A Fast Bayesian Algorithmic Enhancement For Real Time Bit Error Ratio Test (BERT)

Shu-Yi Jack Wong, Jianwen Zhu

Department of Computer Engineering, University of Toronto, Ontario, Canada M5G 3G4

Abstract

Continuous, real-time bit error ratio (BER) test of modern communication and storage channels is a ubiquitous problem: the noises tend to vary in space and time, and are difficult to fully characterize offline. Traditional method requires time consuming accumulation of samples for which Bayesian method has shown its promise in alleviating by incorporating *a priori* knowledge. However, the method has so far depended on a simplistic linear search algorithm that suffers from long running time which defeats the purpose of sample reduction. We reveal the existence of a convex solution space for the problem generally thought to be non-convex. This results in an improved performance by orders of magnitude.

Introduction

Consider G, a Bernoulli RV (Random Variable) of a yield θ or the probability of success:

$$\theta = P(G = success)$$

If we estimate θ with a finite number of Bernoulli trials

$\theta \in RV \quad \Theta$ $\qquad$ θ itself becomes a RV

Knowing a particular θ, the "Likelihood" of x successes out of n Bernoulli trials is given by the binomial distribution as below:

$$f_{X|\Theta}(x|\theta;n) = \binom{n}{x}\theta^{x}(1-\theta)^{n-x}$$

Conversely, with known x and n but unknown θ

$$\hat{\theta}_{MLE}(x;n) = \arg\max_{\theta} f_{X|\theta}(x|\theta;n) = \frac{x}{n}$$

Maximum Likelihood Estimate (MLE) is well known but there is one caveat:

An accurate estimate demands a large n when θ is extremely small

To avoid potentially the long delay for accumulating a large number of samples, we consider the *a priori* probability density function (PDF) $f_{\Theta}(\theta)$.

By Bayesian method, the *a priori* PDF is fused with new observations, leading to the *a posteriori* PDF:

$$f_{\Theta|X}(\theta|x;n) = \frac{f_{X|\Theta}(x|\theta;n)\cdot f_{\Theta}(\theta)}{f_X(x;n)}$$

$$f_X(x;n) = \int_{\theta} f_{X|\Theta}(x|\theta;n)\cdot f_{\Theta}(\theta)\cdot d\theta \qquad \text{Evidence Function}$$

Finding the most likely θ from the *a posteriori* PDF

$$\hat{\theta}_{MAP}(x;n) = \arg\max_{\theta} f_{\Theta|X}(\theta|x;n)$$

The most likely θ locates at the mode (the peak) of the *a posteriori* PDF

Problem Definitions

When *a priori* knowledge does not necessarily include *a priori* observations, what consists of a reasonable reconstruction of $f_{\Theta}(\theta)$?

It is well known that the *a priori* PDF for a Bernoulli likelihood is a Beta distribution which, with the knowledge of its *mode* (m), could be expressed as below:

$$f_{\Theta}(\theta;a,m) = \frac{\Gamma\left(\frac{a-1}{m}+2\right)}{\Gamma(a)\cdot\Gamma\left[\frac{(a-1)(1-m)}{m}+1\right]}\cdot\theta^{a-1}\cdot(1-\theta)^{\frac{(a-1)(1-m)}{m}}$$

"*a*" is unknown

How is it determined?

$$\hat{a} = \arg\max_{a} f_X(x,a;n,m)$$

Fang et al. [1] proposed finding "*a*" by maximizing the "*evidence function*"

This amounts to the maximization of the following objective function

$$f_X(x,a;n,m) = \frac{\Gamma\left(\frac{a-1}{m}+2\right)}{\Gamma(a)\cdot\Gamma\left[\frac{(a-1)(1-m)}{m}+1\right]}\cdot\frac{\Gamma(x+a)\cdot\Gamma\left[n-x+\frac{(a-1)(1-m)}{m}+1\right]}{\Gamma\left(n+\frac{a-1}{m}+2\right)}$$

The Objective Function !!

The objective function is generally considered non-convex and linear search is thus used for solving it. However, for low BER such as 10^{-12}, the search space for "*a*" is in the order of 10^{13} and long search time becomes inevitable.

Is it possible to shorten the search time for real-time applications?

Proposed Method

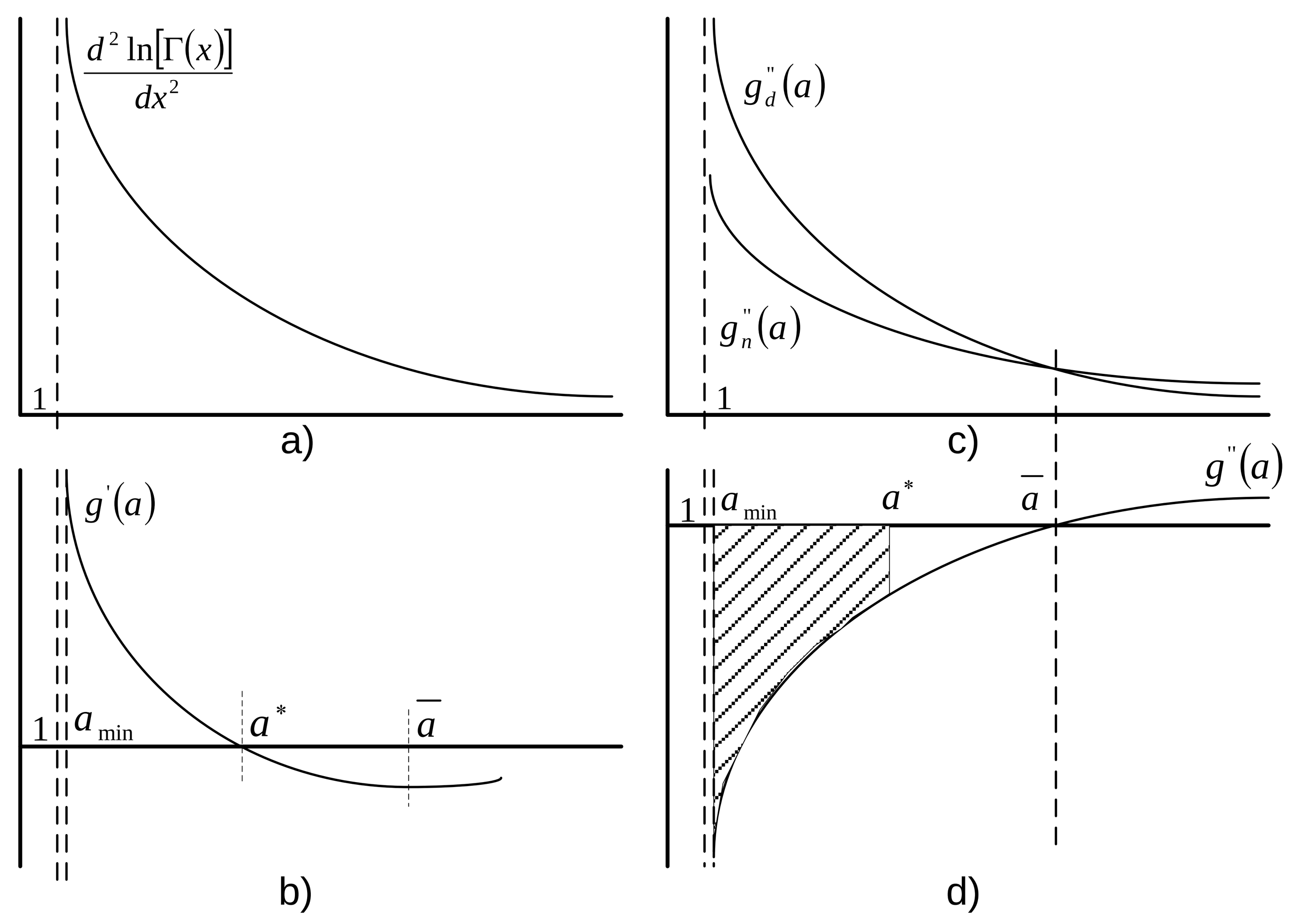
We derive an equivalent maximization problem that is perhaps easier. This allows us to identify the region of non-convexity in order to show that it is in fact irrelevant to the solution space. The maximizer is, therefore, not necessarily bounded to slow algorithm such as linear search.

We Solve $\qquad \hat{a} = \arg\max_{a} \ln\left[f_X(x,a;n,m)\right]$

By applying the approximation of [2]

$$\frac{d\ln[\Gamma(x)]}{dx} \sim \ln(x-1) + \frac{1}{2(x-1)} - \frac{1}{12(x-1)^2}$$

Digamma Approximation



- Figure a) shows the trend for the derivative of Digamma that decreases monotonically without changing polarity. Therefore, it is convex.
- It follows in Figure c) that the second derivative for the numerator $g_n''(a)$ and the denominator $g_d''(a)$ of our new objective function are also convex.
- Figure d) shows that the two functions may still meet at a common point $\bar{a}$ leading to a change in polarity of the second derivative and results in non-convexity.
- Figure b) shows maximization by finding the root a^* for the first derivative $g'(a)$.

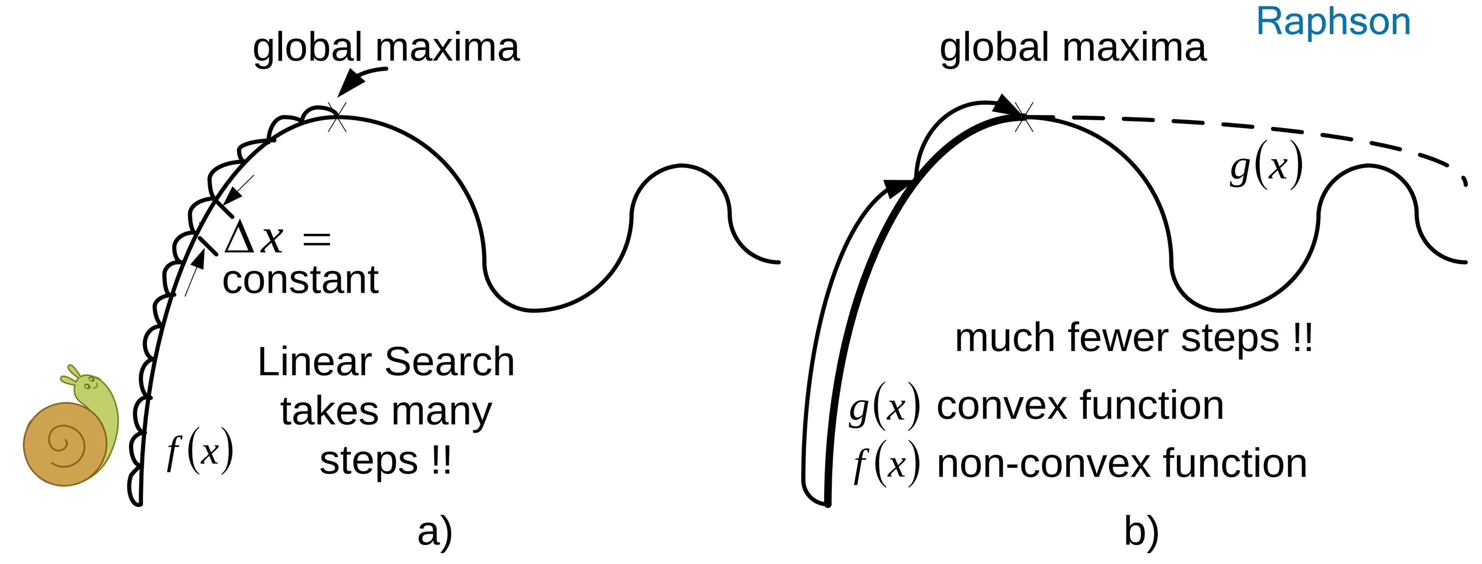
Given $\quad g(a) = \ln\left[f_X(x,a;n,m)\right]$

We prove

$$abs\left[\int_2^{\bar{a}} g''(a)da\right] \geq abs\left[\int_2^{a^*} g''(a)da\right]$$

Solution space is convex !!

Possible to solve with Newton-Raphson



Experimental Results

We report running time results, collected on an implementation using Python 2.7.11 (with Numpy library) executed on a PC with Core i7-3770 CPU running at 3.4GHz. In order to compare our results with the prior work [1], we mimic its test conditions of 1×10^{9} new observations for each test at an *a priori* BER of 5.8×10^{-10} and exact BER of 5.42×10^{-10}.

Running Time Results

Alg \ Bit Rate		2.5Gb/s	5.0Gb/s	8.0Gb/s
	New Observations	0.4s	0.2s	0.125s
Standard BMF-BD [9]	Search Time	948.6s	948.6s	948.6s
	Monte Carlo	2.9s	1.5s	0.9s
	Sample Savings	7.3×	7.3×	7.3×
	Total Savings	0.0031×	0.0015×	0.001×
Proposed	Search Time	0.001s	0.001s	0.001s
	Monte Carlo	6.88s	3.44s	2.15s
	Sample Savings	17.2×	17.2×	17.2×
	Total Savings	17.2×	17.1×	17.1×

Prior art provides only sample savings

The proposed method provides both sample and time savings

Conclusions

- Bayesian method achieves high accuracy with small sample sizes by leveraging on *a priori* knowledge
- We prove that the key objective function has a convex solution space regardless of the function being generally non-convex
- We demonstrate a negligible search time and 17× sample savings by replacing the linear search at the core of the Bayesian method with Newton-Raphson algorithm

References

[1] C. Fang, Q. Huang, F. Yang, X. Zeng, X. Li, and C. Gu, "Efficient Bit Error Rate Estimation for High-Speed Link by Bayesian Model Fusion," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2015*, pp. 1024-1029, IEEE, 2015.

[2] M. Mahmoud, "A New Proof of an Inequality for the Logarithm of the Gamma Function and Its Sharpness," *Turkish Journal of Analysis and Number Theory*, vol 2, no.4, pp. 147-151,2014.

[3] L. Barford, "Sequential Bayesian Bit Error Rate Measurement," *IEEE Transactions on Instrumentation and measurement*, vol. 53, no. 4, pp.947-954, 2004