quad (14) \\ \sigma_{n_q|\vec{h}}^2 \simeq \left( \frac{1}{M} \right)^{2/Q} = \left( \frac{1}{M} \right)^{1/K}. \end{aligned}$$

Figure 6 shows accuracy of above approximations, comparison with 256QAM, and an indication of the achieved energy saving due to inherent diversity.

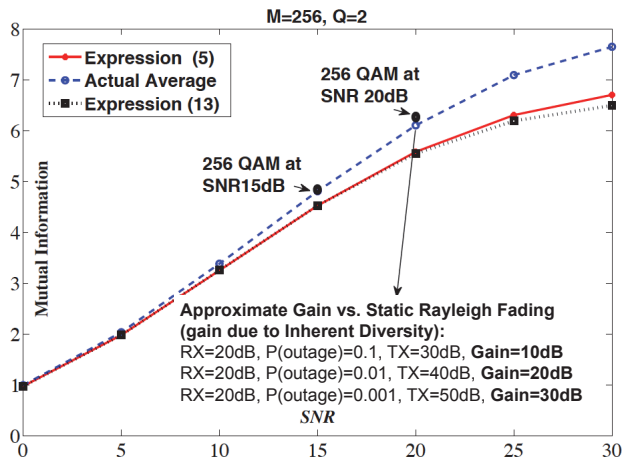


Fig. 6: Example of the accuracy of approximations, comparison with 256QAM, and an indication of the achieved energy saving due to inherent diversity.