

Article

Analysis of Polynomial Nonlinearity Based on Measures of Nonlinearity Algorithms

Mahendra Mallick ^{1,*},†,‡ and Xiaoqing Tian ² 

¹ Independent Consultant, Anacortes, WA 98221, USA

² School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China; tianxiaoqing2017@stu.xjtu.edu.cn

* Correspondence: mmallick.us@gmail.com

† Current address: Anacortes, WA 98221, USA.

‡ The corresponding author contributed most to this work. M. Mallick: Algorithm and Software Development, Analysis of Algorithms and Results, Manuscript Preparation. X. Tian: Software Development, Analysis of Algorithms and Results, Manuscript Preparation.

Received: 31 March 2020; Accepted: 10 June 2020; Published: 17 June 2020



Abstract: We consider measures of nonlinearity (MoNs) of a polynomial curve in two-dimensions (2D), as previously studied in our Fusion 2010 and 2019 ICCAIS papers. Our previous work calculated curvature measures of nonlinearity (MoNs) using (i) extrinsic curvature, (ii) Bates and Watts parameter-effects curvature, and (iii) direct parameter-effects curvature. In this paper, we have introduced the computation and analysis of a number of new MoNs, including Beale's MoN, Linssen's MoN, Li's MoN, and the MoN of Straka, Duník, and Šimandl. Our results show that all of the MoNs studied follow the same type of variation as a function of the independent variable and the power of the polynomial. Secondly, theoretical analysis and numerical results show that the logarithm of the mean square error (MSE) is an affine function of the logarithm of the MoN for each type of MoN. This implies that, when the MoN increases, the MSE increases. We have presented an up-to-date review of various MoNs in the context of non-linear parameter estimation and non-linear filtering. The MoNs studied here can be used to compute MoN in non-linear filtering problems.

Keywords: polynomial curve in 2D; measures of nonlinearity (MoNs); extrinsic curvature; Beale's MoN; Linssen's MoN; Bates and Watts parameter-effects curvature; direct parameter-effects curvature; Li's MoN; MoN of Straka, Duník, and Šimandl; maximum likelihood estimator (MLE); Cramér-Rao lower bound (CRLB)

1. Introduction

The Kalman filter (KF) [1–4] is an optimal estimator (in the minimum mean square error (MMSE) sense) for a filtering problem with linear dynamic and measurement models with additive Gaussian noise. However, many real-world filtering problems are non-linear due to nonlinearity in the dynamic and measurement models. Common real-world non-linear filtering (NLF) problems are bearing-only filtering [5–8], ground moving target indicator (GMTI) filtering [9], passive angle-only filtering in three-dimensions (3D) using an infrared search and track sensor [10–12], etc.

In the early stages of NLF, the extended Kalman filter (EKF) [1–4] was widely used. It was observed in some problems, e.g., falling of a body in earth's atmosphere with high velocity [13,14] and bearing-only filtering [5,7,8] that the EKF performs poorly due to linearization. The high degree of nonlinearity in these problems was the attributed cause for the poor performance of the problem without a quantitative measure of nonlinearity (MoN). To overcome the poor accuracy and convergence problems of the EKF, a number of improved approximate non-linear filters, such as the unscented

Kalman filter (UKF) [14,15], cubature KF (CKF) [16], and particle filter (PF) [8,17] have been proposed during the last two decades.

It is important to address the following questions for NLF problems:

1. Is it possible to find a quantitative MoN for a nonlinear filtering problem?
2. Can we establish a correspondence between the MoN of a NLF problem and the performance of a filtering algorithm?
3. Can we show that the UKF, CKF, or PF gives better results than the EKF, when the degree of nonlinearity (DoN) is high?

Remark 1. *In this paper we consider a parameter estimation problem with polynomial nonlinearity. We hope that insights and results from this analysis would encourage further study of MoN in NLF problems. Next, we describe some historical developments in the field of parameter estimation and NLF.*

Beale in his pioneering work [18] proposed four MoNs for the static non-random parameter estimation problem. Two MoNs were empirical and two were theoretical. Guttman and Meeter [19] and Linssen [20] observed that Beale's method gives lower MoN for highly non-linear problems and proposed a modified MoN. Using differential geometry based curvature measures, Bates and Watts [21,22] and Goldberg et al. [23] extended Beale's work and developed curvature measures of nonlinearity (CMoN) for the static non-random parameter estimation problem. Bates and Watts formulated two CMoN, the parameter-effects curvature and intrinsic curvature [21,24–26].

In [27], we first extended the method of Bates and Watts to the non-linear filtering problem with unattended ground sensor (UGS) to calculate CMoN. Next, we computed the parameter-effects curvature and intrinsic curvature for the bearing-only filtering (BOF) problem [28–31], GMTI filtering problem [30,32,33], video tracking problem [34], and polynomial nonlinearity [35].

In our previous work [35], we considered a polynomial curve in two-dimensions (2D) and calculated CMoN using differential geometry (e.g., extrinsic curvature) [36–38], Bates and Watts parameter-effects curvature [21,25,26], and direct parameter-effects curvature [29]. The computation of these curvatures requires the Jacobian and Hessian of the measurement function [2] evaluated at the true or estimated parameter. The extrinsic curvature uses the true parameter, whereas the other two CMoN use the estimated parameter.

In [35], we obtained the maximum likelihood (ML) estimate [2,39] of the parameter x while using a vector measurement by numerical minimization. In [40], we derived analytic expressions for the ML estimator (MLE) [2,39] and associated variance using a vector measurement. This approach is simple and efficient, since it does not require numerical minimization. We also showed through Monte Carlo simulations in [40] that the variance of the MLE and the Cramér-Rao lower bound (CRLB) [2,41] are nearly the same for different powers of x . We also found that the bias error was small and the mean square error (MSE) [2] was close to the CRLB and variance of the MLE. Our numerical results showed that the average normalized estimation error squared (ANEES) [42] was within the 99% confidence interval most of the time. Hence, the variance of the MLE was in agreement with the estimation error.

Li constructed a combined non-linear function while using the non-linear time evolution function and measurement function in a discrete-time nonlinear filtering problem, and he proposed a global MoN at each measurement time [43]. This MoN minimizes the mean square distance between the combined non-linear function and the set of all affine functions with the same dimension at each measurement time. An un-normalized MoN and a normalized MoN were proposed in [43]. These MoNs can also be unconditional or conditional. The normalized MoN lies in the interval $[0, 1]$. A journal version of the paper with enhancements was published in [44].

The normalized MoN that was proposed in [43] was calculated for non-linear filtering problems, including one with the nearly constant turn motion and a non-linear measurement model [45], a video tracking problem using PF [46], and a hypersonic entry vehicle state estimation problem [47]. In these cases, the normalized MoN were rather low. In [33], we compared the normalized MoN for

the BOF and GMTI filtering problems. Contrary to our expectation, we found that the GMTI filtering problem had a higher conditional normalized MoN than that of the BOF problem in the examples that we investigated.

Using the current mean (e.g., predicted mean) and associated covariance, Duník et al. [48] generate a number of sample points (e.g., sigma points using unscented transform [14]) and transform these points using a non-linear function (e.g., non-linear measurement function or time evolution function). Subsequently, they try to predict the transformed points using a linear transformation and estimate the parameter of the transformation using linear weighted least squares (WLS) [39]. They use the cost function of the WLS evaluated at the estimated parameter as a local MoN.

In [35], we showed analytically and through Monte Carlo simulations that affine mappings with positive slopes exist among the logarithm of the extrinsic curvature, Bates and Watts parameter-effects curvature, direct parameter-effects curvature, MSE, and CRLB. For completeness, we have included these key results from [35] in Section 4. New contributions in this paper include the computation and analysis of following MoNs:

- Beale's MoN [18],
- Least squares based Beale's MoN,
- Linssen's MoN [20],
- Least squares based Linssen's MoN,
- Li's MoN [43,44], and
- MoN of Straka, Duník, and Šimandl [48,49].

It is not possible to derive a mapping analytically between the logarithm of Beale's MoN, Linssen's MoN, Li's MoN, MoN of Straka, Duník, and Šimandl, and the logarithm of the MSE. The numerical results from Monte Carlo simulations also show that affine mappings with positive slopes exist among the logarithm of the MSE and the logarithm of two of these MoNs.

The paper is organized, as follows. Section 2.1 describes the measurement model for polynomial nonlinearity. The MLE for parameter estimation and CRLB using polynomial nonlinearity and a vector measurement is presented in Section 2. Section 3 presents different types of MoN, such as extrinsic curvature based on differential geometry, Beale's MoN, Linssen's MoN, Bates and Watts parameter-effects curvature, direct parameter-effects curvature, Li's MoN, and MoN of Straka, et al. Section 4 discusses mappings among logarithms of extrinsic curvature, parameter-effects curvature, CRLB, and MSE. Section 5 presents the numerical simulation and results. Finally, Section 6 summarizes our contribution and concludes with future work.

Notation Convention: For clarity, we use italics to denote scalar quantities and boldface for vectors and matrices. A lower or upper case Roman letter represents a name (e.g., "s" for "sensor", "RMS" for "root mean square", etc.). We use "==" to define a quantity and $\mathbf{A}'$ denotes the transpose of the vector or matrix $\mathbf{A}$. The n -dimensional identity matrix, m -dimensional null matrix, and $m \times n$ null matrix are denoted by $\mathbf{I}_n$, $\mathbf{0}_m$, and $\mathbf{0}_{m \times n}$, respectively.

2. MLE Parameter Estimation and CRLB

2.1. Measurement Model

We studied CMoN of a polynomial smooth scalar function h of a non-random variable x in [35], where

$$h(x) = ax^n, \quad (1)$$

and a is a non-zero scalar. In scenarios considered, $x > 0$ and $n = 2, 3, 4, 5$.

Remark 2. For MoN of other forms of nonlinearity, such as the bearing-only [27], GMTI [32], and video filtering [34] problems in radar communities, we shall discuss in detail in the end of Section 3.

The measurement model for the polynomial function is given by

$$z_i = h(x) + v_i, \quad i = 1, \dots, N, \quad (2)$$

where v_i is a zero-mean white Gaussian measurement noise with variance σ^2 ,

$$v_i \sim \mathcal{N}(0, \sigma^2). \quad (3)$$

We assume that the measurement noises are independent.

The measurement model can be written in the vector form

$$\mathbf{z} = \mathbf{h}(x) + \mathbf{v}, \quad (4)$$

where

$$\mathbf{z} := \begin{bmatrix} z_1 & z_2 & \dots & z_N \end{bmatrix}', \quad (5)$$

$$\mathbf{v} := \begin{bmatrix} v_1 & v_2 & \dots & v_N \end{bmatrix}', \quad (6)$$

$$\mathbf{h}(x) := h(x)\mathbf{d}, \quad (7)$$

$$\mathbf{d} := \begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}', \quad (8)$$

$$\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}), \quad \mathbf{R} = \mathbf{I}_N \sigma^2. \quad (9)$$

2.2. ML Estimate of Parameter

The likelihood function of x is [2,50,51]

$$\Lambda(x; \mathbf{z}) = p(\mathbf{z}|x) = [(2\pi)^N |\mathbf{R}|]^{-1/2} \exp\{-(1/2)[\mathbf{z} - \mathbf{h}(x)]' \mathbf{R}^{-1} [\mathbf{z} - \mathbf{h}(x)]\}. \quad (10)$$

The maximization of the likelihood in (10) is equivalent to the minimization of the cost function [2,51]

$$J(x) = [\mathbf{z} - \mathbf{h}(x)]' \mathbf{R}^{-1} [\mathbf{z} - \mathbf{h}(x)] = [\mathbf{z} - \mathbf{h}(x)]' [\mathbf{z} - \mathbf{h}(x)] / \sigma^2. \quad (11)$$

The maximum likelihood (ML) estimate $\hat{x}$ of x is obtained by setting the derivative of $J(x)$ to zero [2,51],

$$\frac{dJ(x)}{dx} = 0. \quad (12)$$

From (11) and (12), we obtain

$$[\mathbf{z} - \mathbf{h}(\hat{x})]' \frac{d\mathbf{h}(\hat{x})}{dx} = 0. \quad (13)$$

Because the derivative of $\mathbf{h}(x)$ with respect to x is not zero, we obtain

$$\mathbf{z} - \mathbf{h}(\hat{x}) = \mathbf{0}_{N \times 1}. \quad (14)$$

Hence, the ML estimate satisfies,

$$h(\hat{x})\mathbf{d} = \mathbf{z}. \quad (15)$$

Left-multiplying both sides of (15) by $\mathbf{d}'$, we obtain

$$h(\hat{x})\mathbf{d}'\mathbf{d} = \mathbf{d}'\mathbf{z} = \sum_{i=1}^N z_i. \quad (16)$$

We note that

$$\mathbf{d}'\mathbf{d} = N. \quad (17)$$

Using (1) and (17) in (16) we get

$$a\hat{x}^n = \bar{z}, \quad (18)$$

where $\bar{z}$ is the sample mean of z ,

$$\bar{z} = \frac{1}{N} \sum_{i=1}^N z_i. \quad (19)$$

Thus, from (18), the ML estimate of x is given by

$$\hat{x} = (\bar{z}/a)^{1/n}, \quad n = 2, 3, \dots \quad (20)$$

Remark 3. In general, the MLE for a nonlinear measurement model is biased [51]. We can calculate the variance of $\hat{x}$ under the small error assumption using the linearization approximation. To guarantee the validity of the variance, the bias in the MLE must be calculated. The bias can be numerically calculated using Monte Carlo simulation.

The bias in the MLE is defined by [2,51]

$$b(x) := x - \hat{x}. \quad (21)$$

Remark 4. The ML estimate of x in [35] was obtained by minimizing the cost function in (11) numerically. The estimator in (20) provides simple and efficient way of estimating x using a vector measurement $\mathbf{z}$ without numerical optimization.

2.3. Variance of the MLE

The variance of $\hat{x}$ is given by [51]

$$\sigma_x^2 = (\dot{\mathbf{H}}'\mathbf{R}^{-1}\dot{\mathbf{H}})^{-1}, \quad (22)$$

where

$$\dot{\mathbf{H}} = \left. \frac{d\mathbf{h}(x)}{dx} \right|_{x=\hat{x}}. \quad (23)$$

Using the special form of $\mathbf{R}$ from (9) in (22), we get

$$\sigma_x^2 = \sigma^2(\dot{\mathbf{H}}'\dot{\mathbf{H}})^{-1}. \quad (24)$$

Using (7) in (23), we get

$$\dot{\mathbf{H}} = \left. \frac{dh(x)}{dx} \right|_{x=\hat{x}} \mathbf{d}. \quad (25)$$

Differentiating (1) with respect to x , we obtain

$$\frac{dh(x)}{dx} = anx^{n-1}. \quad (26)$$

Using (26) in (25), we get

$$\dot{\mathbf{H}} = an\hat{x}^{n-1}\mathbf{d}. \quad (27)$$

From (27), we obtain

$$\dot{\mathbf{H}}'\dot{\mathbf{H}} = (an\hat{x}^{n-1})^2 \mathbf{d}'\mathbf{d}, \quad (28)$$

Using (28) and (17) in (24), we obtain

$$\sigma_x^2 = \sigma^2 (\mathbf{H}' \mathbf{H})^{-1} = \frac{\sigma^2}{N(an\hat{x}^{n-1})^2}, \quad (29)$$

$$\sigma_x = \frac{\sigma}{\sqrt{Nan\hat{x}^{n-1}}}. \quad (30)$$

2.4. Cramér-Rao Lower Bound

The CRLB [2,41] for the MSE in the current problem is given by

$$\text{CRLB}_x = \left[\frac{d\mathbf{h}'(x)}{dx} \mathbf{R}^{-1} \frac{d\mathbf{h}(x)}{dx} \right]^{-1}. \quad (31)$$

Remark 5. Calculation of the variance σ_x^2 and CRLB_x are similar. For σ_x^2 , we use the estimate $\hat{x}$ while calculating the Jacobian of the measurement function, whereas, for CRLB_x , we use the true x while calculating the Jacobian of the measurement function.

Using similar procedure, we obtain

$$\text{CRLB}_x = \frac{\sigma^2}{N(anx^{n-1})^2}, \quad (32)$$

$$\sqrt{\text{CRLB}_x} = \frac{\sigma}{\sqrt{Nanx^{n-1}}}. \quad (33)$$

From (30) and (33), we find that, for a given x , the standard deviation (SD) and square root of CRLB are inversely proportional to the power n . Secondly (33) shows that, for a given power, the square root of CRLB decreases as x increases.

3. Measures of Nonlinearity

To explain the key concepts of nonlinearity, consider the scalar function $h(x) = 5 \sin(4x)/x$ shown in Figure 1. We observe in Figure 1 that the function is nearly linear at A and E. If we draw a tangent to the curve at A and E, then the curve is close to the tangent in the neighborhood of A and E. However, tangents to the curve at points B, C, and D differ by large amounts from the curve in the neighborhood of these points. The tangent represents an affine approximation to the curve at a point. We observe that, among points B, C, and D, the curve bends the most at B and the least at point D. If we draw a circle (called the osculating circle) at these points, then the radius of the circle can be used to judge nonlinearity. The rate of bending is high when the radius of the circle is small. In differential geometry [37,38], the curvature κ is inverse of the radius of the osculating circle and, hence, curvature can be viewed as a measure of nonlinearity. The radii of the osculating circles at A and E are nearly infinity and, hence, the curvatures are nearly zero. From Figure 1, we observe that, in general, the nonlinearity of a function can vary with x . Hence, the nonlinearity is a local measure. If the second derivative of a function is non-zero, then the function is non-linear.

In [35,40], we analyzed the CMoN of a polynomial scalar function h of a non-random variable x , as described in Section 2.1. The CMoN were based on the extrinsic curvature using differential geometry, Bates and Watts parameter-effects curvature, and direct parameter-effects curvature. In this paper, we study the following MoNs:

- extrinsic curvature using differential geometry [36–38],
- Beale's MoN [18],
- least squares based Beale's MoN,
- Linssen's MoN [20],

- Least squares based Linssen's MoN,
- parameter-effects curvatures [21,25,29],
- Li's MoN [43,44], and
- MoN of Straka, Duník, and Šimandl [48,49].

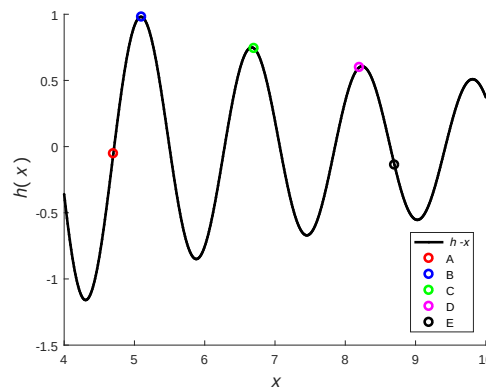


Figure 1. $y = 5 \sin(4x)/x$ versus x .

If a MoN has a high value, then the nonlinearity is high and if it has a low value, then the Therefore, it is impossible to compare them based on numerical values. We can only study their variations.

Consider the m -dimensional vector non-linear function $\mathbf{h}$ of the non-random n -dimensional parameter $\mathbf{x}$. Let $\hat{\mathbf{x}}$ be a known estimate of $\mathbf{x}$. Using the Taylor series expansion of $\mathbf{h}(\mathbf{x})$ about $\hat{\mathbf{x}}$ and keeping the first order term gives

$$\mathbf{h}(\mathbf{x}) \approx \mathbf{T}(\mathbf{x}) = \mathbf{h}(\hat{\mathbf{x}}) + \dot{\mathbf{H}}(\mathbf{x} - \hat{\mathbf{x}}), \quad (34)$$

where $\mathbf{T}(\mathbf{x})$ represents the tangent plane approximation (an affine mapping) to $\mathbf{h}(\mathbf{x})$ and

$$\dot{\mathbf{H}} = \left. \frac{d\mathbf{h}(\mathbf{x})}{d\mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}}. \quad (35)$$

If $m > n$, then $\mathbf{h}$ is an n -dimensional manifold embedded in an m -dimensional space [37,38]. The tangent plane is tangent to the surface $\mathbf{h}$ at $\hat{\mathbf{x}}$. The concept of tangent plane is used in Beale's MoN, Linssen's MoN, Bates and Watts parameter-effects curvatures [21,25], and direct parameter-effects curvature [44].

For polynomial nonlinearity, the CMoN using differential geometry is calculated at the true value x and, hence, it is non-random. The Bates and Watts parameter-effects curvature, direct parameter-effects curvature, Beale's MoN, Li's MoN, and the MoN of Straka et al. are calculated while using an estimate $\hat{x}$ of x . The estimate $\hat{x}$ is obtained from a measurement model involving the measurement function h . Since x is a scalar, we need one or more scalar measurements to estimate x . Table 1 summarizes features of various MoNs.

Table 1. Features of Various MoNs

MoN	Parameters Used	Local/Global	Need?			Basic Idea	Random?
			Measurements	Jacobian/Hessian	Covariances		
Extrinsic Curvature	True	Local	No	Jacobian and Hessian at true value	No	Differential Geometry	Non-random
Beale's MoN	True & estimated	Local	Yes	Jacobian	No	Scaled sum square distance	Random
Linssen's MoN	True & estimated	Local	Yes	Jacobian	No	Root scaled sum square distance	Random
Parameter-effects Curvature	True & estimated	Local	Yes	Jacobian and Hessian at estimated value	No	Differential Geometry	Random
Li's MoN	True	Global	Yes	No	Yes	min. mean square distance	Random
MoN by Straka et al.	True & estimated	Local	Yes	No	No	WLS cost function	Random

The CMoN using differential geometry [36–38] is calculated at the true value x , whereas the Bates and Watts parameter-effects curvature [21,25,26], direct parameter-effects curvature [29], Beale's MoN, Li's MoN, and the MoN of Straka et al. are calculated while using an estimate $\hat{x}$ of x . The estimate $\hat{x}$ is obtained from a measurement model involving the measurement function h . Since x is a scalar, we need one or more scalar measurements to estimate x . Next, we describe various MoN.

3.1. Extrinsic Curvature Using Differential Geometry

The curvature of a circle at every point on the circumference is equal to the inverse of the radius of the circle. Thus, the curvature of a circle is a constant. A circle with a smaller radius bends more sharply and, therefore, has a higher curvature.

We assume that the first and second derivatives of the nonlinear smooth scalar function h exist. The curvature of the curve $y = h(x)$ at a point x is equal to the curvature of the osculating circle at that point. The extrinsic curvature at the point x is defined by [36–38],

$$\kappa(x) := \frac{\left| \frac{d^2 h(x)}{dx^2} \right|}{\left[1 + \left(\frac{dh(x)}{dx} \right)^2 \right]^{3/2}} = \frac{|\ddot{h}(x)|}{[1 + \dot{h}(x)^2]^{3/2}}. \quad (36)$$

The first derivative of h at a point x is given in (26). The second derivative of h with respect to x is given by

$$\ddot{h}(x) = \frac{d^2 h(x)}{dx^2} = an(n-1)x^{n-2}, \quad n = 2, 3, \dots \quad (37)$$

Thus, using $\dot{h}(x)$ and $\ddot{h}(x)$ in (36), we can calculate the extrinsic curvature $\kappa(x)$ at any point x by

$$\kappa(x) = \frac{an(n-1)x^{n-2}}{[1 + (an)^2 x^{2(n-1)}]^{3/2}}. \quad (38)$$

3.2. Beale's MoN

Consider the nonlinear measurement model for the non-random n -dimensional parameter $\mathbf{x}$

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{v}, \quad (39)$$

where $\mathbf{z}$, $\mathbf{h}$, and $\mathbf{v}$ are the measurement, non-linear measurement function, and measurement noise, respectively. Let $\hat{\mathbf{x}}$ be an estimate of $\mathbf{x}$. Subsequently, a Taylor series expansion of $\mathbf{h}(\mathbf{x})$ about

$\hat{\mathbf{x}}$ and keeping the first order term is as (34). Suppose we choose m vectors $\mathbf{x}_i, i = 1, 2, \dots, m$ in the neighborhood of $\mathbf{x}$. Then Beale's first empirical MoN [18] is given by

$$\hat{N}_{\mathbf{x}} = \rho^2 \frac{\sum_{i=1}^m \|\mathbf{h}(\mathbf{x}_i) - \mathbf{T}(\mathbf{x}_i)\|^2}{\sum_{i=1}^m \|\mathbf{h}(\mathbf{x}_i) - \mathbf{h}(\hat{\mathbf{x}})\|^4}, \quad (40)$$

where ρ is the standard radius and it is defined by

$$\rho^2 := \|\mathbf{z} - \mathbf{h}(\hat{\mathbf{x}})\|^2 / (n(N - n)). \quad (41)$$

Guttman and Meeter [19] observed that the empirical MoN underestimates severe nonlinearity. When m approaches infinity, the empirical MoN $\hat{N}_{\mathbf{x}}$ approaches the theoretical MoN $N_{\mathbf{x}}$.

3.3. Least Squares Based Beale's MoN

Consider the scalar function h for polynomial nonlinearity, as described in (1). As described in Beale's MoN, we choose m points $x_i, i = 1, 2, \dots, m$ in the neighborhood of x . Let

$$y_i = ax_i^n, \quad i = 1, 2, \dots, m. \quad (42)$$

An affine mapping as approximation to y_i is given by

$$L(x_i) = A + Bx_i, \quad i = 1, 2, \dots, m. \quad (43)$$

We compute A and B by minimizing the cost function

$$J(A, B) := \sum_{i=1}^m (y_i - A - Bx_i)^2 \quad (44)$$

by the method of least squares (LS) [2,39]. The LS minimization of the cost function yields [52]

$$\hat{B} = (C_{xy} - \bar{x} \bar{y}) / (C_{xx} - \bar{x}^2), \quad (45)$$

$$\hat{A} = \bar{y} - \hat{B} \bar{x}, \quad (46)$$

where

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i, \quad \bar{y} = \frac{1}{m} \sum_{i=1}^m y_i, \quad (47)$$

$$C_{xx} = \frac{1}{m} \sum_{i=1}^m x_i^2, \quad C_{xy} = \frac{1}{m} \sum_{i=1}^m x_i y_i. \quad (48)$$

Then we can use the affine mapping with $\hat{A}$ and $\hat{B}$ in Beale's MoN.

3.4. Linssen's MoN

In order to correct the deficiency in Beale's MoN, Linssen proposed a modification to obtain the following MoN [20]

$$M^* = \sqrt{\rho^2 \frac{\sum_{i=1}^m \|\mathbf{h}(\mathbf{x}_i) - \mathbf{T}(\mathbf{x}_i)\|^2}{\sum_{i=1}^m \|\mathbf{h}(\hat{\mathbf{x}}) - \mathbf{T}(\mathbf{x}_i)\|^4}}. \quad (49)$$

3.5. Least Squares Based Linssen's MoN

Using the same procedure as in Section 3.3, we can use an affine mapping with $\hat{A}$ and $\hat{B}$ as an approximation to y_i in computing Linssen's MoN.

3.6. Parameter-Effects Curvatures

The parameter-effects curvature and intrinsic curvature defined by Bates and Watts [21,25,26] are associated with a non-linear parameter estimation problem and are defined at the estimated parameter. We note that in (1), $h: \mathbb{R} \rightarrow \mathbb{R}$. Since h is a scalar function, the intrinsic curvature of Bates and Watts $K^N(\hat{x})$ [21] or the direct intrinsic curvature $\beta_\delta^N(\hat{x})$ [29] is zero. Thus, only the parameter-effects curvature of Bates and Watts $K^T(\hat{x})$ and the direct parameter-effects curvature $\beta_\delta^T(\hat{x})$ are non-zero. Since the intrinsic curvature is zero, for simplicity in notation, we drop the superscript “T” from the parameter-effects curvature and they are given by

$$K(\hat{x}) := \frac{\|\ddot{\mathbf{H}}\delta^2\|}{\|\dot{\mathbf{H}}\delta\|^2} = \frac{\|\ddot{\mathbf{H}}\|}{\|\dot{\mathbf{H}}\|^2}, \quad (50)$$

$$\beta_\delta(\hat{x}) := \frac{\|\ddot{\mathbf{H}}\delta^2\|}{\|\dot{\mathbf{H}}\delta\|} = \frac{\|\ddot{\mathbf{H}}\|\|\delta\|}{\|\dot{\mathbf{H}}\|}, \quad (51)$$

where

$$\ddot{\mathbf{H}} = \frac{d^2 \mathbf{h}(x)}{dx^2} \Big|_{x=\hat{x}}, \quad (52)$$

$$\delta := x - \hat{x}. \quad (53)$$

From (26), we get

$$\ddot{\mathbf{H}} = an(n-1)\hat{x}^{n-2}\mathbf{d}. \quad (54)$$

Hence, from (27) and (52), we obtain

$$\|\dot{\mathbf{H}}\| = an\hat{x}^{n-1}\sqrt{N}. \quad (55)$$

$$\|\ddot{\mathbf{H}}\| = an(n-1)\hat{x}^{n-2}\sqrt{N}. \quad (56)$$

Substitution of results from (55) and (56) in (50) and (51) gives

$$K(\hat{x}) = \frac{n-1}{na\sqrt{N}} \frac{1}{\hat{x}^n}, \quad (57)$$

$$\beta_\delta(\hat{x}) = \frac{(n-1)|\delta|}{\hat{x}}. \quad (58)$$

We note that the extrinsic curvature in (36) is evaluated at the true x , while the parameter-effects curvatures $K(\hat{x})$ in (50) and $\beta_\delta(\hat{x})$ in (51) are evaluated at the estimate $\hat{x}$. Because $\hat{x}$ is a random variable, $K(\hat{x})$ and $\beta_\delta(\hat{x})$ are random variables. When we perform Monte Carlo simulations and estimate x from measurements, $\hat{x}$ varies among Monte Carlo runs. Therefore, $K(\hat{x})$ and the set of all linear $\beta_\delta(\hat{x})$ vary with Monte Carlo runs.

3.7. Li's MoN

For a scalar random variable x , the un-normalized MoN proposed by Li [43,44] represents the square root of the minimum mean square distance between the nonlinear measurement function h and the set of all affine functions L ,

$$J = \min(E\{(L(x) - h(x))^2\})^{1/2}, \quad (59)$$

where $L(x) = Ax + B$. The scalar parameters A and B are determined in the minimization process. For the current problem, where x is non-random, the un-normalized MoN J and normalized MoN ν are given, respectively, by

$$J = \sigma_h \sqrt{1 - \frac{c_{hx}^2}{\sigma_h^2 \sigma_x^2}}. \quad (60)$$

$$\nu = J/\sigma_h = \sqrt{1 - \frac{c_{hx}^2}{\sigma_h^2 \sigma_x^2}}. \quad (61)$$

Given $\hat{x}$ and σ_x (30), the unscented transformation (UT) [14,15], cubature transformation (CT) [16], or Monte Carlo method [8] can be used to compute σ_h^2 and c_{hx} . We find that the UT gives good results in calculating the two MoNs. Next we describe computing J and ν using the UT. We use $\kappa_{UT} = 2$ [14]. The three weights and sigma points are given, respectively, by

$$w_0 = 2/3, \quad w_1 = 1/6, \quad w_2 = 1/6, \quad (62)$$

$$\chi_0 = \hat{x}, \quad \chi_1 = \hat{x} + \sqrt{3}\sigma_x, \quad \chi_2 = \hat{x} - \sqrt{3}\sigma_x. \quad (63)$$

The measurement transformed points are

$$h_i = a \chi_i^n, \quad i = 0, 1, 2. \quad (64)$$

Then the mean and variance of h are given by

$$\bar{h} = \sum_{i=0}^2 w_i h_i, \quad (65)$$

$$\sigma_h^2 = \sum_{i=0}^2 w_i (h_i - \bar{h})^2, \quad (66)$$

The cross-covariance c_{hx} is computed by

$$c_{hx} = \sum_{i=0}^2 w_i (h_i - \bar{h})(\chi_i - \hat{x}). \quad (67)$$

3.8. MoN of Straka, Duník, and Šimandl

Straka, Duník, and Šimandl presented two local MoNs in [48,49]. Given the estimate $\hat{x}$ and variance σ_x^2 , these MoNs use a number of points $\chi_i, i = 1, 2, \dots, m$ in the neighborhood of $\hat{x}$. We analyze the first MoN proposed by the authors. The transformed points by the non-linear function h are given by

$$z_i = h(\chi_i), \quad i = 1, 2, \dots, m. \quad (68)$$

Define

$$\mathbf{Z} := \begin{bmatrix} z_1 & z_2 & \dots & z_m \end{bmatrix}', \quad (69)$$

$$\mathbf{X} := \begin{bmatrix} \chi_1 & \chi_2 & \dots & \chi_m \end{bmatrix}'. \quad (70)$$

A linear approximation to $\mathbf{Z}$ is $\mathbf{X}\theta$, where θ is a scalar parameter to be estimated. The cost function that is proposed in [48,49] to determine θ is given by

$$J_1(\theta) := (\mathbf{Z} - \mathbf{X}\theta)' \mathbf{W} (\mathbf{Z} - \mathbf{X}\theta), \quad (71)$$

where the weight-matrix $\mathbf{W}$ is given by

$$\mathbf{W} = \text{diag}(d_1, d_2, \dots, d_m), \quad (72)$$

$$d_i = (\chi_i - \hat{x})^2, \quad i = 1, 2, \dots, m. \quad (73)$$

The LS estimate [39] that minimizes the cost function is given by

$$\hat{\theta}_{\text{LS}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{Z}. \quad (74)$$

For this problem, the LS estimate in (74) reduces to

$$\hat{\theta}_{\text{LS}} = \left(\sum_{i=1}^m d_i \chi_i^2 \right)^{-1} \sum_{i=1}^m d_i \chi_i z_i. \quad (75)$$

The cost function J_1 evaluated at $\hat{\theta}_{\text{LS}}$ is treated as a local MoN η , given by

$$\eta = J_1(\hat{\theta}_{\text{LS}}). \quad (76)$$

Remark 6. We have calculated the average MoN for the bearing-only filtering [27], GMTI [32], and video filtering [34] problems. The MoN is presented in the table below (Table 2). From this table we find that the degree of nonlinearity of the bearing-only filtering problem is about two orders of magnitude higher than that of the GMTI or video filtering problem. This implies that a simple filter, such as the EKF or UKF, is sufficient for the GMTI or video filtering problem, but an advanced filter, such as the PF, is needed for the BOF [17] problem.

Table 2. MoNs for the bearing-only, GMTI, and video filtering problems.

Curvature Type	Bearing-Only	GMTI	Video
Parameter-effects	$(300\text{--}1200) \times 10^{-4}$	$(0.8\text{--}1.2) \times 10^{-4}$	0.245×10^{-4}
Intrinsic	$(69\text{--}149) \times 10^{-4}$	0.2×10^{-4}	0.066×10^{-4}
Total	$(369\text{--}1349) \times 10^{-4}$	$(1.0\text{--}1.4) \times 10^{-4}$	0.312×10^{-4}

4. Mapping between CMoN and MSE in Polynomial NonLinearity

The nonlinearity of the problem imposes challenges in parameter estimation. We analyze the CMoN and MSE of the non-linear estimation problem to discover relationships among them. For the current problem, CMoN are measured by the parameter-effects curvature in (57) and the direct parameter-effects curvature in (58). In general, CMoN depend on the first and second derivatives of the non-linear function calculated at the parameter estimate and on the norm of the estimation error for $\beta_\delta(\hat{x})$. Therefore, the CMoN will depend the type of estimator (e.g., ML) used to obtain parameter estimate. The extrinsic curvature (38) depends on the first and second derivatives of the non-linear function evaluated at the true x .

4.1. MSE and Sample MSE

We estimate the x coordinate using noisy measurements at a discrete set $\{x_k\}_{k=1}^{N_x}$ of values. Let $\hat{x}_{k,m}$ denote the estimate of x_k in the m th Monte Carlo run. Subsequently, the error $\tilde{x}_{k,m}$ in $\hat{x}_{k,m}$ is defined by

$$\tilde{x}_{k,m} := x_k - \hat{x}_{k,m}, \quad k = 1, 2, \dots, N_x, \quad m = 1, 2, \dots, M, \quad (77)$$

where M is the number of Monte Carlo runs. The MSE at x_k is given by

$$\text{MSE}_k = E[(\tilde{x}_{k,m})^2], \quad k = 1, 2, \dots, N_x. \quad (78)$$

The sample MSE (SMSE) at x_k is defined by

$$\text{SMSE}_k := \frac{1}{M} \sum_{m=1}^M (\tilde{x}_{k,m})^2, \quad k = 1, 2, \dots, N_x. \quad (79)$$

Let $L_{\text{CRLB}}(x)$ denote the $\log_{10}$ of the CRLB,

$$L_{\text{CRLB}}(x) := \log_{10} \text{CRLB}_x. \quad (80)$$

Taking the log of CRLB_x in (32) we get

$$L_{\text{CRLB}}(x) = \log_{10} \left(\frac{\sigma^2}{Nn^2a^2} \right) - 2(n-1) \log_{10} x. \quad (81)$$

4.2. MSE and Parameter-Effects Curvature

Let $L_K(x)$ denote the log of the expected value of $K(\hat{x})$ in (57). Then

$$L_K(x) := \log_{10} \{E[K(\hat{x})]\}. \quad (82)$$

In order to compute $L_K(x)$, we first approximate the expectation in (82) by assuming $\sigma_{\hat{x}} \ll x$, which holds for the case investigated in our paper,

$$E\{K(\hat{x})\} = \frac{(n-1)}{na} E\left(\frac{1}{\hat{x}^n}\right) \approx \frac{(n-1)}{na} \frac{1}{[E(\hat{x})]^n} \approx \frac{(n-1)}{na} \frac{1}{x^n}. \quad (83)$$

The last step of the above equation follows from an assumption that the estimator is nearly unbiased. Now, taking the logarithm, we have

$$L_K(x) = \log_{10} \left(\frac{n-1}{na} \right) - n \log_{10} x. \quad (84)$$

Now, from Equations (84) and (81), we can see that there is an affine mapping between $L_{\text{CRLB}}(x)$ and $L_K(x)$. That is,

$$L_{\text{CRLB}}(x) = \alpha_1^K L_K(x) + \alpha_0^K, \quad (85)$$

where

$$\begin{aligned} \alpha_1^K &= \frac{2(n-1)}{n}, \\ \alpha_0^K &= \log_{10} \left(\frac{\sigma^2}{Nn^2a^2} \right) - \frac{2(n-1)}{n} \log_{10} \left(\frac{n-1}{na} \right). \end{aligned} \quad (86)$$

We observe that α_1^K is positive and, hence, $L_K(x)$ and $L_{\text{CRLB}}(x)$ have the same sign of the non-zero slopes. As a result, $K(\hat{x})$ and CRLB have the same sign of the non-zero slopes.

4.3. MSE and Direct Parameter-Effects Curvature

The expression for the direct parameter-effects curvature $\beta_{\delta}(\hat{x})$ [29,30] is given by (58). Similar to the previous section, we define

$$L_{\beta}(x) := \log_{10} (E\{\beta_{\delta}(\hat{x})\}). \quad (87)$$

Now, taking the expected value of β , we have

$$E\{\beta(\hat{x})\} \approx \frac{(n-1)}{x} E\{|\delta|\} = \frac{(n-1)}{x} E\{|\hat{x} - x|\}. \quad (88)$$

The RHS of (88) can be simplified by assuming that $\hat{x}$ is unbiased and that it achieves the CRLB. Additionally, we approximate this error to be Gaussian and the variance of $\hat{x}$ is given in (29). Then,

$$E\{|\hat{x} - x|\} = \sqrt{\text{CRLB}_x} \sqrt{\frac{2}{\pi}}. \quad (89)$$

Substituting (89) into (88) and using (32) for CRLB_x we have

$$\mathbb{E}[\beta(\hat{x})] \approx \frac{(n-1)}{na} \sigma \sqrt{\frac{2}{N\pi}} x^{-n}. \quad (90)$$

Thus,

$$L_\beta(x) \approx \log_{10} \left[\frac{(n-1)\sigma \sqrt{\frac{2}{N\pi}}}{na} \right] - n \log_{10} x. \quad (91)$$

From (91) and (81) we can write the affine mapping

$$L_{\text{CRLB}}(x) = \alpha_1^\beta L_\beta(x) + \alpha_0^\beta, \quad (92)$$

where

$$\begin{aligned} \alpha_1^\beta &= \frac{2(n-1)}{n}, \\ \alpha_0^\beta &= \log_{10} \left(\frac{\sigma^2}{Nn^2a^2} \right) - \frac{2n-2}{n} \log_{10} \left[\frac{(n-1)\sigma\sqrt{2}}{na\sqrt{N\pi}} \right]. \end{aligned} \quad (93)$$

We also observe that α_1^β is positive and, hence, $L_\beta(x)$ and $L_{\text{CRLB}}(x)$ have the same sign of the non-zero slopes. As a result, $\beta_\delta(\hat{x})$ and CRLB have the same sign of the non-zero slopes.

4.4. Extrinsic Curvature

The expression for extrinsic curvature for our problem is given in (38). Similar to previous sections, we define

$$L_\kappa(x) := \log_{10}(\kappa(x)). \quad (94)$$

Taking the log of (94), we have

$$\begin{aligned} L_\kappa(x) &= \log_{10}(\kappa(x)) \\ &= \log_{10} \left[an(n-1)x^{n-2} \right] - \frac{3}{2} \log_{10} \left[1 + (anx^{n-1})^2 \right] \\ &\approx \log_{10} \left[an(n-1)x^{n-2} \right] - \frac{3}{2} \log_{10} \left[(anx^{n-1})^2 \right] \\ &= \log_{10} \left[\frac{n-1}{(an)^2} \right] - (2n-1) \log_{10} x. \end{aligned} \quad (95)$$

Note that the second last expression is a valid approximation for $x > 2$. From (95) and (84) it is easy to establish the affine mapping

$$L_K(x) = \gamma_1^K L_\kappa(x) + \gamma_0^K, \quad (96)$$

where

$$\begin{aligned}\gamma_1^K &= \frac{n}{2n-1}, \\ \gamma_0^K &= \log_{10} \left(\frac{n-1}{na} \right) - \frac{n}{2n-1} \log_{10} \left[\frac{n-1}{(an)^2} \right]\end{aligned}\quad (97)$$

Similarly, from (95) and (91) we can establish the affine relationship

$$L_\beta(x) = \gamma_1^\beta L_\kappa(x) + \gamma_0^\beta, \quad (98)$$

where

$$\begin{aligned}\gamma_1^\beta &= \frac{n}{2n-1}, \\ \gamma_0^\beta &= \log_{10} \left[\frac{(n-1)\sigma\sqrt{2}}{na\sqrt{N\pi}} \right] - \frac{n}{2n-1} \log_{10} \left[\frac{n-1}{(an)^2} \right].\end{aligned}\quad (99)$$

Using similar arguments used in previous sections, we infer that the extrinsic curvature and parameter-effects curvature have the same sign of the non-zero slopes. Similarly, the extrinsic curvature and direct parameter-effects curvature have the same non-zero slopes.

4.5. Estimation of CMoN and SMSE by Monte-Carlo Simulations

Let $\bar{K}(\hat{x}_k)$ and $\bar{\beta}_\delta(\hat{x}_k)$ denote the sample means of the Bates and Watts and direct parameter-effects curvatures calculated from M Monte Carlo runs. Subsequently,

$$\bar{K}(\hat{x}_k) := \frac{1}{M} \sum_{m=1}^M K(\hat{x}_{k,m}), \quad k = 1, 2, \dots, N_x, \quad (100)$$

$$\bar{\beta}_\delta(\hat{x}_k) := \frac{1}{M} \sum_{m=1}^M \beta_\delta(\hat{x}_{k,m}), \quad k = 1, 2, \dots, N_x. \quad (101)$$

Correspondingly, we define

$$b_k := \log_{10} \text{SMSE}_k, \quad k = 1, 2, \dots, N_x, \quad (102)$$

$$c_k := \log_{10} \bar{K}(\hat{x}_k), \quad k = 1, 2, \dots, N_x, \quad (103)$$

$$d_k := \log_{10} \bar{\beta}_\delta(\hat{x}_k), \quad k = 1, 2, \dots, N_x. \quad (104)$$

Define

$$\mathbf{b} := \begin{bmatrix} b_1 & b_2 & \dots & b_{N_x} \end{bmatrix}', \quad (105)$$

$$\mathbf{c} := \begin{bmatrix} c_1 & c_2 & \dots & c_{N_x} \end{bmatrix}', \quad (106)$$

$$\mathbf{d} := \begin{bmatrix} d_1 & d_2 & \dots & d_{N_x} \end{bmatrix}'. \quad (107)$$

Suppose that an affine mapping exists between $\mathbf{b}$ and $\mathbf{c}$. Subsequently,

$$b_k = \hat{a}_1^K c_k + \hat{a}_0^K + e_k, \quad k = 1, 2, \dots, N_x, \quad (108)$$

where e_k is a random noise. Afterwards, we can write (108) in the matrix-vector form by

$$\mathbf{b} = \mathbf{H}_c \boldsymbol{\alpha} + \mathbf{e}, \quad (109)$$

where

$$\boldsymbol{\alpha} := \begin{bmatrix} \hat{a}_1^K & \hat{a}_0^K \end{bmatrix}', \quad (110)$$

$$\mathbf{e} := \begin{bmatrix} e_1 & e_2 & \dots & e_{N_x} \end{bmatrix}', \quad (111)$$

$$\mathbf{H}_c := \begin{bmatrix} c_1 & 1 \\ c_2 & 1 \\ \dots & \dots \\ c_{N_x} & 1 \end{bmatrix}. \quad (112)$$

Given $\mathbf{b}$ and $\mathbf{H}_c$, we can estimate $\boldsymbol{\alpha}$ using the linear least squares (LLS).

We can similarly define the affine mapping between other variable pairs. Altogether, we consider the following four:

1. between $\mathbf{b}$ ($\log_{10}(\text{SMSE}_k)$) and $\mathbf{c}$ ($\log_{10}(\bar{K}(\hat{x}_k))$) for each power of the polynomial function, as in (85),
2. between $\mathbf{c}$ ($\log_{10}(\bar{K}(\hat{x}_k))$) and $\log_{10}(\kappa(x_k))$ (94) for each power of the polynomial function, as in (96), and
3. between $\mathbf{d}$ ($\log_{10}(\bar{\beta}_\delta(\hat{x}_k))$) and $\log_{10}(\kappa(x_k))$ (94) for each power of the polynomial function, as in (98).

5. Numerical Simulation and Results

We follow the same simulation scenario as used in our previous work [35]. We use $a = 0.6$ and $n = 2, 3, 4, 5$ and a number of uniformly spaced x coordinates with the spacing of 0.1 in the interval $[2, 7]$. The measurement noise standard deviation (σ) is 0.5. The dimension of the measurement vector is 10 or 20. The results are based on 1000 Monte Carlo runs. Figure 2 shows $\log_{10}(h(x))$ versus x .

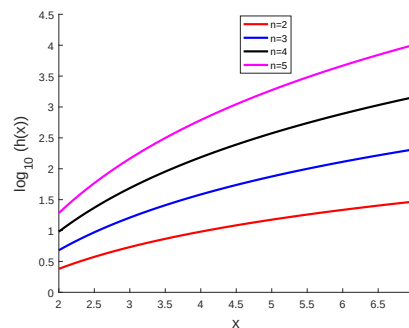


Figure 2. $\log_{10}(h(x))$ versus x .

To assess the accuracy of the MLE, we compute the sample bias, sample MSE, ANEES [42], and CRLB [2,41,51]. Let $x_{k,i} = x_k$, $\hat{x}_{k,i}$, and $\sigma_{k,i}^2$ denote the true parameter, ML estimate, and associated variance, respectively, at the k th point in the i th Monte Carlo run. The sample bias in the estimate at the k th point is defined by [9]

$$\hat{b}_k := \frac{1}{M} \sum_{i=1}^M (x_{k,i} - \hat{x}_{k,i}), \quad (113)$$

where M is the number of Monte Carlo runs. The sample root MSE (RMSE) [9] and ANEES [2,9,42] at the k th point are defined, respectively, by

$$\text{RMSE}_k := \left[\frac{1}{M} \sum_{i=1}^M (x_{k,i} - \hat{x}_{k,i})^2 \right]^{1/2}, \quad (114)$$

$$\text{ANEES}_k := \frac{1}{M} \sum_{i=1}^M (x_{k,i} - \hat{x}_{k,i})^2 / \sigma_{k,i}^2. \quad (115)$$

Figure 3 presents the sample bias for different powers of x . We observe from Figure 3 that the bias is small when compared with the true value of x and the bias decreases with increase in the power of x . In Figure 4, we have plotted the $\sqrt{\text{CRLB}}$ and the average of σ_x over Monte Carlo runs. Figure 4 shows that, for each power of x , the $\sqrt{\text{CRLB}}$ and the average of σ_x are on top of each other and it is not possible to distinguish them in the figure.

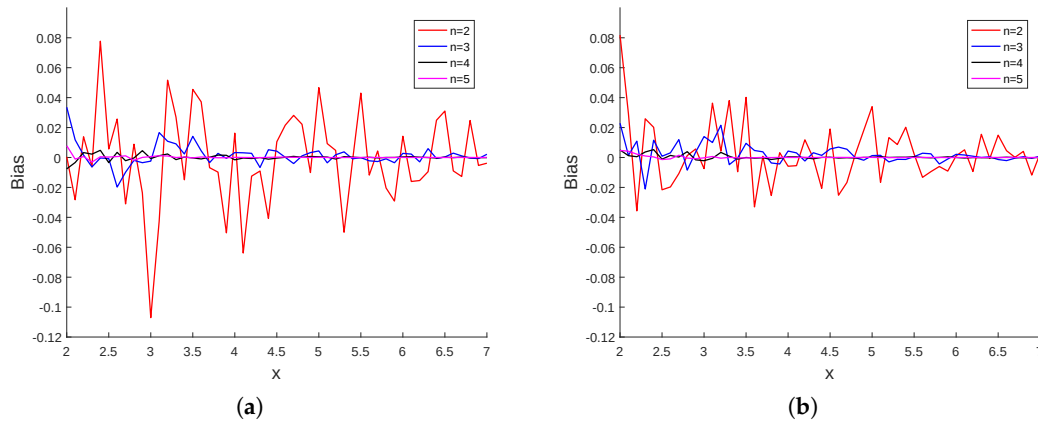


Figure 3. (a) Sample bias vs. x using 10 scalar measurements and (b) sample bias vs. x using 20 scalar measurements.

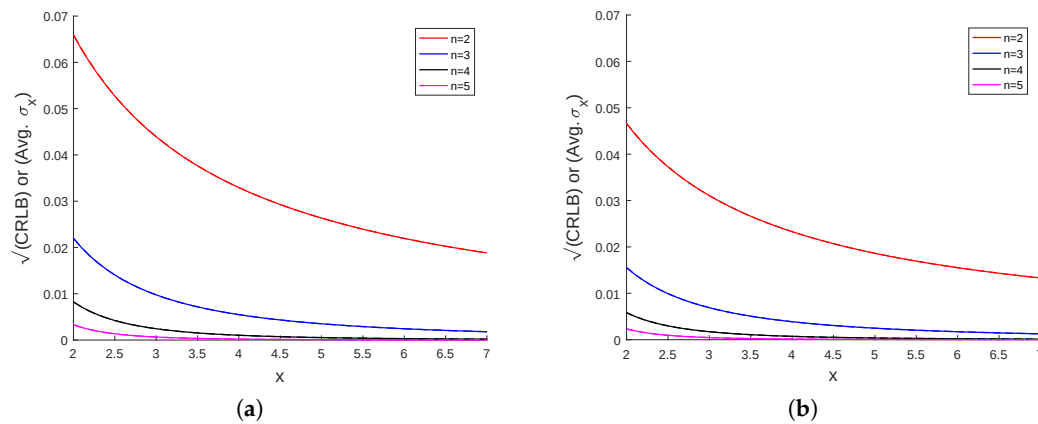


Figure 4. (a) $\sqrt{\text{CRLB}}$ or (Avg. σ_x) vs. x using 10 scalar measurements and (b) $\sqrt{\text{CRLB}}$ or (Avg. σ_x) vs. x using 20 scalar measurements.

Figure 5 presents $\sqrt{\text{CRLB}}$ and RMSE for each power of x . Solid and dashed lines in Figure 5 represent the $\sqrt{\text{CRLB}}$ and RMSE, respectively, for each power of x . We see from Figure 5 that corresponding values of $\sqrt{\text{CRLB}}$ and RMSE are close to each other for each power of x . In Figures 3–5, the bias, $\sqrt{\text{CRLB}}$, σ_x , and RMSE for 20 measurements are smaller than corresponding values for 10 measurements.

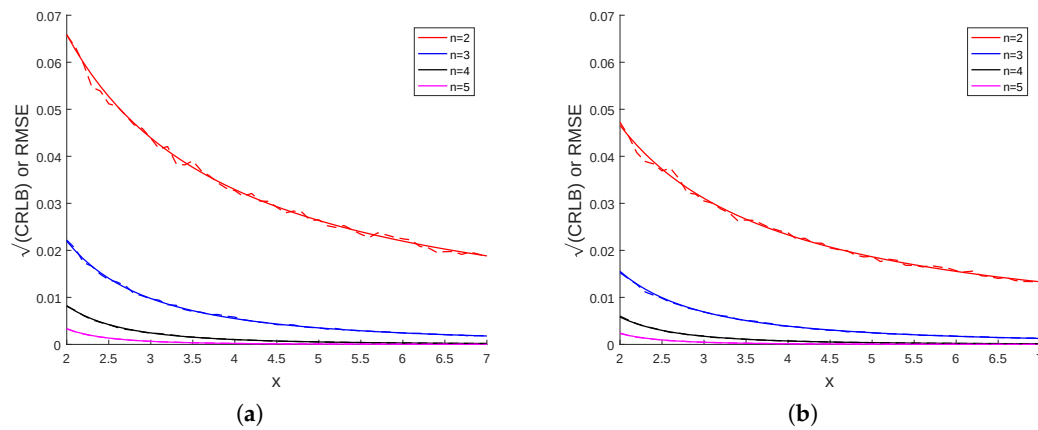


Figure 5. (a) $\sqrt{\text{CRLB}}$ or RMSE vs. x using 10 scalar measurements and (b) $\sqrt{\text{CRLB}}$ or RMSE vs. x using 20 scalar measurements.

We present the ANEES [42] in Figure 6 for different powers of x with 99% confidence bounds. We see from Figure 6 that the ANEES lies within the 99% confidence bounds. This shows that the variance σ_x^2 calculated using the MLE is consistent with the estimation error.

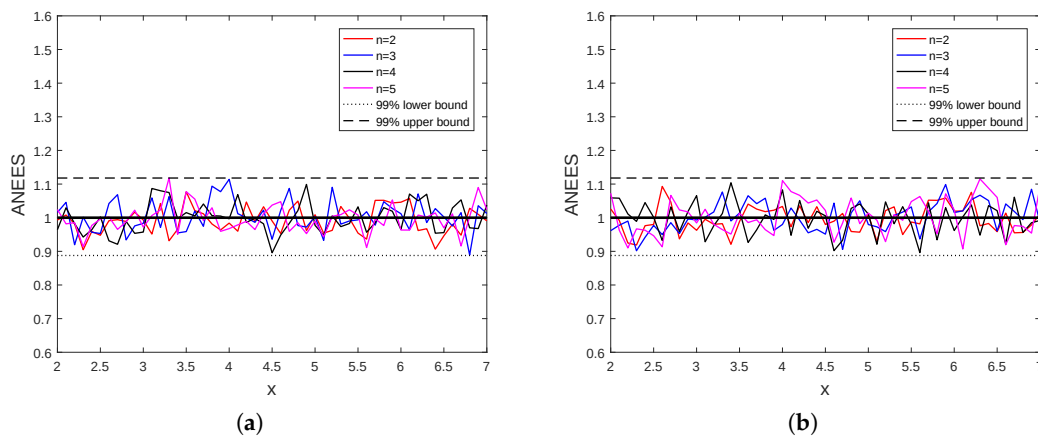


Figure 6. (a) ANEES vs. x using 10 scalar measurements and (b) ANEES vs. x using 20 scalar measurements.

Figure 7 presents the logarithm of the extrinsic curvature $\log_{10}(\kappa(x))$ versus x . The extrinsic curvature is completely determined by the first and second derivatives of the non-linear function h and it is evaluated while using the true x .

In Figures 8–18, we present results using 10 scalar measurements. We have also generated results using 20 scalar measurements. In order to limit the number of figures, we have not presented figures with 20 scalar measurements. The CRLB, variance of estimation error, all MoNs, and MSE follow the same trend. However, the corresponding values compared with 20 measurements are reduced due to improved estimation accuracy.

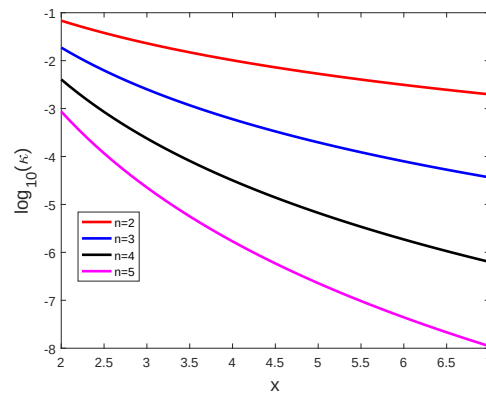


Figure 7. Logarithm of the extrinsic curvature $\log_{10}(\kappa(x))$ versus x .

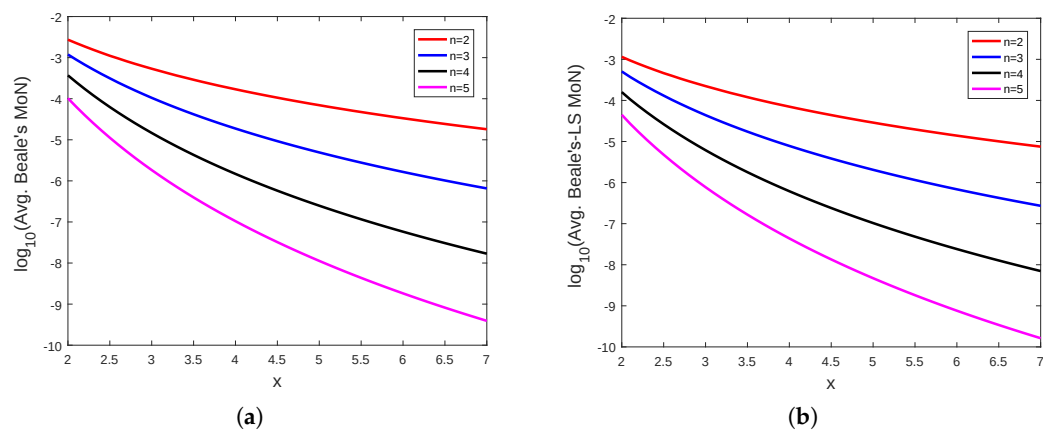


Figure 8. (a) Logarithm of Beale's MoN ($\log_{10}(\text{Avg. Beale's MoN})$) vs. x and (b) logarithm of Beale's MoN using LS ($\log_{10}(\text{Avg. Beale's-LS MoN})$) vs. x with 10 scalar measurements.

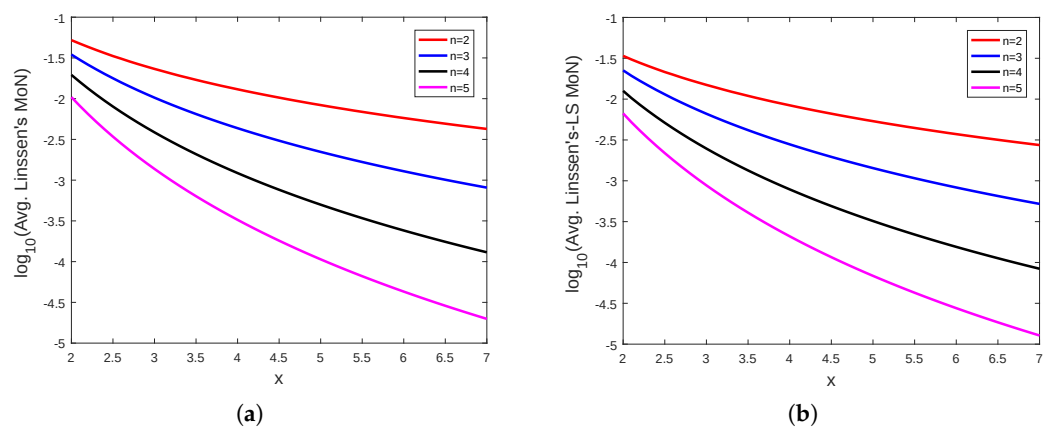


Figure 9. (a) Logarithm of Linssen's MoN ($\log_{10}(\text{Avg. Linssen's MoN})$) vs. x and (b) logarithm of Linssen's MoN using LS ($\log_{10}(\text{Avg. Linssen's-LS MoN})$) vs. x with 10 scalar measurements.

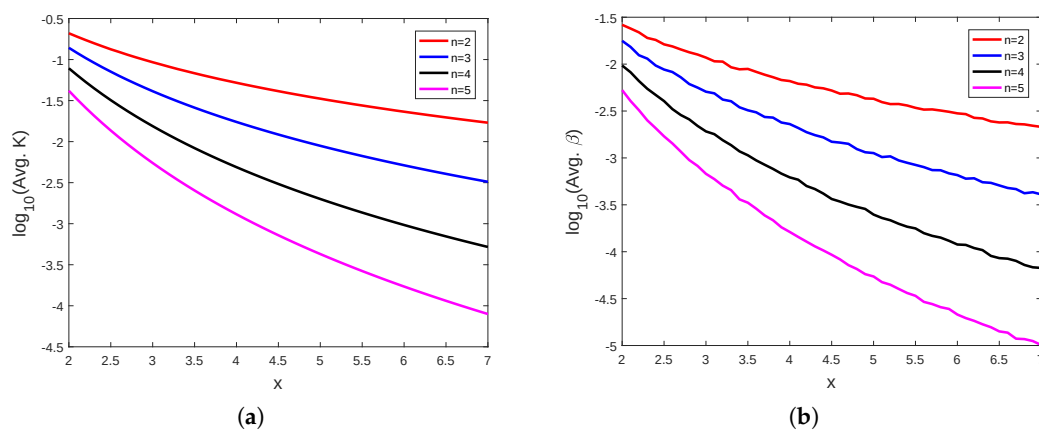


Figure 10. (a) Logarithm of Bates and Watts parameter-effects curvature ($\log_{10}(\text{Avg. } K)$) vs. x and (b) logarithm of direct parameter-effect curvature ($\log_{10}(\text{Avg. } \beta)$) vs. x using 10 scalar measurements.

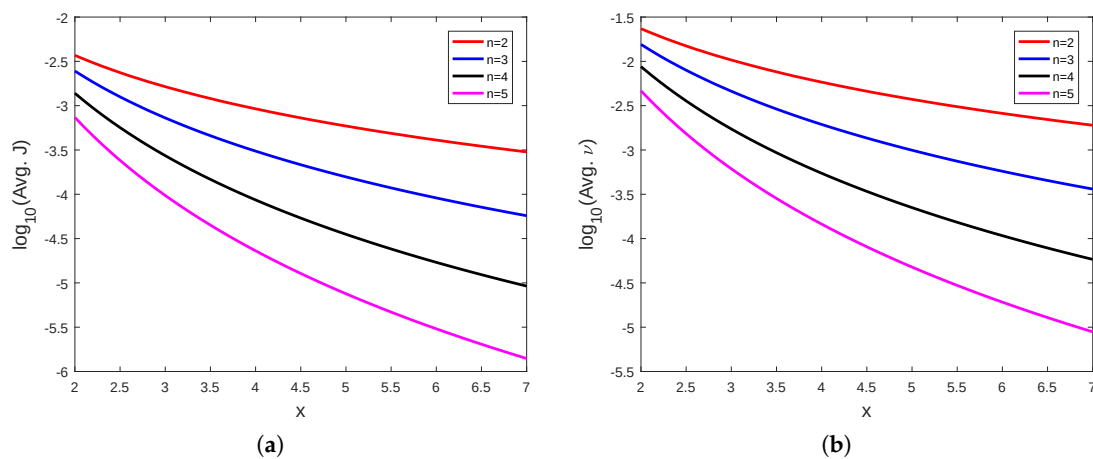


Figure 11. (a) Logarithm of Li's un-normalized MoN ($\log_{10}(\text{Avg. } J)$) vs. x and (b) logarithm of Li's normalized MoN ($\log_{10}(\text{Avg. } \nu)$) vs. x using 10 scalar measurements.

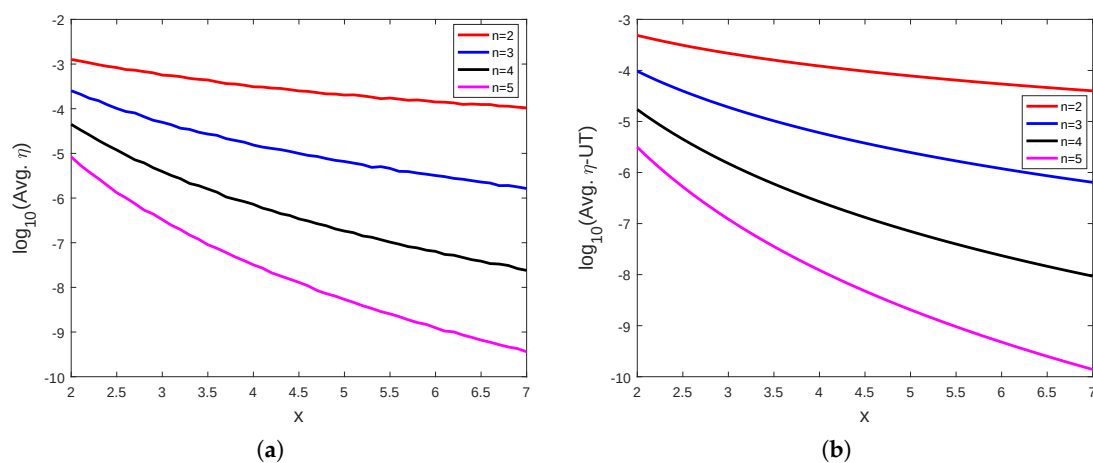


Figure 12. (a) Logarithm of MoN of Straka et al. ($\log_{10}(\text{Avg. } \eta)$) vs. x and (b) logarithm of MoN of Straka et al. with UT ($\log_{10}(\text{Avg. } \eta\text{-UT})$) vs. x using 10 scalar measurements.

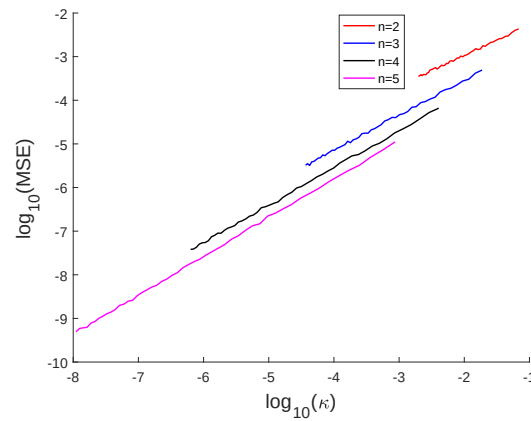


Figure 13. $\log_{10}(\text{MSE})$ vs. logarithm of extrinsic curvature ($\log_{10}(\kappa)$) using 10 scalar measurements.

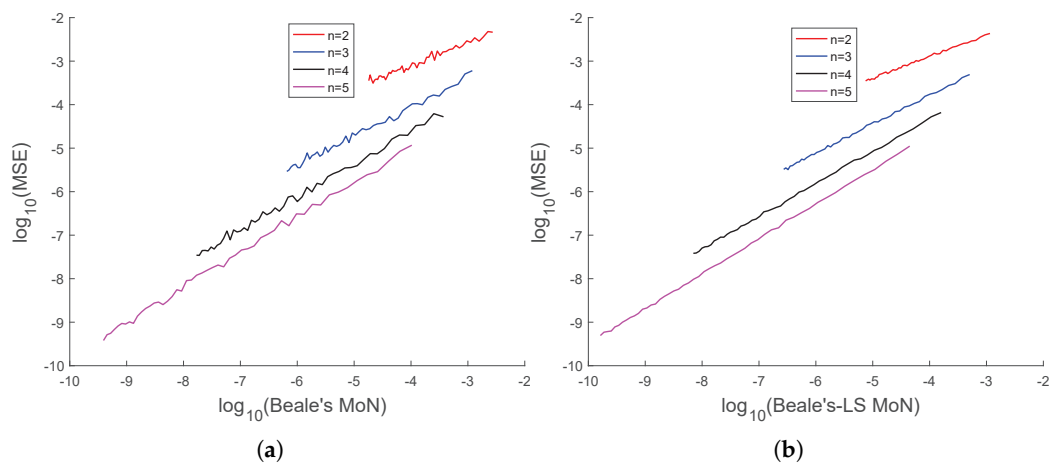


Figure 14. (a) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Avg. Beale's MoN})$ and (b) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Avg. Beale's MoN using LS})$ using 10 scalar measurements.

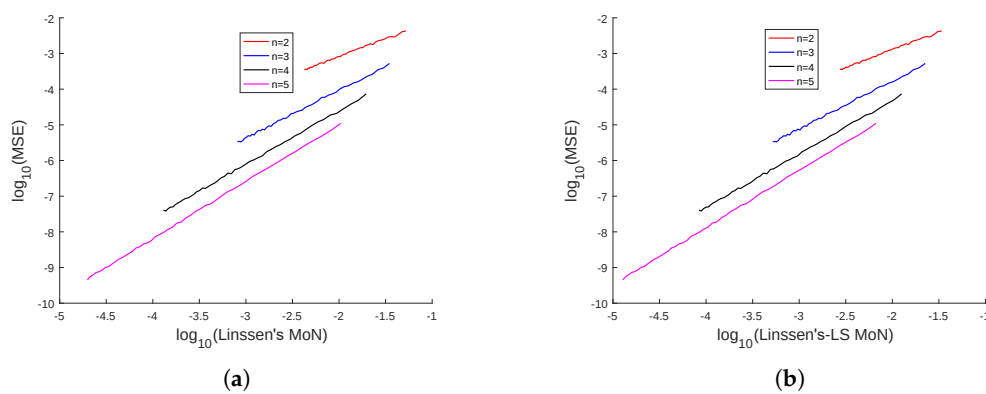


Figure 15. (a) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Linssen's MoN})$ and (b) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Linssen's-LS MoN})$ using 10 scalar measurements.

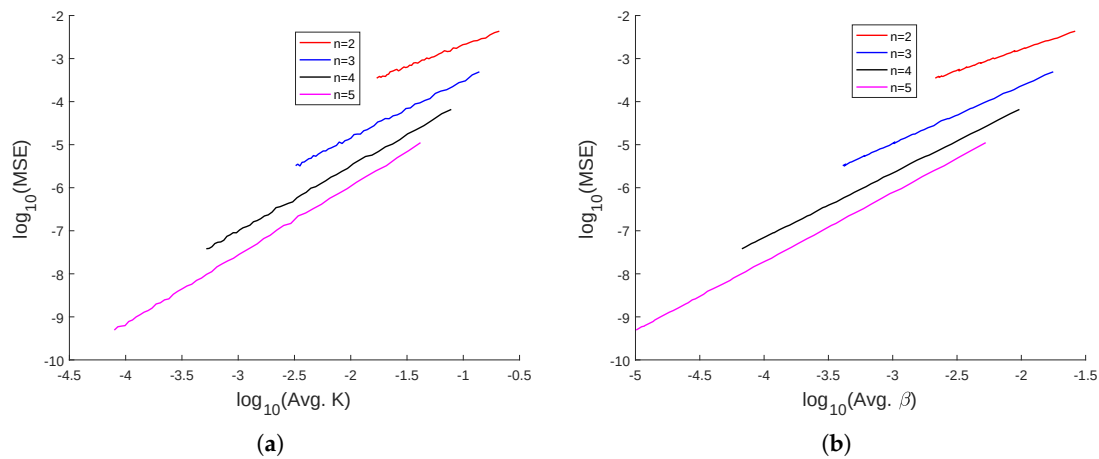


Figure 16. $\log_{10}(\text{MSE})$ vs. logarithm of parameter-effects curvatures. (a) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Avg. } K)$ and (b) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Avg. } \beta)$ using 10 scalar measurements.

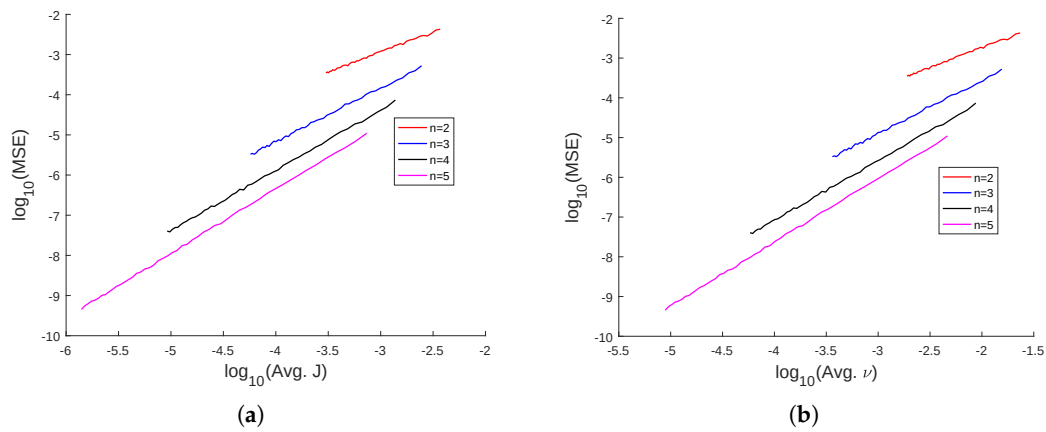


Figure 17. $\log_{10}(\text{MSE})$ vs. logarithm of Li's MoN. (a) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Avg. } J)$ and (b) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Avg. } \nu)$ using 10 scalar measurements.

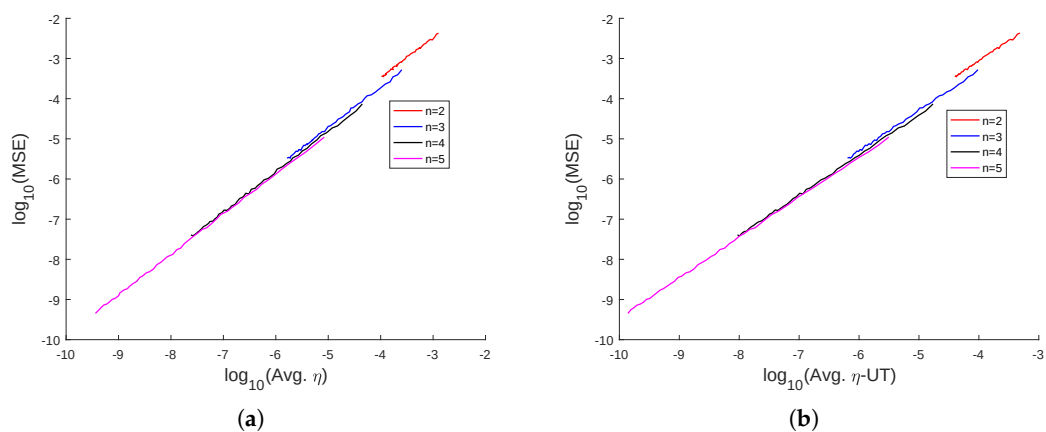


Figure 18. $\log_{10}(\text{MSE})$ vs. logarithm of MoN of Straka et al. (a) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Avg. } \eta)$ and (b) $\log_{10}(\text{MSE})$ vs. $\log_{10}(\text{Avg. } \eta\text{-UT})$ using 10 scalar measurements.

In [35], we had shown analytically, and through Monte Carlo simulation, that affine mappings exist among $\log_{10}(\text{MSE})$, $\log_{10}(\kappa)$, $\log_{10}(\text{Avg. } K)$, and $\log_{10}(\text{Avg. } \beta)$. In Figures 13–18, we have plotted the $\log_{10}(\text{MSE})$ versus $\log_{10}$ of various MoNs using 10 scalar measurements. These figures show that the $\log_{10}(\text{MSE})$ varies with $\log_{10}(\text{MoN})$ according to an affine mapping with a positive slope.

This implies that the MSE increases as an MoN increases. We obtain similar results for the case of 20 scalar measurements.

The above results demonstrate that, for the polynomial nonlinearity problem analyzed, any of the seven MoNs analyzed is suitable metrics to quantify the MSE, which represents the complexity of a parameter estimation problem. Further research is needed to study the applicability of these MoNs in real-world non-linear filtering problems.

6. Conclusions

We considered a polynomial curve in 2D and derived analytic expressions for the ML estimate and associated variance of the independent variable x using a vector measurement. The ML estimate is used to evaluate the Jacobian and Hessian of the measurement function appearing in the computation of Bates and Watts and direct parameter-effects curvatures, Beale's MoN, and Linssen's MoN. Our numerical results show that the variance of the estimated parameter and the Cramér-Rao lower bound (CRLB) are nearly the same for different powers of x . The average normalized estimation error squared (ANEES) lies within the 99% confidence interval, which indicates that the ML based variance is consistent with the estimation error.

We used seven MoNs, including the extrinsic curvature using differential geometry, Beale's MoN (and its least squares variant), Linssen's MoN (and its least squares variant), Bates and Watts parameter-effects curvature, direct parameter-effects curvature, Li's MoN, and the MoN of Straka, Duník, and Šimandl. If a MoN has a high value, then the nonlinearity is high. All of the MoNs show the same type of variation with x and the power of the polynomial. Secondly, as the logarithm of a MoN increases, the logarithm of the MSE also increases linearly for each MoN. This implies that, as a MoN increases, and then the MSE increases. These results are quite surprising, given the fact that these MoNs are derived based on completely different theoretical considerations. The second feature of our analysis is useful in establishing that a MoN in our study can be considered as a candidate metric for quantifying the MSE that represents the complexity of a parameter estimation problem. Our future work will study other practical parameter estimation and non-linear filtering problems.

Author Contributions: Formal analysis, M.M. and X.T.; methodology, M.M.; software, M.M. and X.T.; writing, M.M. and X.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The author thanks Sanjeev Arulampalam of Defence Science and Technology Organisation, Edinburgh SA, Australia for useful discussions and insightful comments in improving the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Anderson, B.D.O.; Moore, J.B. *Optimal Filtering*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1979.
2. Bar-Shalom, Y.; Li, X.; Kirubarajan, T. *Estimation with Applications to Tracking and Navigation*; Wiley: New York, NY, USA, 2001.
3. Gelb, A. (Ed.) *Applied Optimal Estimation*; MIT Press: Cambridge, MA, USA, 1974.
4. Jazwinski, A.H. *Stochastic Processes and Filtering Theory*; Academic Press: New York, NY, USA, 1970.
5. Aidala, V.J. Kalman Filter Behaviour in Bearings-only Tracking Applications. *IEEE Trans. Autom. Control* **1979**, *AES-15*, 29–39.
6. Aidala, V.J.; Hammel, S.E. Utilization of modified polar coordinates for bearings-only tracking. *IEEE Trans. Autom. Control* **1983**, *28*, 283–294. [[CrossRef](#)]
7. Arulampalam, S.; Ristic, B. Comparison of the particle filter with range-parameterised and modified polar EKFs for angle-only tracking. In Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets, Orlando, FL, USA, 24–28 April 2000; Volume 4048, pp. 288–299.
8. Ristic, B.; Arulampalam, S.; Gordon, N. *Beyond the Kalman Filter*; Artech House: Boston, MA, USA, 2004.

9. Mallick, M.; Arulampalam, S. Comparison of Nonlinear Filtering Algorithms in Ground Moving Target Indicator (GMTI) Target Tracking. In Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, USA, 5 January 2004; Volume 5204, pp. 630–647.
10. Blackman, S.S.; White, T.; Blyth, B.; Durand, C. Integration of Passive Ranging with Multiple Hypothesis Tracking (MHT) for Application with Angle-Only Measurements. In Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets, Orlando, FL, USA, 15 April 2010; Volume 7698, pp. 769815-1–769815-11.
11. Mallick, M.; Morelande, M.; Mihaylova, L.; Arulampalam, S.; Yan, Y. Comparison of Angle-only Filtering Algorithms in 3D using Cartesian and Modified Spherical Coordinates. In Proceedings of the 15th International Conference on Information Fusion, Singapore, 9–12 July 2012.
12. Mallick, M.; Morelande, M.; Mihaylova, L.; Arulampalam, S.; Yan, Y. Angle-only Filtering in Three Dimensions, Chapter 1. In *Integrated Tracking, Classification, and Sensor Management: Theory and Applications*; Mallick, M., Krishnamurthy, V., Vo, B.-N., Eds.; Wiley/IEEE: Hoboken, NJ, USA, 2013.
13. Athans, M.; Wishner, R.P.; Bertolini, A. Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements. *IEEE Trans. Autom. Control* **1968**, *13*, 504–518. [[CrossRef](#)]
14. Julier, S.; Uhlmann, J.; Durrant-Whyte, H.F. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Trans. Autom. Control* **2000**, *AC-45*, 477–482. [[CrossRef](#)]
15. Julier, S.J.; Uhlmann, J.K. Unscented filtering and nonlinear estimation. *Proc. IEEE* **2004**, *92*, 401–422. [[CrossRef](#)]
16. Arasaratnam, I.; Haykin, S. Cubature Kalman filters. *IEEE Trans. Autom. Control* **2009**, *54*, 1254–1269. [[CrossRef](#)]
17. Arulampalam, S.; Maskell, S.; Gordon, N.; Clapp, T. A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking. *IEEE Trans. Signal Process.* **2002**, *50*, 174–188. [[CrossRef](#)]
18. Beale, E.M.L. Confidence regions in non-linear estimation. *J. R. Stat. Soc. Ser. B* **1960**, *22*, 41–76. [[CrossRef](#)]
19. Guttman, I.; Meeter, D.A. On Beale's measures of non-linearity. *Technometrics* **1965**, *7*, 623–637. [[CrossRef](#)]
20. Linssen, H.N. Nonlinearity measures: A case study. *Stat. Neerl.* **1975**, *29*, 93–99. [[CrossRef](#)]
21. Bates, D.M.; Watts, D.G. Relative curvature measures of nonlinearity. *J. R. Stat. Society. Ser. B (Methodol.)* **1980**, *42*, 1–25. [[CrossRef](#)]
22. Bates, D.M.; Watts, D.G. Parameter transformations for improved approximate confidence regions in nonlinear least squares. *Ann. Stat.* **1981**, *9*, 1152–1167. [[CrossRef](#)]
23. Goldberg, M.L.; Bates, D.M.; Watts, D.G. Simplified Methods of Assessing Nonlinearity. *Am. Stat. Assoc. Proc. Bus. Eco. Statistics Section* **1983**, 67–74.
24. Bates, D.M.; Hamilton, D.C.; Watts, D.G. Calculation of intrinsic and parameter-effects curvatures for nonlinear regression models. *Commun. Stat. Simul. Comput.* **1983**, *12*, 469–477. [[CrossRef](#)]
25. Bates, D.M.; Watts, D.G. *Nonlinear Regression Analysis and its Applications*; John Wiley: New York, NY, USA, 1988.
26. Seber, G.A.; Wild, C.J. *Nonlinear Regression*; John Wiley: New York, NY, USA, 1989.
27. Mallick, M. Differential geometry measures of nonlinearity with application to ground target tracking. In Proceedings of the 7th International Conference on Information Fusion, Stockholm, Sweden, 28 June–1 July 2004.
28. Scala, B.F.; Mallick, M.; Arulampalam, S. Differential Geometry Measures of Nonlinearity for Filtering with Nonlinear Dynamic and Linear Measurement Models. In Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets, San Diego, CA, USA, 28–30 August 2007.
29. Mallick, M.; Scala, B.F.L.; Arulampalam, M.S. Differential geometry measures of nonlinearity for the bearing-only tracking problem. In Proceedings of the SPIE Conference on Signal Processing, Sensor Fusion, and Target Recognition XIV, Orlando, FL, USA, 25 May 2005; Volume 5809, pp. 288–300.
30. Mallick, M.; Scala, B.F.L.; Arulampalam, S. Differential geometry measures of nonlinearity with applications to target tracking. In Proceedings of the Fred Daum Tribute Conference, Monterey, CA, USA, 24 May 2007.
31. Duník, J.; Straka, O.; Mallick, M.; Blasch, E. Survey of nonlinearity and non-Gaussianity measures for state estimation. In Proceedings of the 19th International Conference on Information Fusion, Heidelberg, Germany, 5 July 2016; pp. 1845–1852.

32. Mallick, M.; Scala, B.F.L. Differential geometry measures of nonlinearity for ground moving target indicator (GMTI) filtering. In Proceedings of the 2005 8th International Conference on Information Fusion, Philadelphia, PA, USA, 25–28 July 2005; pp. 219–226.
33. Mallick, M.; Ristic, B. Comparison of Measures of Nonlinearity for Bearing-only and GMTI Filtering. In Proceedings of the 20th International Conference on Information Fusion, Xi'an, China, 10–13 July 2017.
34. Mallick, M.; Scala, B.F.L. Differential geometry measures of nonlinearity for the video tracking problem. In Proceedings of the SPIE Conference on Signal Processing, Sensor Fusion, and Target Recognition XV, Orlando (Kissimmee), FL, USA, 17 May 2006; Volume 6235, pp. 246–258.
35. Mallick, M.; Yan, Y.; Arulampalam, S.; Mallick, A. Connection between Differential Geometry and Estimation Theory for Polynomial Nonlinearity in 2D. In Proceedings of the 13th International Conference on Information Fusion, Edinburgh, UK, 26–29 July 2010.
36. Available online: <http://en.wikipedia.org/wiki/curvature> (accessed on 5 March 2020).
37. Carmo, M.D. *Differential Geometry of Curves and Surfaces*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1976.
38. Struik, D.J. *Lectures on Classical Differential Geometry*, 2nd ed.; Dover Publications: Mineola, NY, USA, 1988.
39. Mendel, J.M. *Lessons in Estimation Theory for Signal Processing Communications, and Control*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1995.
40. Mallick, M. Improvements to Measures of Nonlinearity Computation for Polynomial Nonlinearity. In Proceedings of the 8th International Conference on Control, Automation and Information Sciences (ICCAIS 2019), Chengdu, China, 23–26 October 2019.
41. Trees, H.L.V.; Bell, K.L.; Tian, Z. *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory*, 2nd ed.; Wiley: New York, NY, USA, 2013.
42. Li, X.R.; Zhao, Z.; Jilkov, V.P. Estimator's credibility and its measures. In Proceedings of the IFAC 15th World Congress, Barcelona, Spain, 21–26 July 2002.
43. Li, X.R. Measure of Nonlinearity for Stochastic Systems. In Proceedings of the 15th International Conference on Information Fusion, Singapore, 9–12 July 2012; pp. 1073–1080.
44. Liu, Y.; Li, X.R. Measure of Nonlinearity for Estimation. *IEEE Trans. Signal Process.* **2015**, *63*, 2377–2388. [[CrossRef](#)]
45. Sun, T.; Xin, M.; Jia, B. Nonlinearity-based adaptive sparse-grid quadrature filter. In Proceedings of the American Control Conference (ACC), Chicago, IL, USA, 1–3 July 2015; pp. 2499–2504.
46. Wang, P.; Blasch, E.; Li, X.R.; Jones, E.; Hanak, R.; Yin, W.; Beach, A.; Brewer, P. Degree of nonlinearity (DoN) measure for target tracking in videos. In Proceedings of the 9th International Conference on Information Fusion, Heidelberg, Germany, 5 July 2016; pp. 1390–1397.
47. Sun, T.; Xin, M. Hypersonic entry vehicle state estimation using nonlinearity-based adaptive cubature Kalman filters. *Acta Astronaut.* **2017**, *134*, 221–230. [[CrossRef](#)]
48. Duník, J.; Straka, O.; Šimandl, M. Nonlinearity and non-Gaussianity measures for stochastic dynamic systems. In Proceedings of the 16th International Conference on Information Fusion, Istanbul, Turkey, 9 July 2013; pp. 204–211.
49. Straka, O.; Duník, J.; Šimandl, M. Unscented Kalman filter with controlled adaptation. In Proceedings of the 16th IFAC Symposium on System Identification, Brussels, Belgium, 11–13 July 2012; pp. 906–911.
50. Lu, Q.; Bar-Shalom, Y.; Willett, P.; Zhou, S. Nonlinear observation models with additive Gaussian noises and efficient MLEs. *IEEE Signal Process. Lett.* **2017**, *24*, 545–549. [[CrossRef](#)]
51. Schweppe, F.C. *Uncertain Dynamic Systems*; Prentice Hall: Englewood Cliffs, NJ, USA, 1973.
52. Walpole, R.E.; Myers, R.H.; Myers, S.L.; Ye, K. *Probability and Statistics for Engineers and Scientists*, 9th ed.; Prentice Hall: Englewood Cliffs, NJ, USA, 2011.

