Using a Gaussian Channel Twice

Bernhardsson, Bo

Published in:
International Symposium on Information Theory

2013

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

IEEE Information Symposium on Information Theory July 7-12, 2013. Istanbul, Turkey

Using a Gaussian Channel Twice

Bo Bernhardsson
Dep. of Automatic Control
Lund University, Sweden
Email: bob@control.lth.se

***Abstract*—The problem of communicating one bit over a memoryless Gaussian channel with an energy constraint is discussed. It is assumed that the channel is allowed to be used only two times. An ideal feedback channel is also supposed available. The optimal feedback strategy and the bit-error probability are derived. It is shown that feedback gives a significant performance gain and that the optimal strategy is discontinuous. It is also shown that most of the performance increase can be obtained even with a one-bit feedback channel.** [1]

I. Introduction

Shannon observed in [1] that feedback will not improve the capacity when communicating over a memory-less channel. This conclusion relies on the definition of capacity as a limiting case with arbitrary long blocks and no decoding delay constraints. Several authors have since then analysed different effects of feedback, see for instance [2], [3], [4], [5] and [6]. The current paper is inspired by the interesting results of [7] where it is shown that the Shannon-limit on -1.6dB energy per bit can be obtained even for the case of block length one, if a noise-free feedback channel is available. The obtained scheme however still has potentially unbounded decoding delay. To understand the benefits of feedback in the case of finite block lengths the extreme case with a decoding delay of two channel uses seems natural to study

A. Problem

We want to transmit the message $m \in \{0, 1\}$ where either message is equally likely. The coder/transmitter is assumed to send real numbers u_k using side-information from a causal noise-free feedback channel, see Fig. 1.

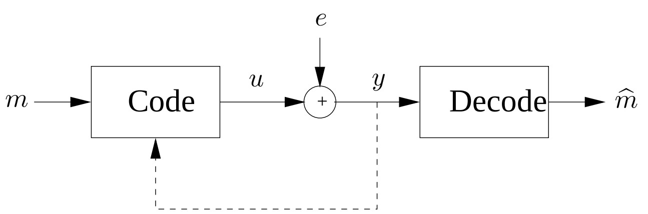

Fig. 1: The system studied in the paper. The feedback channel is assumed noise-free.

The signal is transmitted over a channel with additive Gaussian noise with unit variance. At time k the coder hence sends $u_k(y_{[0,k-1]}, m)$ and the decoder receives $y_k = u_k + e_k$, where $e_k \sim N(0, 1)$ is Gaussian noise. We will analyse the case with two transmissions, i.e.

$$\begin{aligned} y_0 &= u_0(m) + e_0 \\ y_1 &= u_1(y_0, m) + e_1. \end{aligned} \tag{1}$$

The task of this paper is to find optimal coder functions u_0 and u_1 so that an average energy constraint,

$$E(u_0^2 + u_1^2) \le S_{max} \tag{2}$$

is satisfied for a given level S_{max} and a decoder which minimizes the bit error probability

$$P^e = P(\widehat{m} \neq m).$$

We will use the notation $\varphi(t) = (2\pi)^{-\frac{1}{2}} e^{-t^2/2}$ and $Q(x) = \int_x^\infty \varphi(t)dt$. It is well known (e.g. [8]) that the optimal bit error rate without feedback is given by

$$P^e_{\text{no feedback}} = Q(\sqrt{S_{max}}), \tag{3}$$

which can be achieved by antipodal signaling $u_0 = \pm\sqrt{S_{max}}$ and $u_1 = 0$. There is no performance benefit with splitting the energy into several transmissions.

B. Optimal Decoder

The bit error probability is minimized by the Maximum Likelihood-decoder, which chooses the decoded message as

$$\widehat{m} = \arg\max_{i \in \{0,1\}} P(y_0, y_1 \mid m = i).$$

The decoder will output the message m that maximizes the posterior probability

$$\begin{aligned} &\log P(y \mid m) \\ &\quad = -\log(2\pi) - \frac{1}{2}(y_0 - u_0(m))^2 - \frac{1}{2}(y_1 - u_1(y_0, m))^2. \end{aligned}$$

The first transmission should be antipodal, i.e. $E(u_0) = \frac{1}{2}(u_0(1) + u_0(0)) = 0$, since a nonzero constant $E(u_0)$ does not carry any information and just wastes energy since $E(u_0^2) = E(u_0 - E(u_0))^2 + (E(u_0))^2$. We will use the notation

$$\begin{aligned} x_0 &:= u_0(1) = -u_0(0) \\ x_1(y_0) &:= \frac{1}{2}(u_1(y_0, 1) - u_1(y_0, 0)). \end{aligned}$$

[1]The research was supported by the ELLIIT Excellence Center and by the Swedish Research Council through the LCCC Linneaus Center.

Without loss of generality we assume that $x_0 \geq 0$, $u_1(y,1) \geq 0$ and $u_1(y,0) \leq 0$ for all y.

The decoded bit m is determined by the sign of

$$\begin{aligned}\log \frac{P(y \mid m=1)}{P(y \mid m=0)} &= \frac{1}{2}\big(-(y_0-x_0)^2+(y_0+x_0)^2 \\ &\quad -(y_1-u_1(y_0,1))^2+(y_1-u_1(y_0,0))^2\big) \\ &= 2y_0x_0 + 2\left(y_1 - \frac{u_1(y_0,1)+u_1(y_0,0)}{2}\right)x_1(y_0)\end{aligned}$$

If $m=1$ the bit is correctly decoded if

$$y_0x_0 + (x_1(y_0)+e_1)x_1(y_0) > 0, \qquad (4)$$

where $y_0 = x_0 + e_0$. For $m=0$ it is correctly decoded if

$$-y_0x_0 + (x_1(y_0)-e_1)x_1(y_0) > 0, \qquad (5)$$

where $y_0 = -x_0 + e_0$. It follows from (4) and (5) that the bit error probability is given by

$$\begin{aligned}P^e &= \frac{1}{2}\int_{-\infty}^{\infty} Q\left(\frac{y_0x_0}{x_1(y_0)} + x_1(y_0)\right)\varphi(y_0-x_0)dy_0 \\ &\quad + \frac{1}{2}\int_{-\infty}^{\infty} Q\left(\frac{-y_0x_0}{x_1(y_0)} + x_1(y_0)\right)\varphi(y_0+x_0)dy_0 \qquad (6) \\ &= \frac{1}{2}\int_{-\infty}^{\infty}\left(Q\left(\frac{y_0x_0}{x_1(y_0)}+x_1(y_0)\right)\right. \\ &\quad \left. + Q\left(\frac{y_0x_0}{x_1(-y_0)}+x_1(-y_0)\right)\right)\varphi(y_0-x_0)dy_0. \qquad (7)\end{aligned}$$

The expected energy in the left hand side of (2) is given by

$$S := x_0^2 + \frac{1}{2}\int_{-\infty}^{\infty} u_1^2(y_0,1)\varphi(y_0-x_0) + u_1^2(y_0,0)\varphi(y_0+x_0)dy. \qquad (8)$$

C. Optimal Encoder

For each y_0, minimizing the integrand in S over $u_1(y_0,1)$ and $u_1(y_0,0)$ subject to the constraint

$$\frac{1}{2}(u_1(y_0,1) - u_1(y_0,0)) = x_1(y_0)$$

is a convex quadratic optimization problem with a linear constraint. We can therefore use the following result:

Lemma *Assume that $W > 0$ and that A has full row rank. The minimum of x^TWx subject to $Ax = b$ is then obtained for $x = W^{-1}A^T(AW^{-1}A^T)^{-1}b$ and is given by $x^TWx = b^T(AW^{-1}A^T)^{-1}b$.*

Using this we obtain that

$$\begin{bmatrix} u_1(y_0,1) \\ u_1(y_0,0)\end{bmatrix} = \frac{2x_1(y_0)}{\varphi(y_0-x_0)+\varphi(y_0+x_0)}\begin{bmatrix}\varphi(y_0+x_0) \\ -\varphi(y_0-x_0)\end{bmatrix} \qquad (9)$$

and the energy (8) becomes

$$\begin{aligned}S &= x_0^2 + 2\int_{-\infty}^{\infty}\frac{x_1^2(y_0)}{\varphi^{-1}(y_0-x_0)+\varphi^{-1}(y_0+x_0)}dy_0 \\ &= x_0^2 + 2\int_0^{\infty}\frac{x_1^2(y_0)+x_1^2(-y_0)}{\varphi^{-1}(y_0-x_0)+\varphi^{-1}(y_0+x_0)}dy_0. \qquad (10)\end{aligned}$$

Since the integrands in (7) and (10) are unchanged if $x_1(y_0)$ is changed to $x_1(-y_0)$ we can assume that $x_1(y_0)$ is symmetric in y_0. From (9) we can then conclude that

$$u_1(y_0,1) = -u_1(-y_0,0) = \frac{2}{1+e^{2y_0x_0}}x_1(y_0).$$

To find $x_0 \geq 0$ and a symmetric function $x_1(\cdot) \geq 0$ minimizing P_e under the energy constraint $S \leq S_{max}$ we introduce a Lagrange-multiplier $\lambda > 0$ obtain the following result.

Theorem *The optimal feedback strategy x_0, $u_1(\cdot)$ can be found by solving*

$$\min_{x_0,x_1(\cdot)} L(x_0,x_1(\cdot)) = \min_{x_0,x_1(\cdot)} P^e(x_0,x_1(\cdot)) + \lambda S(x_0,x_1(\cdot)). \qquad (11)$$

and using

$$u_1(y_0,1) = -u_1(-y_0,0) = \frac{2}{1+e^{2y_0x_0}}x_1(y_0).$$

For a given x_0 we can find $x_1(y_0)$ from the implicit equation

$$x_1 \exp\left(\frac{y_0^2x_0^2}{2x_1^2} + \frac{x_1^2}{2}\right) - \frac{\cosh(y_0x_0)}{2\sqrt{2\pi}\lambda} = 0. \qquad (12)$$

The optimal x_1 equals either zero or the largest real root of (12), depending on which case gives the smallest value of the integrand in

$$\begin{aligned}P^e = \int_0^{\infty}&\left(Q\left(\frac{y_0x_0}{x_1(y_0)}+x_1(y_0)\right)\varphi(y_0-x_0)\right. \\ &\left. + Q\left(\frac{-y_0x_0}{x_1(y_0)}+x_1(y_0)\right)\varphi(y_0+x_0)\right)dy_0. \qquad (13)\end{aligned}$$

Proof

To find the optimal communication scheme we will fix the Lagrange multiplier λ and for each x_0 optimize the integrand of $P^e + \lambda S$ over $x_1(y_0)$ for each y_0. The optimal x_0 is then found by a one-dimensional search. The procedure is repeated for different values of λ resulting in a curve of achievable bit-error P^e vs power S. The resulting P^e and S are continuous functions of λ, from which it follows that the method of Lagrange multipliers used actually finds the pareto-optimal boundary of the (convex) domain of achievable (P^e, S).

Optimizing L over $x_1(y_0)$ can be done separately for each y_0. We therefore seek the infimum of

$$\begin{aligned}&Q\left(\frac{y_0x_0}{x}+x\right)\varphi(y_0-x_0) + Q\left(\frac{-y_0x_0}{x}+x\right)\varphi(y_0+x_0) \\ &\quad + \frac{4x^2\lambda}{\varphi^{-1}(y_0-x_0)+\varphi^{-1}(y_0+x_0)} \qquad (14)\end{aligned}$$

with respect to $x := x_1(y_0)$. Using $Q'(x) = -\varphi(x)$, we see there are stationary points when

$$\begin{aligned}0 = \frac{dL}{dx} &= -\frac{1}{\pi}\exp\left(-\frac{y_0^2x_0^2}{2x^2} - \frac{x^2}{2} - \frac{y_0^2}{2} - \frac{x_0^2}{2}\right) \\ &\quad + \frac{8x\lambda}{\varphi^{-1}(y_0-x_0)+\varphi^{-1}(y_0+x_0)}.\end{aligned}$$

This is an implicit equation in $x = x_1(y_0)$ for each y_0 which can be simplified to

$$x \exp\left(\frac{y_0^2 x_0^2}{2x^2} + \frac{x^2}{2}\right) - \frac{\cosh(y_0 x_0)}{2\sqrt{2\pi}\lambda} = 0,$$

where the left hand side has the same sign as dL/dx. It is easily seen that there are at most two real positive solutions of (12) and that the sign of the derivative goes from positive to negative to positive. This means that the smallest value of L is taken either at $x = 0$ or at the largest solution x^* of (12). For $x = 0+$ the expression (14) becomes

$$\varphi(|y_0| + x_0).$$

This value should therefore be compared with

$$Q\left(\frac{y_0 x_0}{x^*} + x^*\right)\varphi(y_0 - x_0) + Q\left(-\frac{y_0 x_0}{x^*} + x^*\right)\varphi(y_0 + x_0) + \frac{4x^{*2}\lambda}{\varphi^{-1}(y_0 - x_0) + \varphi^{-1}(y_0 + x_0)},$$

where x^* is the largest solution from the implicit equation above and the alternative with smallest result should be chosen. An alternative is to directly minimize

$$Q\left(\frac{y_0 x_0}{x} + x\right)\varphi(y_0 - x_0) + Q\left(-\frac{y_0 x_0}{x} + x\right)\varphi(y_0 + x_0) + \frac{4x^2\lambda}{\varphi^{-1}(y_0 - x_0) + \varphi^{-1}(y_0 + x_0)}$$

over x.

Note that the solution $x_1(y_0) = 0$ corresponds to that the second transmission is not used. Close analysis shows that the value of x for which the minima occurs can be a discontinuous function of y, see Figure 5.

II. Results

Figures 2-3 compare achievable performance for optimal transmission without use of feedback (top blue) and optimal transmission with use of feedback(black). Also shown is a suboptimal feedback scheme (red) corresponding to a constant $x_1(y_0) \equiv x_1$ (such as used in [7]). There is a significant performance gain of many dBs using feedback. The performance gain increases with SNR. The suboptimal scheme with constant $x_1(y_0) \equiv x_1$ (which for each λ was optimized jointly with x_0) is rather close to optimal, except for the low SNR regime where the optimal scheme outperforms the suboptimal with some tenths of dBs. Notice also that the feedback scheme obtainable with one-bit feedback (red dashed) captures most of the performance gain with feedback. The one-bit feedback scheme was obtained by assuming the feedback to give information about whether or not $|y_0| \leq a$. The level a was found by straight-forward search. We have not been able to prove that this is the optimal use of the one-bit feedback channel, but think it is true.

The optimal use of power in the second transmission, determined by $x_1(y_0)$, is interesting. The function $x_1(y_0)$ turns out to be discontinuous, showing that the second transmission should not be used if the first output y_0 is far away from zero. The discontinuity is most pronounced in the low SNR regime, for high SNR the discontinuity threshold moves to very high levels of y_0, corresponding to turning off the 2nd transmission only at exteremly unlikely outcomes from the first transmission. Note that for low SNR the second transmission is used mainly when y_0 is close to zero. A majority of the power is used for the first transmission. The optimal $u_1(y_0, 1)$ and $u_1(y_0, 0)$ for $S_{max} = 2.42$ is illustrated in Fig 6.

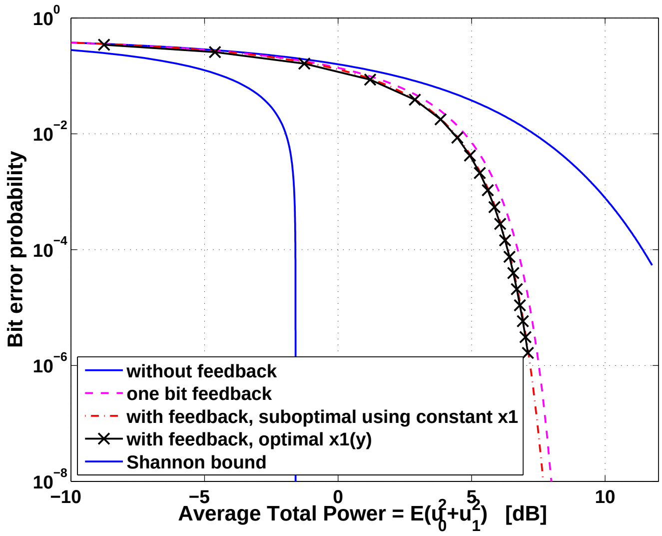

Fig. 2: Bit error probability versus average power: Optimal transmission without use of feedback (full), one-bit feedback scheme (dashed) suboptimal feedback scheme (dash-dotted), optimal feedback scheme (full-x), Shannon bound for infinite-block transmissions (full). Notice the significant performance gain with feedback, even using only one-bit feedback.

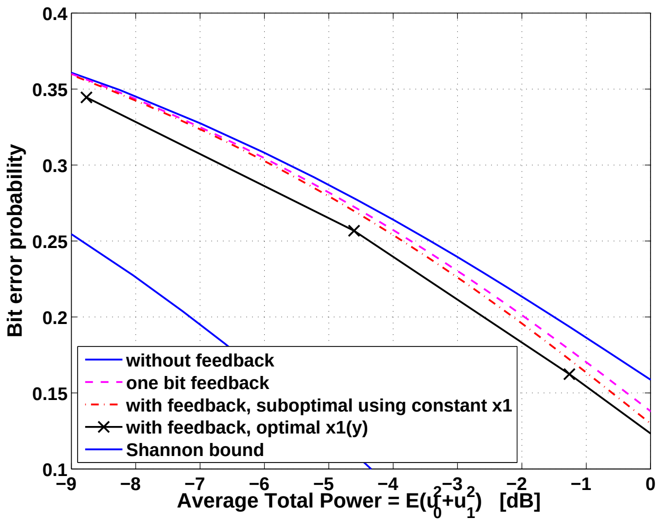

Fig. 3: Zoom of previous figure. There is a performance cost of 0.5-1dB with the one-bit feedback channel, compared to using an infinite-capacity feedback channel.

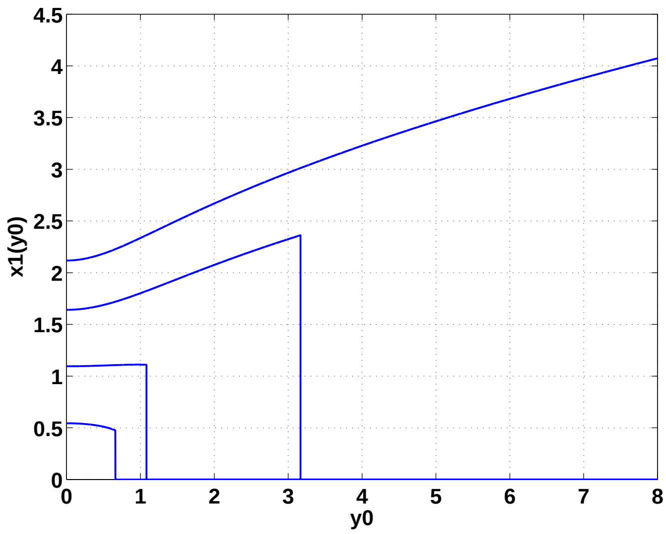

Fig. 4: The optimal $x_1(y_0)$, which is discontinuous, is shown for four cases corresponding to BER of 0.26(lowest), 0.09, 0.02, 0.004 (highest) respectively. The discontinuity of the upper curve is outside the visable range. For the four different cases we have $x_0 = 0.48, 0.89, 1.19, 1.39$ respectively and the total power $S_{max} = 0.35, 1.32, 2.42, 3.12$.

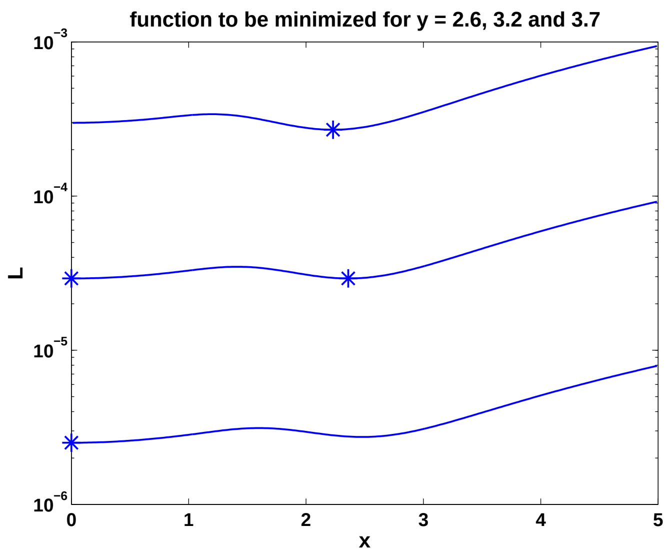

Fig. 5: The minima of the function in (14) change when y_0 changes. This results in a discontinuity of the function $x_1(y_0)$. The figure corresponds to $x_0 = 1.19, S_{max} = 2.42$ and three values of y_0 around 3.2, compare Fig 4.

III. Conclusion

The optimal feedback scheme for transmission of one bit of information over a energy constrained Gaussian channel has been found for the case when the Gaussian channel can be used two times. The optimal scheme is discontinuous, but a continuous simpler suboptimal scheme can be found

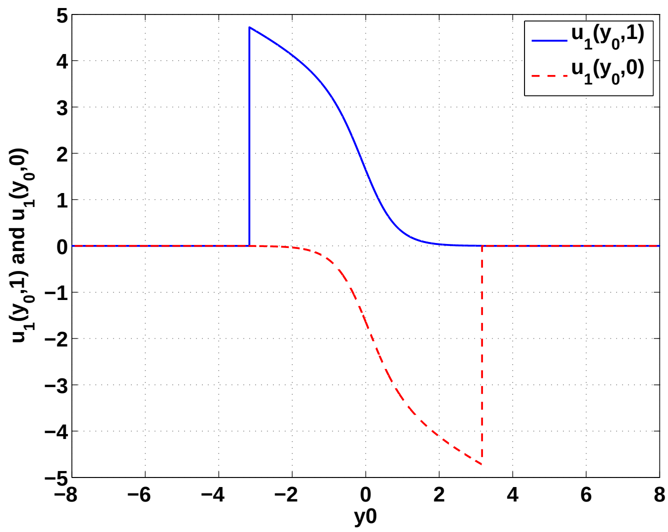

Fig. 6: The optimal functions $u_1(y_0, 1)$ and $u_1(y_0, 0)$ when $S_{max} = 2.42$ and $x_0 = 1.19$.

with rather similar performance. The generalization to longer decoding delay constraints is open.

References

[1] C. Shannon, "The zero error capacity of a noisy channel,," *Transactions on Information Theory*, vol. IT-2, pp. 8–19, sep 1956.

[2] M. Horstein, "Sequential transmission using noiseless feedback," *Information Theory, IEEE Transactions on*, vol. 9, no. 3, pp. 136 – 143, jul 1963.

[3] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–i: No bandwidth constraint," *Information Theory, IEEE Transactions on*, vol. 12, no. 2, pp. 172 – 182, apr 1966.

[4] J. Schalkwijk, "A coding scheme for additive noise channels with feedback–ii: Band-limited signals," *Information Theory, IEEE Transactions on*, vol. 12, no. 2, pp. 183 – 189, apr 1966.

[5] R. Gallager and B. Nakiboglu, "Variations on a theme by Schalkwijk and Kailath," *Information Theory, IEEE Transactions on*, vol. 56, no. 1, pp. 6 –17, jan. 2010.

[6] O. Shayevitz and M. Feder, "Optimal feedback communication via posterior matching," *Information Theory, IEEE Transactions on*, vol. 57, no. 3, pp. 1186 –1222, march 2011.

[7] Y. Polyanskiy, H. Poor, and S. Verdu, "Minimum energy to send k bits through the Gaussian channel with and without feedback," *Information Theory, IEEE Transactions on*, vol. 57, no. 8, pp. 4880 –4902, aug. 2011.

[8] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.