# Turbo Decoders Which Adapt to Noise Distribution Mismatch [1]

*Xiaoling Huang* and *Nam Phamdo*

Department of Electrical and Computer Engineering
State University of New York at Stony Brook
Stony Brook, NY 11794-2350
xhuang@sbee.sunysb.edu, phamdo@sbee.sunysb.edu

## Abstract

This paper investigates the sensitivity of Turbo decoder performance to mismatch of the noise distribution, and proposes a simple on-line procedure for estimating the unknown noise distribution from each block of received signal. This procedure consists of (i) quantization of the received signal, and (ii) estimation of the noise distribution from the histogram of the quantized received signal. For Gaussian and Laplacian noise, the proposed procedure leads to decoder performances which are comparable (within 0.1 dB at BER $10^{-4}$) to the case where the noise distribution is *known exactly*.

**Keywords**: Turbo codes, robustness, generalized Gaussian noise.

**Submitted to IEEE Communications Letters, February 1998, revised August 1998**

[1]This work is supported in part by Nippon Telegraph and Telephone Corporation, Northrop Grumman Corporation, Line Imaging Systems, and the New York State Strategic Partnership for Industrial Resurgence (SPIR) Program.

## List of Tables



## List of Figures

## I Introduction

Since the invention of Turbo codes [1]-[3] five years ago, the coding research community has been brewing with interest and excitement. This is due to the record breaking performances of Turbo codes which come very close (within 0.5 dB at $10^{-4}$ bit error rate (BER)[1]) to the Shannon theoretical limit. These "near optimum" performances are achieved by a parallel concatenated encoder which consists of two or more recursive systematic convolutional encoders separated by one or more interleaver(s). Decoding of Turbo codes is achieved by an iterative procedure based upon a modified Bahl et al. algorithm [1]. Unlike traditional minimum distance decoders which typically implement the Viterbi algorithm, Turbo decoders require knowledge of the channel signal-to-noise ratio (SNR) and the noise distribution (probability density function (pdf)).

Recently, Summers and Wilson [4], Jordan and Nichols [5] and Reed and Asenstorfer [6] investigated the sensitivity of the Turbo decoder to SNR mismatch. These works showed that the Turbo decoder is robust to SNR mismatch of $-2$ dB to $+6$ dB. Methods for on-line estimating the SNR were proposed in [4] and [6]. And the results indicate that "blind" Turbo decoder performs nearly identical to the Turbo decoder which has exact knowledge of the channel SNR. In [4] and [6],the authors assumed an additive white Gaussian noise (AWGN) channel and that the decoder knows that the noise distribution is Gaussian (though it does not know the variance of the distribution).

In this paper, we investigate the sensitivity of Turbo decoders to noise distribution mismatch, propose a simple method to estimate the noise distribution, and report some simulation results.

## II System Model

We consider an additive white noise channel in which the noise distribution is unknown. We use the rate-1/2 code presented in the papers by Berrou et al. [1], [2]. The block size of the code is $(n, k) = (131072, 65536)$, the code generator polynomial is (37,21) in octal notation, and the interleaver size is $256 \times 256$. The channel model is as follows:

$$y_i = x_i + z_i, \tag{1}$$

$i = 1, 2, \ldots, n$, where $x_i = \pm 1$ is the channel input, $z_i$ is an independent and identically distributed noise sequence with unknown pdf $f(z)$, and $y_i$ is the channel output. We assume that $f(z)$ has zero mean and even symmetry. For purpose of simulation, we will use a class of generalized Gaussian distributions:

$$f(z) = \frac{\alpha\eta(\alpha,\sigma)}{2\Gamma(1/\alpha)} \exp(-[\eta(\alpha,\sigma)|z|]^{\alpha}), \quad -\infty < z < \infty, \tag{2}$$

where $\eta(\alpha, \sigma) = \sigma^{-1}[\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}]^{1/2}$, $\sigma^2$ is the noise variance, $\Gamma(\cdot)$ is the Gamma function defined by $\Gamma(p) = \int_0^\infty x^{p-1}e^{-x}dx$. The parameter $\alpha \in (0, \infty)$ determines the pdf shape. The interested reader is referred to [8] for plots of the generalized Gaussian pdf for various values of $\alpha$. This class of distributions is chosen because it is a large class which contains several commonly used distributions as special cases. When $\alpha = 2$, $f(z)$ in

(2) is the Gaussian distribution. When $\alpha = 1$, it is the Laplacian distribution. As $\alpha \to \infty$, $f(z)$ approaches a uniform distribution.

## III Distribution Mismatch and Estimation

We first investigate the sensitivity of the Turbo decoder to distribution mismatch. In Figure 1, we plot simulation results in which the true noise pdf is Gaussian ($\alpha = 2$), but the decoder assumes a generalized Gaussian noise with parameter $\hat{\alpha}$. The decoder has exact knowledge of the channel SNR. The simulation is performed for 64 blocks of data, each of which consists of $k = 65536$ "information" bits. The decoder, which implements the modified Bahl et al algorithm described in [1]-[3], stops after 20 iterations. We do not terminate the code and hence the initial state probabilities of the backward algorithm are set to be uniform (see equation (36b) in [2]). When $\hat{\alpha} = 2$, the decoder is "matched" to the noise. Observe that at $10^{-4}$ BER, there is a mismatch loss of 0.75 dB when $\hat{\alpha} = 0.5$, 0.32 dB when $\hat{\alpha} = 1$, 0.31 dB when $\hat{\alpha} = 3$ and 0.65 dB when $\hat{\alpha} = 3.5$.

Next, we propose a simple method for "blind" Turbo decoding (without exact knowledge of the noise pdf and channel SNR). This method involves (i) quantization of the channel output $y_i$, and (ii) on-line estimation of the conditional probability mass function (pmf) of the quantized signal.

The purpose of quantization is to reduce the problem of estimating the conditional pdf to estimating the conditional pmf which, in general, is much easier. We use an $N$-level uniform quantizer (assume $N$ is an integer power of 2). When $N \leq 16$, the quantization thresholds are chosen as:

$$T_j = \begin{cases} -\infty & \text{if } j = 0 \\ j/2^{q-2} - 2 & \text{if } j = 1, 2, \ldots, 2^q - 1 \\ +\infty & \text{if } j = 2^q \end{cases}, \tag{3}$$

where, $q = \log_2 N$. When $N \geq 32$, the following thresholds are used:

$$T_j = \begin{cases} -\infty & \text{if } j = 0 \\ j/2^{q-4} - 8 & \text{if } j = 1, 2, \ldots, 2^q - 1 \\ +\infty & \text{if } j = 2^q \end{cases}. \tag{4}$$

Note that the quantization stepsize is partially dependent on the number of levels. This is to keep the quantizer support region within a reasonable range. Though it is not relevant, we place the quantization levels at the midpoint of each quantization region. We chose this quantizer because it is "symmetric" around $-1, 0$, and 1. The reason for this will soon become clear.

Let $\hat{y}_i$ be the quantized received signal. We now use $\{\hat{y}_i\}_{i=1}^n$ as input to the Turbo decoder. With the quantized input, the Turbo decoder now requires the conditional pmfs: $p(\hat{y}_i|x_i = -1)$ and $p(\hat{y}_i|x_i = +1)$. We now describe how these two pmfs are estimated from the quantized observation $\{\hat{y}_i\}_{i=1}^n$.

For each block of $n = 131072$ received quantized samples $\{\hat{y}_i\}_{i=1}^n$, we calculate the histogram, $h(\hat{y})$, of $\hat{y}_i$. Note that $\hat{y}$ takes on one of $N$ possible values. Next, we symmetrize the histogram by calculating its even part: $h_s(\hat{y}) = (h(\hat{y}) + h(-\hat{y}))/2$. This is possible because the quantizer is symmetric around 0. We then estimate $p(\hat{y}|x = +1)$ by copying a portion of $h_s(\hat{y})$ and then symmetrizing it around +1 (due to the assumption that $f(z)$ has even symmetry). The left tail end of the distribution is then obtained by a linear approximation. When $N \geq 32$, the exact formula for $p(\hat{y}|x = +1)$ is given as follows:

$$p(\hat{y}|x = +1) = \begin{cases} h_s(\hat{y}) & \text{if } \hat{y} \geq +1 \\ h_s(2 - \hat{y}) & \text{if } -6 \leq \hat{y} < +1 \\ (\hat{y} + a)h_s(a)/2 & \text{if } \hat{y} < -6 \end{cases}, \quad (5)$$

Where $a = 8 - \frac{16}{N}$. For $N \leq 16$, the exact formula for $p(\hat{y}|x = +1)$ is simpler:

$$p(\hat{y}|x = +1) = \begin{cases} h_s(\hat{y}) & \text{if } \hat{y} \geq +1 \\ h_s(2 - \hat{y}) & \text{if } -2 \leq \hat{y} < +1 \end{cases}. \quad (6)$$

If $p(\hat{y}|x = +1)$ is zero, we set it to a small number. $p(\hat{y}|x = -1)$ is estimated in a similar manner and will be symmetric to $p(\hat{y}|x = +1)$. This "copy-and-symmetrize" method is possible due to the symmetry of the quantizer around $\pm 1$.

In Figure 2, we plot the true and estimated pmfs for Gaussian noise at 1.0 dB SNR. Visually, the estimated pmfs are nearly identical to the true pmfs.

## IV Simulation Results

First, the influence of the quantizer on decoding performance is examined. We simulate the Turbo decoder for an AWGN channel in which the channel output $\{y_i\}_{i=1}^n$ is quantized using an $N$-level quantizer with $N = 4, 8, 16, 32, 64$, and $\infty$ (unquantized). It is assumed that the decoder has the *exact* conditional pmfs: $P(\hat{y}|x = -1)$ and $P(\hat{y}|x = +1)$. At BER $= 10^{-4}$, there is a performance loss of 0.04 dB when $N = 64$ as compared to the unquantized case ($N = \infty$). Similarly, when $N = 32, 16, 8, 4$, the loss is 0.15 dB, 0.05 dB, 0.15 dB and 0.55 dB, respectively. The non-monotonicity of the loss (as a function of $N$) is due to the different quantizer stepsize used in (3) and (4).

There is a fundamental tradeoff in the selection of $N$. For small $N$, the estimation of the conditional pmfs is easy because there are few values to estimate, but the irreversible loss of information due to the quantization (hard-decision demodulation) is high. On the other hand, for large $N$, the information loss is low but the pmf estimation becomes less accurate. In light of this discussion, we chose $N = 64$ because it is large enough to yield almost no loss of information and it is small enough to get a reasonably accurate estimation of the conditional pmfs.

In Figure 3, we present simulation results of the Turbo decoder for Gaussian, Laplacian, uniform and generalized Gaussian ($\alpha = 0.5$) noise. For each noise distribution, we plot the decoder performance for three cases: (a) The decoder knows the exact noise distribution and it does not quantize the received signal (True-Pdf); (b) The decoder knows the exact noise distribution but it quantizes the received signal with

$N = 64$ (True-Pmf); (c) The decoder does not know the noise distribution and it decodes the quantized received signal using the method described in Section III (Estimated-Pmf). In this way, one can determine whether the performance loss (if there is any) is due mainly to the quantization or to the inaccuracy in the pmf estimation.

For Gaussian and Laplacian noise, the Turbo decoder performances are comparable (within 0.1 dB at BER $= 10^{-4}$) to the case where the noise distribution is known exactly (matched decoder). For uniform and generalized Gaussian noise ($\alpha = 0.5$), the Turbo decoder performance losses are about 0.36 dB and 0.39 dB respectively, compared to the performances of the matched decoders. For the uniform noise, the main performance loss is caused by the quantization. This is because the quantizer given by (4) is not a particularly "good" choice for uniform noise at low SNR. For the generalized Gaussian noise ($\alpha = 0.5$), the main performance loss is from the inaccuracy of the pmf estimation. We attribute this to the difficulty of estimating the noise distribution as SNR goes lower. Note that the "knees" of the performance curves are steeper for large values of $\alpha$.

In Table 1, we summarize the performance of the proposed blind decoder and compare it to (i) the mismatched decoder in which the decoder assumes Gaussian noise ($\hat{\alpha} = 2$) and there is no quantization, (ii) the matched decoder ($\hat{\alpha} = \alpha$) without quantization and (iii) the Shannon capacity limit of the *binary-input* channel. The Shannon limit is calculated using the Blahut algorithm [7].

| True $\alpha$ | Mismatched Decoder ($\hat{\alpha} = 2$) | Proposed "Blind" Decoder | Matched Decoder ($\hat{\alpha} = \alpha$) | Shannon Capacity Limit |
|---|---|---|---|---|
| 0.5 | 4.63 | $-2.27$ | $-2.66$ | $-3.97$ |
| 1.0 | 1.55 | 0.43 | 0.38 | $-0.54$ |
| 1.5 | 0.84 | 0.75 | 0.72 | 0.07 |
| 2.0 | 0.65 | 0.70 | 0.65 | 0.19 |
| 2.5 | 0.64 | 0.64 | 0.62 | 0.18 |
| 3.0 | 0.63 | 0.59 | 0.58 | 0.13 |
| $\infty$ | 0.66 | $-0.24$ | $-0.60$ | $-1.23$ |

Table 1: SNR(dB) Required to Achieve BER $= 10^{-4}$.

The majority of current research on Turbo codes deals with reducing the size of the interleaver (hence reducing the code latency). It is reasonable to ask: How would the proposed method perform when there are fewer observation available for pmf estimation? To answer this question, we have performed a simulation of the blind decoder in which only $m = 8192$ out of $n = 131072$ samples in each block are used for pmf estimation. For an AWGN channel, this decoder achieves BER of $10^{-4}$ at 0.84 dB SNR. Hence, we expect the performance of the proposed blind decoder to hold up even for interleaver size as small as $64 \times 64$.

## V Conclusion

In this paper, we investigate the robustness of the Turbo decoder to noise distribution mismatch and propose a simple "copy-and-symmetrize" method for on-line estimation of the noise distribution. The proposed procedure leads to "relatively good" performance for the class of generalized Gaussian distributions. We observe, from the last two columns of Table 1, that Turbo codes also perform very well for non-Gaussian noise.

**Acknowledgement**

The authors thank an anonymous reviewer and Todd Summers for valuable comments which lead to a major improvement of this paper.

## References

[1] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes," *Proceedings of 1993 International Conference on Communications*, pp. 1064-1070.

[2] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo Codes," *IEEE Transactions on Communications,* Vol. 44, No. 10, pp. 1261-1271, October 1996.

[3] S. Benedetto and G. Montorsi, "Unveiling Turbo Codes: Some Results on Parallel Concatenated Coding Schemes," *IEEE Transactions on Information Theory,* Vol. 42, No. 2, pp. 409-428, March 1996.

[4] T. A. Summers and S. G. Wilson, "SNR Mismatch and On-Line Estimation in Turbo Decoding," *IEEE Transactions on Communications*, Vol. 46, No. 4, pp. 421-423, April 1998.

[5] M. A. Jordan and R. A. Nichols, "The Effects of Channel Characteristics on Turbo Code Performance," *Proc. of Milcom 96*, pp. 17-21.

[6] M. C. Reed and J. Asenstorfer, "A Novel Variance Estimator for Turbo-Code Decoding," In *International Conference on Telecommunications*, pp. 173-178, April 1997.

[7] R. E. Blahut, "Computation of Channel Capacity and Rate-Distortion Functions," *IEEE Transactions on Information Theory*, Vol. IT-18, No. 4, pp. 460-473, July 1972.

[8] N. Farvardin and V. Vaishampayan, "Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source-Channel Coding," *IEEE Transactions on Information Theory*, Vol. IT-33, No. 6, pp. 827-838, November 1987.

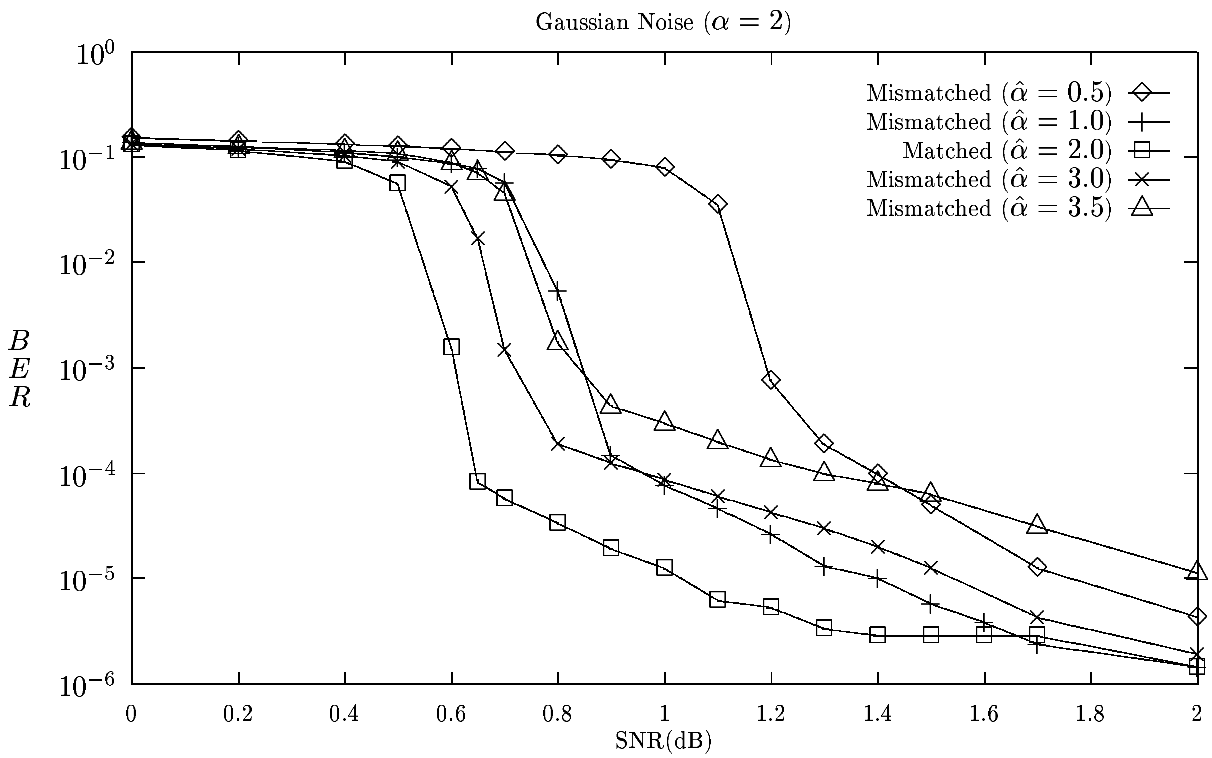


Figure 1: Performances of Turbo Decoders Under Pdf-Matched and Pdf-Mismatched Conditions; True Noise Pdf is Gaussian ($\alpha = 2$); Decoder Assumes Noise Pdf is Generalized Gaussian with Parameter $\hat{\alpha}$.

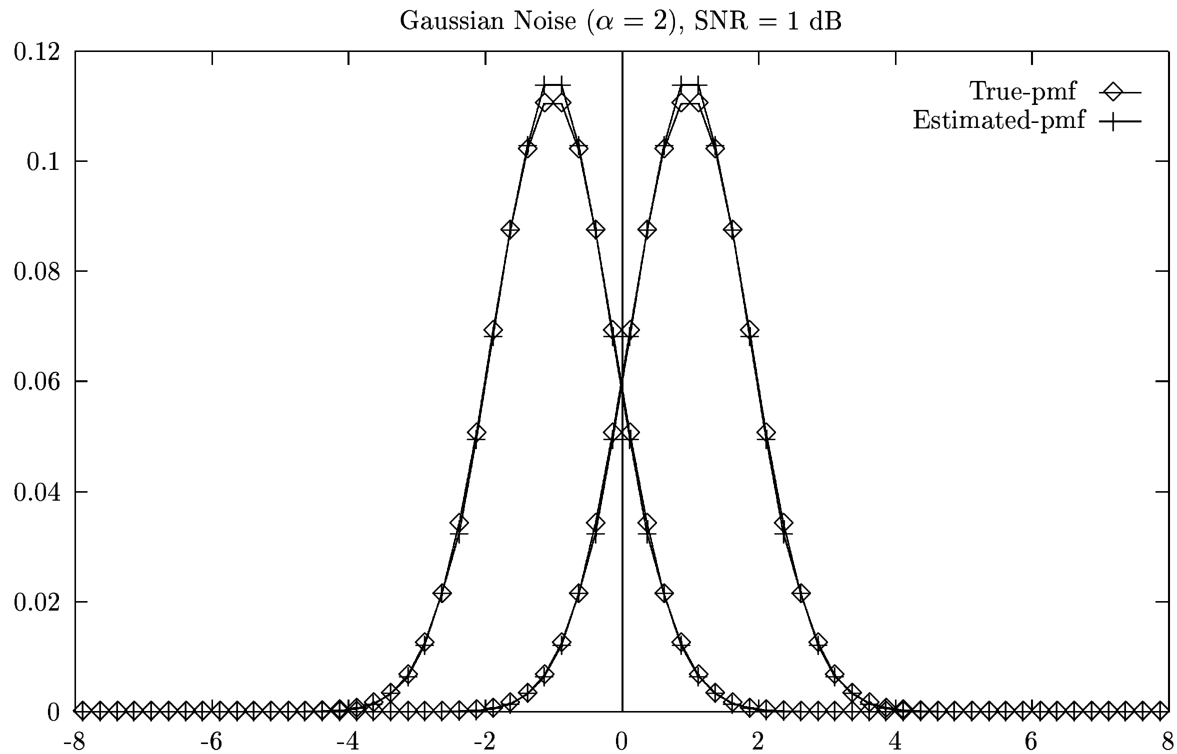


Figure 2: True and Estimated Conditional Probability Mass Functions $p(\hat{y}|x = -1)$ and $p(\hat{y}|x = +1)$; Gaussian Noise at 1.0 dB SNR; Estimated PMF is Obtained from a Block of $n = 131072$ Quantized Received Signals.

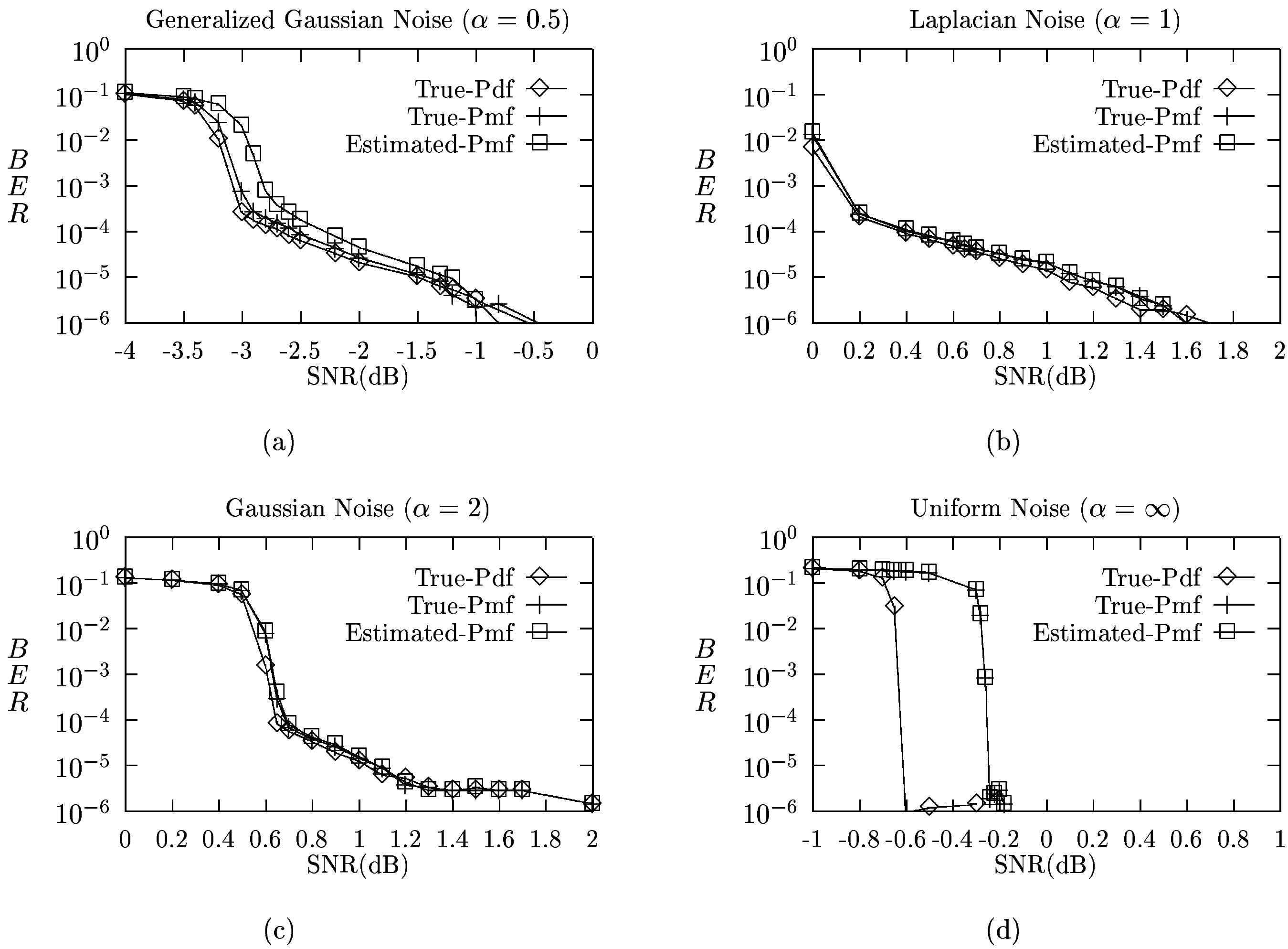


Figure 3: Decoding Performance for Generalized Gaussian Noise with Parameter $\alpha$ Based on True Pdf, True Pmf and Estimated Pmf for Rate-1/2 Code; (a) $\alpha = 0.5$, (b) $\alpha = 1$, (c) $\alpha = 2$, (d) $\alpha = \infty$.