ASYMPTOTIC OPTIMAL QUANTIZER DESIGN FOR DISTRIBUTED BAYESIAN ESTIMATION

Xia Li[⋆] *Jun Guo*[⋆] *Uri Rogers*[†] *Hao Chen*[⋆]

[⋆] Electrical and Computer Engineering, Boise State University
[†] Engineering, Eastern Washington University

ABSTRACT

In this paper, we address the optimal quantizer design problem for distributed Bayesian parameter estimation with one-bit quantization at local sensors. A performance limit obtained for any distributed parameter estimator with a known prior is adopted as a guidance for quantizer design. Aided by the performance limit, the optimal quantizer and a set of noisy observation models that achieve the performance limit are derived. Further, when the performance limit may not be achievable for some applications, we develop a near-optimal estimator which consists of a dithered noise and a single threshold quantizer. In the scenario where the parameter is Gaussian and signal-to-noise ratio is greater than -1.138 dB, we show that one can construct such an estimator that achieves approximately 99.65% of the performance limit.

***Index Terms*—Distributed Bayesian Estimation, One-bit Quantization, Quantizer Design, Cramer-Rao Lower Bound, Asymptotic Performance Limit**

I. INTRODUCTION

Distributed estimation is a classical research problem in wireless sensor network [1]–[3]. One of the typical network structures in distributed estimation is the parallel network, where a number of local sensors send their compressed observations to a fusion center (FC), and the FC makes an estimate. Due to constraints of sensor power and transmission bandwidth between local sensors and FC, the sensor observations are often quantized into one or a few bits based on their quantization rules. One-bit quantization is employed in the scenarios where the bandwidth and sensor power are severely limited [2]–[8].

Performance of a distributed estimator is largely determined by the quantizer at sensors and the estimator at the FC. One of the most widely used performance metrics for an estimator is its mean squared error (MSE). For any unbiased estimator, MSE is lower bounded by its Cramér-Rao lower bound (CRLB) [9]. Since the maximum likelihood (ML) and Maximum a Posteriori (MAP) estimation achieve the CRLB asymptotically [9], [10], this bound is often employed as a metric for performance evaluation and optimization in distributed estimation systems [3], [8]. Compared to centralized estimation settings where the unquantized data are directly accessible, distributed estimation suffers a logarithm rate loss in estimation performance with the use of some uniform dithered quantizers [4], [11]. To further improve estimation performance, the design of local quantizers with nonidentical and identical quantization rules are investigated, respectively in [6] and [12]. Quantizer design and its performance in distributed Bayesian estimation framework using Posterior Cramér-Rao lower bound (PCRLB) were discussed in [8]. However, the applicability of such approaches are rather limited as PCRLB cannot be achieved under some circumstances [10]. Overall, the problem of optimal quantizer design is shown to be extremely complicated in many applications [13], and the design of optimal distributed estimators is still an open problem in general except for a few very special cases. Some progress in addressing the optimization problem in distributed estimation were made where the sensor data are conditionally independent and identically distributed (i.i.d.). In [3], [7], the performance limit (PL) with identical one-bit quantizer was derived under the minimax criterion. Recently, the PL under Bayesian setting with a known prior distribution was derived in [14]. Compared to the naïve full precision PL which assumes the unquantized sensor data are available at the FC, the proposed distributed PLs are shown to be comparatively much tighter in most meaningful signal-to-noise ratio (SNR) regions [3], [14].

In this paper, we consider the problem of estimating a random scalar parameter θ from N i.i.d. noisy one-bit quantized distributed sensor messages with identical quantizers. Note, while nonidentical quantizers may offer better performance, they are sensitive to system changes, complicated to design, optimize, and implement in practical applications. We show in this work that, surprisingly, in some practical applications, the Bayesian PL in [14] may actually be achieved or nearly achieved by a dithered Sign quantizer, i.e., a dithered threshold quantizer with threshold 0 and the dithering noise follows a certain distribution. We derive the pdf of optimal dithered noise as well as the near-optimal noise, if the former is hard to obtain. For example, when the parameter θ is a Gaussian random variable and the observations are contaminated by additive Gaussian noises, we show that the performance achieved by a simple dithered Sign quantizer is very close to the PL.

II. DISTRIBUTED BAYESIAN ESTIMATION WITH ONE-BIT QUANTIZATION

We investigate a distributed estimation problem in wireless sensor networks where the goal is to estimate a random scalar parameter θ from a set of N conditionally i.i.d. sensor observations $\boldsymbol{X} = [X_1, X_2, \cdots, X_N]$ with X_i be the observation at sensor i, $i \in \{1, 2, \ldots, N\}$ such that

$$f(\mathbf{X}|\theta) = \prod_{i=1}^{N} f(X_i|\theta),$$

where $f(\mathbf{X}|\theta)$ and $f(X_i|\theta)$ are known probability density functions (pdfs), and θ has a prior pdf $p_\theta(\theta)$ defined on an open support $\Theta = \{\theta \in \Re : p_\theta(\theta) > 0\} = (a, b)$, where a, b (can be $-\infty$ or ∞) are known lower and upper bounds for θ, respectively. A widely considered case of such model is the location estimation problem, where

$$X_i = \theta + W_i, \tag{1}$$

and W_i is the additive i.i.d. zero mean observation noise with pdf $f_W(\cdot)$. We assume the observed data X_i is quantized to U_i, a one-bit quantized message (0 or 1) with quantization rule $\gamma_i : \Re \rightarrow [0,1]$, such that

$$\gamma_i(X_i) = \Pr(U_i = 1|X_i).$$

We consider all possible local quantization rules similar to those investigated in [3] by letting $\gamma_i(x) \in [0,1]$ be a real number between 0 and 1 and assume the channels between the sensors and the FC are noiseless such that FC receives U_i without error. Moreover, notice the advantage of identical quantization rules, we consider only identical quantizers such that $\gamma = \gamma_1 = \gamma_2 = \cdots = \gamma_N$.

In the literature, the Sign quantizer $\gamma_S(x)$ is widely employed with the output be the sign of observation such that

$$\gamma_S(x) = \begin{cases} 1 & X_i \geq 0 \\ 0 & X_i < 0. \end{cases}$$

While the Sign quantizer is simple to implement, its performance is not always satisfactory, especially when SNR is high [3], [7]. In practice, dithering is often used in quantization [15]–[17]. Under the Sign quantization scheme, if a dithering noise $W_{d,i}$ is added so that $U_i = 1$ when $X_i + W_{d,i} > 0$, the resulting quantization rule can be described as a probabilistic quantizer $\gamma_i(X_i) = \Pr(X_i + W_{d,i} > 0) = \Pr(W_{d,i} > -X_i) = 1 - F_{W_{d,i}}(-X_i)$, where $F_{W_{d,i}}(\cdot)$ is the cumulative distribution function (cdf) of $W_{d,i}$. Therefore, a monotonic non-decreasing quantizer γ_i can be implemented by a dithered Sign quantizer with a suitable noise $W_{d,i}$ with pdf $f_{W_{d,i}}(x) = \frac{dF_{W_{d,i}}(x)}{dx} = \frac{d(1-\gamma_i(-x))}{dx} = \gamma_i'(-x)$ and vice versa.

As the estimation made at the FC is based on $\{U_1, U_2, \ldots, U_N\}$, the estimation performance is completely determined by the conditional probability mass function (pmf) $\kappa(\theta) = \Pr(U_i = 1|\theta)$ given by

$$\kappa(\theta) = E_{x|\theta}\left[\Pr(U_i = 1|X_i)\right] = \int_x \left[\gamma(x) f(x|\theta)\right] dx,$$

which is a function of the quantization rule $\gamma(x)$ and observation model $f(\mathbf{X}|\theta)$. To ensure the existence of an consistent estimator, $\kappa(\theta)$ has to be unique for any θ, i.e., $\kappa(\theta_1) \neq \kappa(\theta_2)$, $\forall \theta_1 \neq \theta_2$.

II-A. Asymptotic Estimation Performance and the Performance Limit of a Distributed Bayesian Estimator

When the size of sensor network is reasonably large, the normalized asymptotic MSE $\epsilon(\theta, \kappa, f) = N \cdot \text{MSE}$ for efficient Bayesian estimators (MLE, MAP are asymptotic efficient) is given by [10]

$$\begin{aligned}\epsilon(\kappa, f) &\approx N \cdot E_\theta\left(\text{CRLB}(\theta, \kappa, f)\right) \\ &\approx E_\theta\left(\frac{1}{I_i(\theta, \kappa, f)}\right) \quad (\text{as } N \rightarrow \infty) \\ &= \int_{-\infty}^{+\infty} \left[p_\theta(\theta) \frac{1}{I_i(\theta, \kappa, f)}\right] d\theta \end{aligned} \tag{2}$$

where $I_i(\theta, \kappa, f) = \frac{\left(\frac{d}{d\theta}\kappa(\theta)\right)^2}{\kappa(\theta)(1-\kappa(\theta))}$ is the Fisher information (FI) at sensor i with a differentiable κ on θ [3]. Among all possible observation models, the best performance is achieved with the perfect observation where each sensor observes $X_i = \theta$ in a noiseless, deterministic manner. It was shown that if $\int_a^b p_\theta(\theta)^{\frac{1}{3}} d\theta < \infty$, the asymptotic PL ϵ_{PL} for any one-bit identical quantizer and any arbitrary observation model is [14]

$$\epsilon_{PL} = \frac{\left(\int_a^b p_\theta(\theta)^{\frac{1}{3}} d\theta\right)^3}{\pi^2} \tag{3}$$

with $\epsilon \geq \epsilon_{PL}$ for any distributed estimators regardless of its observation model. This PL can be achieved by applying a probabilistic quantizer

$$\kappa_o(x) = \gamma_o(x) = \frac{1}{2}\left(1 - \cos\left(\pi \frac{\int_a^\theta p_\theta(t)^{\frac{1}{3}} dt}{\int_a^b p_\theta(\theta)^{\frac{1}{3}} d\theta} x\right)\right) \tag{4}$$

for the perfect observation case. By flipping the sensor outputs from 0 to 1 or 1 to 0, it can be shown that the same estimation performance is also achievable by the probabilistic quantizer $1 - \gamma_o(x)$.

III. OPTIMAL QUANTIZER DESIGNS

Since the case of perfect observation rarely exists in practical applications, we consider the cases where the observations are contaminated by additive noise, and try to design the optimal Bayesian estimator. Instead of comparing the estimation performance to the naïve full precision bound (which is often too loose and not achievable), we rely on the Bayesian PL to determine the optimality of the estimator. With Bayesian PL serves as a benchmark for all observation models and all possible quantizers, if one can find a quantizer that achieves the PL, then no other quantizer can perform better, and the optimal design problem is solved. If such optimal quantizer is difficult to obtain, the problem is considered to be almost solved if one can find a quantizer that has a negligible gap between its performance and the PL. In the following, we discuss the observation models and focus especially on the dithered Sign quantizer design to achieve or approach the optimal performance (PL), whenever is possible.

III-A. Achieving the PL with a Sign Quantizer

First, we consider a special case of the localization model (1) when the observation noise $W_{d,i} \sim f_W = f_{W_o}$, where f_{W_o} is given by

$$1 - F_{W_o}(-\theta) = \kappa_o(\theta) \leftrightarrow f_{W_o}(\theta) = \frac{dF_{W_o}(\cdot)}{d\theta} = \kappa_o'(-\theta), \tag{5}$$

by applying a Sign quantizer at the sensors, the resulting conditional probability is thus $\kappa(\theta) = P(U_i = 1|\theta) = P(\theta + W_i > 0) = 1 - F_W(-\theta) = 1 - F_{W_o}(-\theta) = \kappa_o(\theta)$, where $\kappa_o(\theta)$ denotes the optimal pmf achieved under perfect observation $X_i = \theta$ as in equation (4). Therefore, when the observation noise follows $f_W(x) = f_{W_o}(x) = \kappa_o'(-x)$, the PL is achieved with a Sign quantizer.

III-B. Achieving the PL with a Dithered Sign Quantizer

Next, we analyze the more general case where the pdf of observation noise $f_W(\cdot)$ is not the same as $f_{W_o}(\cdot)$. Recall that *any* monotonic nondecreasing $\gamma(x)$ can be implemented equivalently as a dithered Sign quantizer with a dithering noise $W_{d,i}$, where the corresponding pdf is $f_{W_d}(x) = \gamma'(-x)$. With dithering noise, the dithered observation becomes $\tilde{X}_i = X_i + W_{d,i} =$

$\theta + W_i + W_{d,i} = \theta + \tilde{W}_i$, where $\tilde{W}_i = W_i + W_{d,i}$. Therefore, it is still possible to achieve the PL as long as $f_{\tilde{W}}$, the pdf of $\tilde{W}_i$ is the same as f_{W_o}. Using the relationship between pdfs of two summed random variables, such equivalence can be described as, $f_{\tilde{W}} = f_W(\cdot) * f_{W_d}(\cdot) = f_{W_o}(\cdot)$, where $f_{W_d}(\cdot)$ is the pdf of additive dithered noise $W_{d,i}$, and $*$ denotes the convolution operator. Taking the Fourier transformation on both sides of the equation, the resulting $\mathcal{F}_W(\omega)\mathcal{F}_{W_d}(\omega) = \mathcal{F}_{W_o}(\omega)$ leads to $f_{W_d}(x) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}_{W_o}(\omega)}{\mathcal{F}_W(\omega)}\right)$, where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier transformation (FT) and the Inverse Fourier transformation (IFT), respectively. Since $f_W(\cdot)$ and $f_{W_o}(\cdot)$ are pdfs, i.e., $\mathcal{F}_W(0) = \mathcal{F}_{W_o}(0) = 1$, and the dithered noise is therefore maintains $\mathcal{F}_{W_d}(0) = \int_{-\infty}^{+\infty} f_{W_d}(x)\,dx = 1$. For $f_{W_d}(\cdot)$ to be a valid pdf, the only constraint is $f_{W_d}(\cdot) \geq 0$. If such a $\mathcal{F}^{-1}\left(\frac{\mathcal{F}_{W_o}(\omega)}{\mathcal{F}_W(\omega)}\right)$ is valid, then the optimization problem is completely solved.

However, this analytical approach cannot be implemented in many practical applications. For example, in the case when $\mathcal{F}^{-1}\left(\frac{\mathcal{F}_{W_o}(\cdot)}{\mathcal{F}_W(\cdot)}\right)$ is not a valid pdf, the dithering noise $f_{W_d}(\cdot)$ which achieves PL does not exist. Moreover, it is often too difficult to obtain $\mathcal{F}_{W_o}(\cdot)$ or $\mathcal{F}_{W_i}(\cdot)$, e.g., when $f_{W_o}(\cdot)$ is a complicated function, its FT may not be that easily calculated as we will show in Section IV. Fortunately, in some cases, design approaches exploiting the structure of the observation models can be employed to obtain sub-optimal quantizers which perform closely to the PL.

IV. DISTRIBUTED ESTIMATION OF A GAUSSIAN RANDOM PARAMETER

We now consider the important case when the parameter θ is Gaussian distributed. Without loss of generality, we assume $\theta \sim \mathcal{N}(0,1)$ with zero mean and unit variance. Consider the localization model (1) where the observation noise is also Gaussian with $W_i \sim \mathcal{N}\left(0, \sigma_w^2\right)$, we aim to design the optimal PL achieving quantizer (if possible) as well as the near-optimal quantizer with a performance close to the PL.

Note the existence of $p_\theta(\theta)^{\frac{1}{3}}$ integral satisfies the condition for the derived PL (3) in [14]. We first consider the perfect observation case where $\sigma_w^2 = 0 \leftrightarrow X = \theta$, establish the PL ε_g (e.g. subscript g) as the performance benchmark, and derive the corresponding optimal $\kappa_o(\theta)$ as well as the optimal dithering noise pdf $f_{W_o}(\cdot)$ that can be used to implement a dithered Sign quantizer to achieve ε_g.

Corollary 1. *[Optimal Dithered Noise in Distributed Bayesian Estimation with Gaussian Prior and Identical Sign Quantizer]*

For distributed Bayesian estimation with local sensors observation model (1)*, if the parameter θ has Gaussian pdf $p_\theta(\theta) = \Phi(\theta)$, the best performance for distributed Bayesian estimation is $\varepsilon_g = \frac{6\sqrt{3}}{\pi}$. When the Sign quantizer is applied for each sensor observation with the zero observation noise ($X_i = \theta$, $W_i = 0$), the optimal dithered noise $W_{d,i}$ (added to X_i or θ) that achieves the performance limit ε_g has the following distribution*

$$f_{W_{o,g}}(\theta) = \left[\frac{\pi}{2\sqrt{3}} \sin\left(\pi Q\left(\frac{\theta}{\sqrt{3}}\right)\right)\right]\Phi\left(\frac{\theta}{\sqrt{3}}\right), \qquad (6)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{+\infty}\exp\left(-\frac{t^2}{2}\right)dt$ and $\Phi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ is the pdf of standard Gaussian.

The proof of corollary 1 is given in VII.

Remark 2. When the prior on θ is standard Gaussian, formula (6) suggests the optimal dithered noise also has some characteristics of a Gaussian distribution, where $\left[\left(\frac{\pi}{2\sqrt{3}}\right)\sin\left(\pi Q\left(\frac{\theta}{\sqrt{3}}\right)\right)\right]$ acts as a scaling factor on Gaussian pdf $\Phi\left(\frac{\theta}{\sqrt{3}}\right)$.

Remark 3. In a different perspective, if the observation noise satisfies $W_i \sim f_{W_{o,g}}(\cdot)$ in (1), then ε_g is obtained via the use of a Sign quantizer. For the general case where $W_i \nsim f_{W_{o,g}}(\cdot)$, ε_g is still achievable if there exists a dithered noise $W_{d,i}$ satisfies $W_i + W_{d,i} \sim f_{W_{o,g}}(\cdot)$, or $f_W(\cdot) * f_{W_d}(\cdot) = f_{W_{o,g}}(\cdot)$. However, such $f_{W_d}(\cdot)$ may not exist due to pdf legitimate requirement on $f_{W_d}(\cdot)$. The validity of $f_{W_d}(\cdot)$ is not easy to examine due to the difficulty in calculating FT of $f_{W_{o,g}}(\cdot)$ and IFT of $f_{W_d}(\cdot) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}_{W_o}(\cdot)}{\mathcal{F}_W(\cdot)}\right)$.

Therefore, for the general case where the observation noise variance $\sigma_w^2 \neq 0$, we consider a special class of dithering noise where $W_{d,i} \sim \mathcal{N}\left(0, \sigma_d^2\right)$. Since the sum of two independent Gaussian is still Gaussian, we have $\tilde{W}_i = W_i + W_{d,i} \sim \mathcal{N}\left(0, \sigma_{\tilde{w}}^2\right)$ where $\sigma_{\tilde{w}}^2 = \sigma_w^2 + \sigma_d^2$, and the optimization problem reduces to find the optimal $\sigma_{\tilde{w}}^2 = \sigma_o^2 \geq \sigma_w^2$ and the optimal dithering noise $W_{d,i} \sim \mathcal{N}\left(0, \sigma_o^2 - \sigma_w^2\right)$ that minimizes the normalized MSE.

When $X = \theta + \tilde{W}$, the FI per sensor is given by $I(\theta) = \frac{\left(\Phi\left(\frac{\theta}{\sigma_{\tilde{w}}}\right)\right)^2}{\sigma_{\tilde{w}}^2 Q\left(\frac{\theta}{\sigma_{\tilde{w}}}\right)Q\left(-\frac{\theta}{\sigma_{\tilde{w}}}\right)}$ [3] and the asymptotic MSE (2) is therefore

$$\epsilon\left(\sigma_{\tilde{w}}^2\right) = \int_{-\infty}^{+\infty} \frac{\sigma_{\tilde{w}}^2 \Phi(\theta)\, Q\left(\frac{\theta}{\sigma_{\tilde{w}}}\right) Q\left(-\frac{\theta}{\sigma_{\tilde{w}}}\right)}{\left(\Phi\left(\frac{\theta}{\sigma_{\tilde{w}}}\right)\right)^2} d\theta.$$

With a numerical search, we found that with a optimal $\sigma_{\tilde{w}}^2 = \sigma_o^2 = 1.14^2$, a minimum normalized MSE $\epsilon\left(\sigma_o^2\right)$ is achieved with $\epsilon\left(\sigma_o^2\right) = 3.319$. Compared to the PL ε_g, the optimal Gaussian dithered quantizer is about $\frac{3.308}{3.319} = 99.65\%$ efficient as the PL. Since PL is the performance upper bound for all possible observation models, the Gaussian dithered Sign quantizer with noise $\mathcal{N}\left(0, \sigma_o^2 - \sigma_w^2\right)$ is therefore "almost" asymptotically optimal (with an efficiency at least 99.65% of the optimal one), as long as $\sigma_w^2 \leq \sigma_o^2 = 1.14^2$ or equivalently, the SNR of the sensor observations is greater than $10\log\frac{1}{1.14^2} = -1.138$dB. Here, we denote σ_o^2 as the near-optimal Gaussian noise variance. Thus, the asymptotic optimal distributed design problem is solved, and the corresponding asymptotic optimal quantizer is a dithered quantizer with Gaussian dithering noise $\mathcal{N}\left(0, 1.14^2 - \sigma_w^2\right)$.

In this paper, we solved the asymptotic optimal distributed estimator design problem for a wide range of SNRs between $10\log\frac{1}{\sigma_o^2}$dB $= -1.138$dB to ∞dB. When SNR is lower than-1.138dB, our numerical simulation results suggest that the optimal dithering noise is 0.

V. EXPERIMENTAL RESULTS

V-A. Near-Optimal Gaussian Noise under Distributed Bayesian

When the parameter θ is Gaussian distributed with $\theta \sim \mathcal{N}(0,1)$, we now validate our observation of the Gaussian characteristic in $f_{W_{o,g}}(\theta)$ by showing the pdf comparison between the optimal dithered noise $f_{W_{o,g}}(\theta)$ and the near-optimal Gaussian noise $\mathcal{N}\left(0, \sigma_o^2\right)$ in Figure 1(b). The similarity between these two pdfs further indicates the near-optimal performance of the additive Gaussian noise $\mathcal{N}\left(0, \sigma_o^2\right)$.

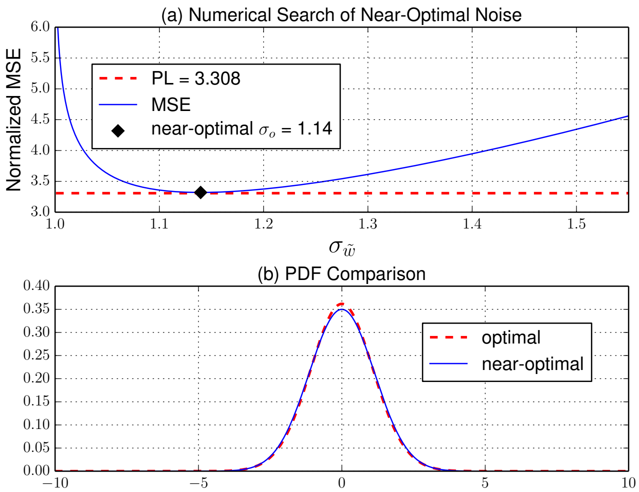

Fig. 1. (a) Numerical search of near-optimal Gaussian noise: $\mathcal{N}\left(0, 1.14^2\right)$, MSE $= 3.319$. (b) PDF comparison between optimal noise $f_{W_{o,g}}(\theta)$ and near-optimal Gaussian noise $\mathcal{N}\left(0, 1.14^2\right)$.

V-B. Monte Carlo Simulation of Asymptotic MSE under MLE and MAP

Next, we perform Monte Carlo simulations of MSE under MLE and MAP for different Gaussian dithered Sign quantizer designs. When the dithering observation is contaminated by $\tilde{W} \sim \mathcal{N}\left(0, \sigma_{\tilde{w}}^2\right)$, MLE of θ is given by

$$\hat{\theta} = -\sigma_{\tilde{w}} Q^{-1}\left(\frac{\sum_{i=1}^{i=N} U_i}{N}\right),$$

and the MAP of θ is determined by

$$\begin{aligned}\hat{\theta} :=& \arg\max_{\theta}[\left(\sum_{i=1}^{i=N} U_i\right)\log\left(Q\left(-\frac{\theta}{\sigma_{\tilde{w}}}\right)\right) \\ &+ \left(\left(N - \sum_{i=1}^{i=N} U_i\right)\log\left(Q\left(\frac{\theta}{\sigma_{\tilde{w}}}\right)\right)\right) + \log\left(\Phi\left(\theta\right)\right)],\end{aligned}$$

where Q^{-1} denotes the inverse of complementary cdf of Gaussian distribution [9].

Figure 2 demonstrates the effectiveness of the proposed near-optimal dithered quantizer. When the system SNR $= 0.915$dB ($\sigma_w^2 = 0.9^2$), the estimation performance diverges with a Sign quantizer without dithering noise. The near-optimal performance is achieved if we apply a dithered Sign quantizer with dithering Gaussian noise $W_{d,i} \sim \mathcal{N}\left(0, 1.14^2 - 0.9^2\right)$. For situations where SNR < -1.138dB ($\sigma_w^2 > 1.14^2$), the PL cannot be reached by a Sign quantizer. It is noticed that the MAP outperforms MLE due to its use of prior distribution of θ. Also, when number of sensors N is small, the performance of both MLE and MAP are better than the PL, due to PL is an asymptotic performance bound for sufficiently large N. For the case when sensor number is small (non-asymptotic), the Bayesian PL is not tight.

VI. Conclusion

We considered the problem of parameter estimation in distributed system with i.i.d. sensor observations and independent identical quantizers under Bayesian criterion where the prior probability density function of the parameter is known. We developed design approaches for optimal and near-optimal distributed estimation systems when the sensors observations are quantized to one-bit messages. For the case where the signals are contaminated by additive Gaussian noise, we developed a near-optimal dithered quantizer with a performance close to the PL, and validated our result via Monte Carlo simulations.

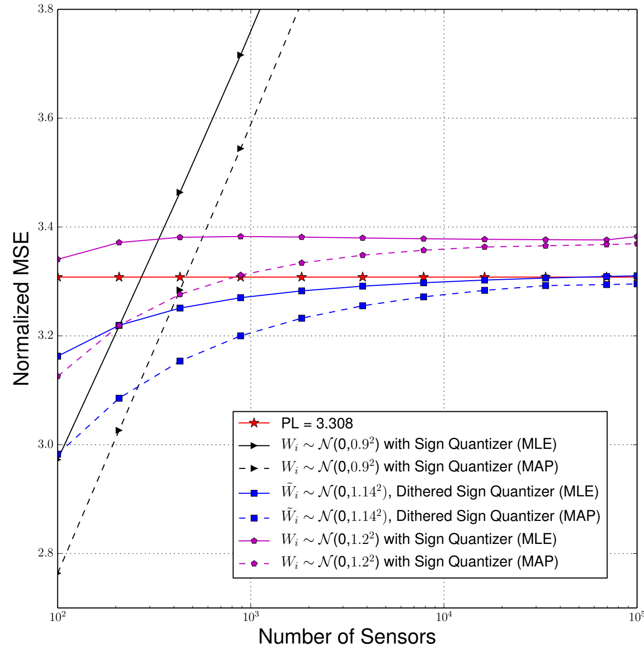

Fig. 2. Monte Carlo simulation of MSE under MLE and MAP: near-optimal $\mathcal{N}\left(0, 1.14^2\right)$.

VII. Proof of Corollary

From (3), the best achievable performance for distributed Bayesian estimation with standard Gaussian observation $\theta \sim \mathcal{N}(0,1)$ is $\varepsilon_g = \frac{6\sqrt{3}}{\pi}$, which is achieved under noise free condition $X = \theta$. Under such setting, let $\kappa(\theta) = \gamma(\theta) = \Pr(U_i = 1|\theta) = \frac{1+\sin(g(\theta))}{2}$ with $g(\theta) : (a,b) \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ [14], and the optimal dithered noise $f_{W_o}(\cdot)$ from (5) is thus

$$f_{W_o}(\theta) = \gamma_o'(-\theta) = \frac{g'(\theta)\cos\left(g\left(-\theta\right)\right)}{2}. \tag{7}$$

Given perfect Gaussian observation model, the gradient of $g(\theta)$ [14] after applying the optimal C_o is

$$g'(\theta) = \frac{\pi}{3^{\frac{1}{2}}\left(2\pi\right)^{\frac{1}{2}}}\exp\left(\frac{\theta}{6}\right) = \frac{\pi}{\sqrt{3}}\Phi\left(\frac{\theta}{\sqrt{3}}\right), \tag{8}$$

and the integration of (8) gives

$$g(\theta) = \int_{-\infty}^{\theta} g'(t)\, dt + g(a) = \pi Q\left(-\frac{\theta}{\sqrt{3}}\right) - \frac{\pi}{2}. \tag{9}$$

Thus, Corollary 1 is proved by substituting (8) and (9) in (7).

VIII. REFERENCES

[1] J. Gubner, “Distributed estimation and quantization,” *Information Theory, IEEE Transactions on*, vol. 39, no. 4, pp. 1456–1459, Jul 1993.

[2] W.-M. Lam and A. R. Reibman, “Design of quantizers for decentralized estimation systems,” *IEEE Transactions on communications*, vol. 41, no. 11, pp. 1602–1605, 1993.

[3] H. Chen and P. Varshney, “Performance limit for distributed estimation systems with identical one-bit quantizers,” *Signal Processing, IEEE Transactions on*, vol. 58, no. 1, pp. 466–471, Jan 2010.

[4] O. Dabeer and A. Karnik, “Signal parameter estimation using 1-bit dithered quantization,” *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5389–5405, 2006.

[5] V. Megalooikonomou and Y. Yesha, “Quantizer design for distributed estimation with communication constraints and unknown observation statistics,” *IEEE Transactions on Communications*, vol. 48, no. 2, pp. 181–184, 2000.

[6] Z.-Q. Luo, “An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network,” *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 4, pp. 735–744, April 2005.

[7] S. Kar, H. Chen, and P. K. Varshney, “Optimal identical binary quantizer design for distributed estimation,” *Signal Processing, IEEE Transactions on*, vol. 60, no. 7, pp. 3896–3901, 2012.

[8] A. Vempaty, H. He, B. Chen, and P. Varshney, “On quantizer design for distributed bayesian estimation in sensor networks,” *Signal Processing, IEEE Transactions on*, vol. 62, no. 20, pp. 5359–5369, Oct 2014.

[9] S. M. Kay, “Fundamentals of statistical signal processing, volume i: Estimation theory (v. 1),” in *Technometrics*, 1995, no. 37.

[10] H. L. Van Trees, *Detection, estimation, and modulation theory*. John Wiley & Sons, 2004.

[11] O. Dabeer and E. Masry, “Multivariate signal parameter estimation under dependent noise from 1-bit dithered quantized data,” *Information Theory, IEEE Transactions on*, vol. 54, no. 4, pp. 1637–1654, 2008.

[12] A. Ribeiro and G. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks-part ii: unknown probability density function,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 7, pp. 2784–2796, July 2006.

[13] P. Venkitasubramaniam, L. Tong, and A. Swami, “Quantization for maximin are in distributed estimation,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 7, pp. 3596–3605, July 2007.

[14] X. Li, J. Guo, U. Rogers, and H. Chen, “The performance limit for distributed bayesian estimation with identical One-Bit quantizers,” in *2015 IEEE Signal Processing and Signal Processing Education Workshop (SP/SPE) (SPW2015)*, Salt Lake City, USA, Aug. 2015, pp. 331–336.

[15] L. Roberts, “Picture coding using pseudo-random noise,” *Information Theory, IRE Transactions on*, vol. 8, no. 2, pp. 145–154, February 1962.

[16] R. Gray and J. Stockham, T.G., “Dithered quantizers,” *Information Theory, IEEE Transactions on*, vol. 39, no. 3, pp. 805–812, May 1993.

[17] R. Wannamaker, S. Lipshitz, J. Vanderkooy, and J. Wright, “A theory of nonsubtractive dither,” *Signal Processing, IEEE Transactions on*, vol. 48, no. 2, pp. 499–516, Feb 2000.