On Nonlinearity Measuring Aspects of Stochastic Integration Filter

Jindřich Havlík, Ondřej Straka, Jindřich Duník and Jiří Ajgl
European Centre of Excellence - New Technologies for Information Society,
Department of Cybernetics, University of West Bohemia, Technická 8, Pilsen, Czech Republic

Keywords: State Estimation, Bayesian Approach, Stochastic Integration, Measures of Nonlinearity, Measures of Non-Gaussianity.

Abstract: The paper deals with Bayesian state estimation of nonlinear stochastic dynamic systems. The focus is aimed at the stochastic integration filter, which is based on a stochastic integration rule. It is shown that the covariance matrix of the integration error calculated as a byproduct of the rule can be used as a measure of nonlinearity. The measure informs the user about validity of the assumptions of Gaussianity, which is adopted by the stochastic integration filter. It is also demonstrated how to use this information for a prediction of the number of remaining iterations of the rule. The paper also focuses on utilization of the integration error covariance matrix for improving estimates of the mean square error of the estimates, which is produced by the filter.

1 INTRODUCTION

Nonlinear state estimation of discrete-time stochastic dynamic systems is an important field of study, which has undergone a rapid development in the last two decades. The importance of this field stems from its crucial role in many areas such as signal processing, target tracking, satellite navigation, fault detection, and adaptive and optimal control problems. It is an essential part of any decision-making process.

The system state completely characterizes the system at a given time and usually is not directly measurable. The system behavior is described by a stochastic dynamic discrete-time model consisting of a stochastic difference equation characterizing dynamics of the state and by a stochastic algebraic equation representing a relation between the state and the measurement. The solution to the state estimation problem usually follows the Bayesian approach or the optimization approach.

The Bayesian approach is built up on the Bayesian recursive relations (BRRs) (Sorenson, 1974), which provide the state estimate in the form of a conditional probability density function (PDF) of the state conditioned by the measurement. However, the exact solution to the BRRs can be found only for a few special cases, such as a linear system with Gaussian disturbances. For a nonlinear or non-Gaussian system the conditional PDF has a complex shape and the solution usually cannot be found in a closed form. Hence, some approximations must be employed.

A group of approximate methods providing approximate conditional PDF capturing the complexity of the conditional PDF with a great fidelity are called global methods. They are represented namely by the Gaussian sum method (Ristic et al., 2004), the point-mass method (Kramer and Sorenson, 1988), and the particle filter (Doucet et al., 2001). The practical use of the global methods is limited especially due to their extensive computational complexity.

Another group of methods is formed by assuming Gaussian approximation of the conditional PDF provided by the BRRs. Such methods are denoted as Gaussian filters (GFs) (Ito and Xiong, 2000) or local Bayesian methods (Arasaratnam and Haykin, 2009). The GFs are represented namely by the quadrature Kalman filter (Arasaratnam et al., 2007), the cubature Kalman filter (Arasaratnam and Haykin, 2009), and the stochastic integration filter (SIF) (Duník et al., 2013b), which utilizes the iterative stochastic integration rule (SIR) (Genz and Monahan, 1998).

Contrary to the Bayesian approach, the optimization approach assumes minimization of a certain design criterion, usually the mean square error (MSE), leading to filters such as the extended Kalman filter, the divided difference filter (Nørgaard et al., 2000), or the unscented Kalman filter (Julier and Uhlmann, 2004). The optimization is usually feasible by enforcing a linear structure of the filter.

Even though optimization methods and local Bayesian methods were derived by different approaches, finally they provide *formally* same solution

Havlík, J., Straka, O., Duník, J. and Ajgl, J.
On Nonlinearity Measuring Aspects of Stochastic Integration Filter.
DOI: 10.5220/0005983903530361
In *Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2016) - Volume 1*, pages 353-361
ISBN: 978-989-758-198-4

to the state estimation problem, which is computationally reasonable and therefore often used in practice.

The GFs perform well only for mildly nonlinear functions, for which the Gaussianity assumption is approximately valid. For strong nonlinearities the assumption is only a rough approximation. During the last decade several methods quantifying nonlinearity or non-Gaussianity, which have been applied in monitoring the GF assumption validity, have appeared (Mallick, 2004), (Li, 2012) and (Duník et al., 2013a). Such methods are usually based on evaluating effects of nonlinearities in neighborhoods of conditional means or testing unimodal or non-heavy-tailed properties of the conditioned PDFs.

The goal of the paper is to show that the SIF based on SIRs possesses nonlinearity measuring capabilities itself. The information can be used to give the user an idea how strong nonlinear behavior the system exhibits or how much the Gaussian assumption is violated and how much the estimates should be trusted. Further, it will be shown that the information can be utilized to predict computational requirements of the SIRs employed by the filter to achieve a desired accuracy. And above all, the information will be used to make the estimate of the MSE provided by the SIF more precise.

The rest of the paper is organized as follows: Section 2 is devoted to a brief introduction to the nonlinear state estimation, its solution by the GFs, and measures of nonlinearity and non-Gaussianity. The SIR is briefly presented in Section 3. Section 4 presents the revealed capability of the SIF to measure nonlinearity and its application. Next, in Section 5 the utilization of this information to provide an improved estimate of the MSE is outlined and the paper is concluded by Section 6.

2 STATE ESTIMATION AND GAUSSIAN FILTERS

The aim of this section is to formulate the nonlinear state estimation problem, to present its general solution by means of the BRR, to describe the GFs, and to present two measures of nonlinearity and non-Gaussianity proposed in literature.

2.1 Formulation of the Nonlinear State Estimation Problem

Let a discrete-time nonlinear stochastic system be considered in the following state-space form

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{w}_k, \qquad k = 0,1,2,\ldots, \tag{1}$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \qquad k = 0,1,2,\ldots, \tag{2}$$

where the vectors $\mathbf{x}_k \in \mathbb{R}^{n_x}$, and $\mathbf{z}_k \in \mathbb{R}^{n_z}$ represent the state of the system, and the measurement at time instant k, respectively, $\mathbf{f}_k : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ and $\mathbf{h}_k : \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$ are known vector functions, and $\mathbf{w}_k \in \mathbb{R}^{n_x}$, $\mathbf{v}_k \in \mathbb{R}^{n_z}$ are mutually independent state and measurement white noises. The PDFs of the noises are Gaussian with zero means and known covariance matrices (CMs) $\Sigma_k^{\mathbf{w}}$ and $\Sigma_k^{\mathbf{v}}$, respectively, i.e., $p(\mathbf{w}_k) = \mathcal{N}\{\mathbf{w}_k; \mathbf{0}_{n_x\times 1}, \Sigma_k^{\mathbf{w}}\}$[1] and $p(\mathbf{v}_k) = \mathcal{N}\{\mathbf{v}_k; \mathbf{0}_{n_z\times 1}, \Sigma_k^{\mathbf{v}}\}$, respectively, where $\mathbf{0}_{a\times b}$ denotes an $a \times b$ matrix of zeros. The PDF of the initial state is Gaussian and known as well, i.e., $p(\mathbf{x}_0) = \mathcal{N}\{\mathbf{x}_0; \bar{\mathbf{x}}_0, \mathbf{P}_0\}$. The initial state is independent of the noises.

State estimation aims at searching the state $\mathbf{x}_k$ based on measurements up to the time instant ℓ, which will be denoted as $\mathbf{z}^\ell \triangleq [\mathbf{z}_0^T, \mathbf{z}_1^T, \ldots, \mathbf{z}_\ell^T]^T$. Due to the stochastic nature of the system, the state estimate is described by the conditional PDF $p(\mathbf{x}_k|\mathbf{z}^\ell)$. In this paper, the filtering ($k = \ell$) and the one-step prediction ($k = \ell + 1$) problems will be considered only.

To find the filtering estimate $p(\mathbf{x}_k|\mathbf{z}^k)$ the Bayesian approach is used, and the following BRRs provide the solution (Sorenson, 1974):

$$p(\mathbf{x}_k|\mathbf{z}^k) = \frac{p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)}{\int p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)\,\mathrm{d}\mathbf{x}_k}, \tag{3}$$

where the one-step prediction PDF is

$$p(\mathbf{x}_k|\mathbf{z}^{k-1}) = \int p(\mathbf{x}_{k-1}|\mathbf{z}^{k-1})p(\mathbf{x}_k|\mathbf{x}_{k-1})\,\mathrm{d}\mathbf{x}_{k-1}, \tag{4}$$

and $p(\mathbf{x}_k|\mathbf{x}_{k-1}) = p_{\mathbf{w}_{k-1}}(\mathbf{x}_k - \mathbf{f}_{k-1}(\mathbf{x}_{k-1}))$ and $p(\mathbf{z}_k|\mathbf{x}_k) = p_{\mathbf{v}_k}(\mathbf{z}_k - \mathbf{h}_k(\mathbf{x}_k))$.

An analytical solution to (3) and (4) is an intricate functional-domain problem. It can be analytically computed for a few cases only. Such a case is given by linear functions and Gaussian PDFs in the system equations (1) and (2). So, for a nonlinear or a non-Gaussian system, approximate solutions are mostly necessary.

2.2 Gaussian Filters

The GFs suppose that the joint prediction PDF $p(\mathbf{z}_k, \mathbf{x}_k|\mathbf{z}^{k-1})$ is at each time instant Gaussian (Arasaratnam and Haykin, 2009)

$$p(\mathbf{z}_k, \mathbf{x}_k|\mathbf{z}^{k-1}) = \mathcal{N}\left\{ \begin{bmatrix} \mathbf{z}_k \\ \mathbf{x}_k \end{bmatrix}; \begin{bmatrix} \bar{\mathbf{z}}_{k|k-1} \\ \bar{\mathbf{x}}_{k|k-1} \end{bmatrix}, \begin{bmatrix} \mathbf{P}^{\mathbf{zz}}_{k|k-1} & \mathbf{P}^{\mathbf{xz}^T}_{k|k-1} \\ \mathbf{P}^{\mathbf{xz}}_{k|k-1} & \mathbf{P}^{\mathbf{xx}}_{k|k-1} \end{bmatrix} \right\}. \tag{5}$$

[1] For the sake of simplicity all PDFs will be given by their argument, if not stated otherwise, i.e., $p(\mathbf{w}_k) = p_{\mathbf{w}_k}(\mathbf{w}_k)$.

Then the filtering PDF $p(\mathbf{x}_k|\mathbf{z}^k)$ and the one-step predictive PDF $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ are also Gaussian

$$p(\mathbf{x}_k|\mathbf{z}^k) = \mathcal{N}\{\mathbf{x}_k; \bar{\mathbf{x}}_{k|k}, \mathbf{P}^{\mathbf{xx}}_{k|k}\}, \quad (6)$$

$$p(\mathbf{x}_k|\mathbf{z}^{k-1}) = \mathcal{N}\{\mathbf{x}_k; \bar{\mathbf{x}}_{k|k-1}, \mathbf{P}^{\mathbf{xx}}_{k|k-1}\}. \quad (7)$$

The GFs compute the first two moments of (6) and (7), i.e., the conditional means $\bar{\mathbf{x}}_{k|k} = \mathsf{E}[\mathbf{x}_k|\mathbf{z}^k]$ and $\bar{\mathbf{x}}_{k|k-1} = \mathsf{E}[\mathbf{x}_k|\mathbf{z}^{k-1}]$ and the conditional CMs $\mathbf{P}^{\mathbf{xx}}_{k|k} = \mathsf{cov}[\mathbf{x}_k|\mathbf{z}^k]$ and $\mathbf{P}^{\mathbf{xx}}_{k|k-1} = \mathsf{cov}[\mathbf{x}_k|\mathbf{z}^{k-1}]$.

The GF algorithm can be written in the following form:

Algorithm 1: Gaussian Filter.

Step 1: (*initialization*) Set the time instant $k = 0$ and define a priori initial condition by its first two moments

$$\bar{\mathbf{x}}_{0|-1} \triangleq \mathsf{E}[\mathbf{x}_0] = \bar{\mathbf{x}}_0, \quad (8)$$

$$\mathbf{P}^{\mathbf{xx}}_{0|-1} \triangleq \mathsf{E}[(\mathbf{x}_0 - \bar{\mathbf{x}}_{0|-1})(\mathbf{x}_0 - \bar{\mathbf{x}}_{0|-1})^T] = \mathbf{P}_0. \quad (9)$$

Step 2: (*filtering, measurement update*) The filtering mean $\bar{\mathbf{x}}_{k|k}$ and CM $\mathbf{P}^{\mathbf{xx}}_{k|k}$ are computed by the means of

$$\bar{\mathbf{x}}_{k|k} = \bar{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \bar{\mathbf{z}}_{k|k-1}), \quad (10)$$

$$\mathbf{P}^{\mathbf{xx}}_{k|k} = \mathbf{P}^{\mathbf{xx}}_{k|k-1} - \mathbf{K}_k\mathbf{P}^{\mathbf{zz}}_{k|k-1}\mathbf{K}_k^T, \quad (11)$$

where

$$\mathbf{K}_k = \mathbf{P}^{\mathbf{xz}}_{k|k-1}(\mathbf{P}^{\mathbf{zz}}_{k|k-1})^{-1} \quad (12)$$

is the filter gain and the measurement prediction $\bar{\mathbf{z}}_{k|k-1}$ is given by

$$\bar{\mathbf{z}}_{k|k-1} = \mathsf{E}[\mathbf{z}_k|\mathbf{z}^{k-1}]. \quad (13)$$

The predictive CMs $\mathbf{P}^{\mathbf{xz}}_{k|k-1}$ and $\mathbf{P}^{\mathbf{zz}}_{k|k-1}$ are computed as

$$\begin{aligned}\mathbf{P}^{\mathbf{zz}}_{k|k-1} &= \mathsf{E}[(\mathbf{z}_k - \bar{\mathbf{z}}_{k|k-1})(\mathbf{z}_k - \bar{\mathbf{z}}_{k|k-1})^T|\mathbf{z}^{k-1}] \\ &= \mathsf{E}[(\mathbf{h}_k(\mathbf{x}_k) - \bar{\mathbf{z}}_{k|k-1}) \\ &\quad \times (\mathbf{h}_k(\mathbf{x}_k) - \bar{\mathbf{z}}_{k|k-1})^T|\mathbf{z}^{k-1}] + \Sigma^{\mathbf{v}}_k, \quad (14)\end{aligned}$$

$$\mathbf{P}^{\mathbf{xz}}_{k|k-1} = \mathsf{E}[(\mathbf{x}_k - \bar{\mathbf{x}}_{k|k-1})(\mathbf{z}_k - \bar{\mathbf{z}}_{k|k-1})^T|\mathbf{z}^{k-1}]. \quad (15)$$

Step 3: (*prediction, time update*) The predictive mean $\bar{\mathbf{x}}_{k+1|k}$ and CM $\mathbf{P}^{\mathbf{xx}}_{k+1|k}$ are given by

$$\bar{\mathbf{x}}_{k+1|k} = \mathsf{E}[\mathbf{x}_{k+1}|\mathbf{z}^k] = \mathsf{E}[\mathbf{f}_k(\mathbf{x}_k)|\mathbf{z}^k], \quad (16)$$

$$\begin{aligned}\mathbf{P}^{\mathbf{xx}}_{k+1|k} &= \mathsf{E}[(\mathbf{x}_{k+1|k} - \bar{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1|k} - \bar{\mathbf{x}}_{k+1|k})^T|\mathbf{z}^k] \\ &= \mathsf{E}[(\mathbf{f}_k(\mathbf{x}_k) - \bar{\mathbf{x}}_{k+1|k})(\mathbf{f}_k(\mathbf{x}_k) - \bar{\mathbf{x}}_{k+1|k})^T|\mathbf{z}^k] \\ &\quad + \Sigma^{\mathbf{w}}_k. \quad (17)\end{aligned}$$

Let $k = k+1$.
The algorithm then continues by **Step 2**.

The particular GFs vary in the way they compute the integrals determining the predictive characteristics of the measurement (13-15) and the predictive characteristics of the state (16, 17). Mostly, the integrals cannot be calculated analytically. For example, the integrals can be approximated by the SIR, which leads to the SIF.

2.3 Measures of Nonlinearity and Non-Gaussianity

The GFs are designed under the Gaussianity assumption (5), which will become invalid due to the nonlinearities present in the system. This may consequently severely degrade the performance of the filter, disrupt its accuracy or credibility. Hence, it is convenient to monitor working conditions of the filter, which affect validity of the assumption. For this purpose the measures of nonlinearity (MoNL) or non-Gaussianity (MoNG) render an applicable tool. The MoNL and MoNG may be used for sequential monitoring of the GF assumption and an impact of approximations along a trajectory of the GF.

Two examples of MoNL, which will be later used for a comparison, are described below. The described MoNL evaluate the nonlinearity of a function $\mathbf{y} = \mathbf{g}(\mathbf{x})$, where $\mathbf{x}$ is a random variable with a mean $\bar{\mathbf{x}}$ and a CM $\mathbf{P}^{\mathbf{xx}}$ defining a region, where the nonlinearity will be analyzed. The MoNL can be used in the filtering or the predictive step of the GF (**Algorithm 1**). In the filtering step the function $\mathbf{g}(\mathbf{x})$ is equal to is $\mathbf{h}_k(\mathbf{x}_k)$, i.e., $\mathbf{g}(\mathbf{x}) = \mathbf{h}_k(\mathbf{x}_k)$, and in the predictive step $\mathbf{g}(\mathbf{x}) = \mathbf{f}_k(\mathbf{x}_k)$.

Differential Geometry Measure based MoNL

The MoNL based on a differential geometry measure (DGM) was originally proposed in (Bates and Watts, 1988) for the nonlinear parameter estimation and later it was extended for the nonlinear state estimation (Mallick, 2004). The MoNL directly assesses the curvature[2] of the nonlinear function at a given point. The state-independent version of the measure is defined as

$$\mathrm{MoNL}_i^{\mathrm{DGM}} = \frac{\|\mathcal{X}_i^T\ddot{\mathbf{G}}\mathcal{X}_i\|}{\|\dot{\mathbf{G}}\mathcal{X}_i\|}, \quad (18)$$

where $\mathcal{X}_i \in \mathbb{R}^{n_x}$, $\{\mathcal{X}_i\}_{i=1}^q$ is a set of q user-defined points somehow respecting the probabilistic description of $\mathbf{x}$, $\dot{\mathbf{G}} = \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}|_{\mathbf{x}=\bar{\mathbf{x}}}$ is the Jacobian of $\mathbf{g}(\mathbf{x})$, and $\ddot{\mathbf{G}} = \frac{\partial^2 \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}\partial \mathbf{x}}|_{\mathbf{x}=\bar{\mathbf{x}}}$ is the Hessian of $\mathbf{g}(\mathbf{x})$.

Therefore, the measure assesses the contribution of the Taylor series expansion (TSE) second order

[2]Amount by which a curve deviates from a tangent (hyper-)plane.

term[3] relative to the TSE first order term which defines the tangential (linear) approximation of the function. The measure (18) can be evaluated at multiple points $\mathcal{X}_i$ and, its maximum can be selected for assessing nonlinearity of $\mathbf{g}(\mathbf{x})$, i.e.,

$$\mathrm{MoNL}^{\mathrm{DGM}} = \max_{i\in\{1,2,\ldots,q\}} \mathrm{MoNL}_i^{\mathrm{DGM}}. \qquad (19)$$

Least-Squares based MoNL

In (Duník et al., 2013a) the weighted least-squares (WLS) based MoNL was proposed assessing the WLS residue, i.e., the difference between the nonlinear function and its linear (least-squares) approximation defined by parameters

$$\theta^* = \arg\min_{\theta} (\mathbf{Y} - \mathbf{X}\theta)^T \mathbf{W} (\mathbf{Y} - \mathbf{X}\theta), \qquad (20)$$

where $\mathbf{X} = \begin{bmatrix} [\mathcal{X}_0, \ldots, \mathcal{X}_q] \\ 1, \ldots, 1 \end{bmatrix}^T$, $\mathbf{Y} = [\mathcal{Y}_0, \ldots, \mathcal{Y}_q]^T$, $\mathcal{Y}_i = \mathbf{g}(\mathcal{X}_i), \forall i$, $\mathcal{X}_i \in \mathbb{R}^{n_x}$, $\mathcal{Y}_i \in \mathbb{R}^{n_y}$, $\{\mathcal{X}_i\}_{i=1}^q$ is a set of user-defined points, and $\mathbf{W}$ is a diagonal weighting matrix respecting placement of the points $\mathcal{X}_i$. Then, the MoNL is given by

$$\mathrm{MoNL}^{\mathrm{WLS}} = \mathbf{Y}^T [\mathbf{W} - \mathbf{W}\mathbf{X}(\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}]\mathbf{Y}. \qquad (21)$$

The points $\{\mathcal{X}_i\}_{i=1}^q$ are suitably selected (user-defined) weighted points to represent the probabilistic description of $\mathbf{x}$. The largest element can be selected for assessing of nonlinearity $\mathbf{g}(\mathbf{x})$.

3 STOCHASTIC INTEGRATION FILTER

The SIF (Duník et al., 2013b) utilizes the SIR (Genz and Monahan, 1998) for calculating the integrals arising in the GF algorithm (13)-(17). The main advantage of the SIR is that it provides faster rate of convergence than the simple Monte Carlo (MC) integration rule (Genz and Monahan, 1998). The SIF takes this advantage from the SIR and guarantees asymptotically exact or in some cases (linear or polynomial functions $\mathbf{f}_k$ and $\mathbf{h}_k$, depending on the SIR degree) exact calculation of the integrals. The integrals are calculated utilizing approximate description of a random variable representing a state estimate by a set of points and the corresponding weights.

[3]The Hessian can be further decomposed into the component in the tangential plane and the perpendicular component (Mallick, 2004).

The SIF is given by the algorithm of the GF (**Algorithm 1**), where the integrals (13)-(17) are approximately computed using SIR. **Algorithm 2** illustrates the degree-3 SIR for computing an approximate value of the integral

$$\mathbf{I}(\gamma) = \mathrm{E}[\gamma] = \int \gamma(\mathbf{x}) \mathcal{N}(\mathbf{x}; \bar{\mathbf{x}}, \mathbf{P}^{\mathbf{xx}}) \mathrm{d}\mathbf{x}. \qquad (22)$$

for a given function $\mathbf{y} = \gamma(\mathbf{x}), \gamma : \mathbb{R}^{n_x} \to \mathbb{R}^{n_\gamma}$, a mean value $\bar{\mathbf{x}}$ and a covariance matrix $\mathbf{P}^{\mathbf{xx}}$ of $\mathbf{x}$.

Algorithm 2: Degree-3 Stochastic Integration Rule.

Step 1: Select a maximum number of iterations N_{max} or an error tolerance ε.

Step 2: Set the current iteration number $N = 0$, initial value of the integral $\hat{\mathbf{I}}_0(\gamma) = \mathbf{0}_{n_\gamma \times 1}$, initial square error of the integral $\mathbf{V}_0 = \mathbf{0}_{n_\gamma \times n_\gamma}$, and set $\zeta_0 = \bar{\mathbf{x}}$.

Step 3: Repeat (until $N = N_{max}$ or $\max \mathbf{V}_N < \varepsilon^2$)

a) Set $N = N + 1$.

b) Generate a uniformly random orthogonal matrix $\mathbf{Q}_N$ of dimension $n_x \times n_x$ and generate a random number ρ_N from the Chi distribution with $(n_x + 2)$ degrees of freedom, i.e., $\rho_N \sim \mathrm{Chi}(n_x + 2)$.

c) Compute a set of points $\{\zeta_i\}_{i=0}^{2n_x}$ and appropriate weights $\{\omega_i\}_{i=0}^{2n_x}$ according to

$$\zeta_i = \bar{\mathbf{x}} - \rho_N \mathbf{S}^{\mathbf{xx}} \mathbf{Q}_N \mathbf{e}_i, \qquad (23)$$

$$\zeta_{n_x+i} = \bar{\mathbf{x}} + \rho_N \mathbf{S}^{\mathbf{xx}} \mathbf{Q}_N \mathbf{e}_i, \qquad (24)$$

$$\omega_0 = 1 - \frac{n_x}{\rho_N^2}, \ \omega_i = \omega_{n_x+i} = \frac{1}{2\rho_N^2}, \qquad (25)$$

where $i = 1, 2, \ldots, n_x$, $\mathbf{e}_i$ is the i-th column of the $n_x \times n_x$ identity matrix, and $\mathbf{S}^{\mathbf{xx}}$ is a decomposition of matrix $\mathbf{P}^{\mathbf{xx}}$ so that $\mathbf{P}^{\mathbf{xx}} = \mathbf{S}^{\mathbf{xx}} [\mathbf{S}^{\mathbf{xx}}]^T$.

d) Compute the following relations for the approximation of the integral value at current iteration $\mathrm{SR}^{(3)}$ (spherical-radial degree-3 SIR), the updated integral value $\hat{\mathbf{I}}_N(\gamma)$, and the corresponding mean square error estimate $\mathbf{V}_N$, i.e.,

$$\mathrm{SR}^{(3)}(\mathbf{Q}_N, \rho_N) = \sum_{i=0}^{2n_x} \omega_i \gamma(\zeta_i), \qquad (26)$$

$$\mathbf{D} = \frac{\mathrm{SR}^{(3)}(\mathbf{Q}_N, \rho_N) - \hat{\mathbf{I}}_{N-1}(\gamma)}{N}, \qquad (27)$$

$$\hat{\mathbf{I}}_N(\gamma) = \hat{\mathbf{I}}_{N-1}(\gamma) + \mathbf{D}, \qquad (28)$$

$$\mathbf{V}_N = \frac{N-2}{N} \mathbf{V}_{N-1} + \mathbf{D}\mathbf{D}^T. \qquad (29)$$

Step 4: Once the stopping conditions are fulfilled, the approximate value of the integral $\mathbf{I}(\gamma)$ is $\hat{\mathbf{I}}_N(\gamma)$.

Note that the maximum in the stopping condition $\max \mathbf{V}_N < \varepsilon^2$ denotes the maximum over all elements

of $\mathbf{V}_N$ and that the points ζ_i and weights ω_i are specified to satisfy

$$\sum_{i=0}^{2n_x} \omega_i \zeta_i = \bar{\mathbf{x}}, \tag{30}$$

$$\sum_{i=0}^{2n_x} \omega_i (\zeta_i - \bar{\mathbf{x}})(\zeta_i - \bar{\mathbf{x}})^T = \mathbf{P}^{\mathbf{xx}}. \tag{31}$$

The SIF of higher and lower degrees can be found in (Genz and Monahan, 1998).

Algorithm 2 is used for approximative calculations of (13)-(17) using the SIR, i.e., it produces estimates of $\bar{\mathbf{z}}_{k|k-1}$, $\mathbf{P}^{\mathbf{zz}}_{k|k-1}$, $\mathbf{P}^{\mathbf{xz}}_{k|k-1}$, $\bar{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}^{\mathbf{xx}}_{k|k-1}$. The SIR estimates will be further denoted as the original quantity with an extra hat over them, i.e., for example $\hat{\bar{\mathbf{z}}}_{k|k-1}$ is an estimate of $\bar{\mathbf{z}}_{k|k-1}$ produced by the SIR of arbitrary degree and stopping condition. Analogously, $\hat{\mathbf{P}}^{\mathbf{xz}}_{k|k-1}$ is an estimate of $\mathbf{P}^{\mathbf{xz}}_{k|k-1}$. The differences between the SIR estimates and original quantities will be denoted as the quantity with tilde, i.e.,

$$\tilde{\mathbf{z}}_{k|k-1} \triangleq \bar{\mathbf{z}}_{k|k-1} - \hat{\bar{\mathbf{z}}}_{k|k-1}$$

or analogously $\tilde{\mathbf{P}}^{\mathbf{xz}}_{k|k-1} \triangleq \mathbf{P}^{\mathbf{xz}}_{k|k-1} - \hat{\mathbf{P}}^{\mathbf{xz}}_{k|k-1}$. In the following sections an arbitrary degree SIR with an arbitrary stopping condition is assumed. For this reason, the notation $\mathrm{SR}(\mathbf{Q}_N, \rho_N)$ will be used without the superscript indicating the rule degree.

4 MEASURING NONLINEARITY BY SIR

This section will analyze the byproduct of the SIR calculation - the square error $\mathbf{V}_N$ of the integral approximate value $\hat{\mathbf{I}}_N(\gamma)$, which has been overlooked in SIF design so far. First, certain aspects of **Algorithm 2** will be clarified.

4.1 Square Error of Integral Value

The GFs require computation of the integral values in the form of (22). The SIR of arbitrary degree (for example **Algorithm 2** for degree-3) approximates this integral by a sample mean of spherical-radial rules

$$\hat{\mathbf{I}}_N(\gamma) = \frac{1}{N} \sum_{j=1}^{N} \mathrm{SR}(\mathbf{Q}_j, \rho_j), \tag{32}$$

which are given as a sum of n_p weighted semi-random points[4]

$$\mathrm{SR}(\mathbf{Q}_j, \rho_j) = \sum_{i=1}^{n_p} \omega_i \gamma(\zeta_i(\mathbf{Q}_j, \rho_j)), \tag{33}$$

where the notation $\zeta_i(\mathbf{Q}_j, \rho_j)$ was used to explicitly denote dependence of ζ_i on $\mathbf{Q}_j$ and ρ_j. The randomization is handled using

i) the random orthogonal matrix $\mathbf{Q}_j$, which governs rotation in the state space,

ii) the scaling parameters, i.e., $\rho_j \sim \mathrm{Chi}(n_x + 2)$ for the degree-3 rule, which govern spread of the points ζ_i.

The SIR also computes an estimate $\mathbf{V}_N$ of the MSE of the integral value approximation $\hat{\mathbf{I}}_N(\gamma)$. The error of $\hat{\mathbf{I}}_N(\gamma)$ denoted as $\tilde{\mathbf{I}}_N(\gamma) = \mathbf{I}_N - \hat{\mathbf{I}}_N$ has zero mean and its MSE is

$$\begin{aligned} \mathrm{E}\left[\tilde{\mathbf{I}}_N(\gamma)\tilde{\mathbf{I}}_N(\gamma)^T\right] &= \mathsf{var}[\tilde{\mathbf{I}}_N] = \mathsf{var}\left(\hat{\mathbf{I}}_N\right) \\ &= \frac{1}{N}\mathsf{var}\left(\mathrm{SR}(\mathbf{Q}, \rho)\right). \end{aligned} \tag{34}$$

Its estimate by the SIR is

$$\mathbf{V}_N = \frac{1}{N}\hat{\mathsf{var}}\left(\mathrm{SR}(\mathbf{Q}, \rho)\right), \tag{35}$$

where $\hat{\mathsf{var}}(\mathrm{SR}(\mathbf{Q}, \rho))$ is calculated as the sample variance of $\mathrm{SR}(\mathbf{Q}, \rho)$ as

$$\mathbf{V}_N = \frac{1}{N}\frac{1}{N-1}\sum_{i=1}^{N}\Big(\mathrm{SR}(\mathbf{Q}_i, \rho_i) - \hat{\mathbf{I}}_N(\gamma)\Big)\Big(\cdot\Big)^T. \tag{36}$$

The notation $(\alpha)(\cdot)^T$ stands for $(\alpha)(\alpha)^T$.

Theorem. *For a linear function* $\gamma(\mathbf{x}) = \Gamma \cdot \mathbf{x}$*, where* $\Gamma \in \mathbb{R}^{n_\gamma \times n_x}$*, arbitrary mean value* $\bar{\mathbf{x}}$*, a CM* $\mathbf{P}^{\mathbf{xx}}$ *and* $\forall N > 1$*, it holds that* $\mathsf{var}(\mathrm{SR}(\mathbf{Q}, \rho)) = 0$.

Proof. For arbitrary rotation matrix $\mathbf{Q}_j$ and scaling parameter ρ_j, the arbitrary degree spherical-radial SIR is

$$\begin{aligned} \mathrm{SR}(\mathbf{Q}_j, \rho_j) &= \sum_i \omega_i \gamma(\zeta_i(\mathbf{Q}_j, \rho_j)) \\ &= \sum_i \omega_i \Gamma \zeta_i(\mathbf{Q}_j, \rho_j) \\ &= \Gamma \sum_i \omega_i \zeta_i(\mathbf{Q}_j, \rho_j) \\ &= \Gamma \bar{\mathbf{x}}, \end{aligned} \tag{37}$$

where the relation (30) was used. As $\mathrm{SR}(\mathbf{Q}_j, \rho_j)$ given by (37) does not depend on $\mathbf{Q}_j$ and ρ_j for a linear function γ, it holds that $\mathsf{var}(\mathrm{SR}(\mathbf{Q}, \rho)) = 0$. □

[4]Number of points is governed by the degree of the SIR. For example, $n_p = 2n_x + 1$ for the degree-3 rule.

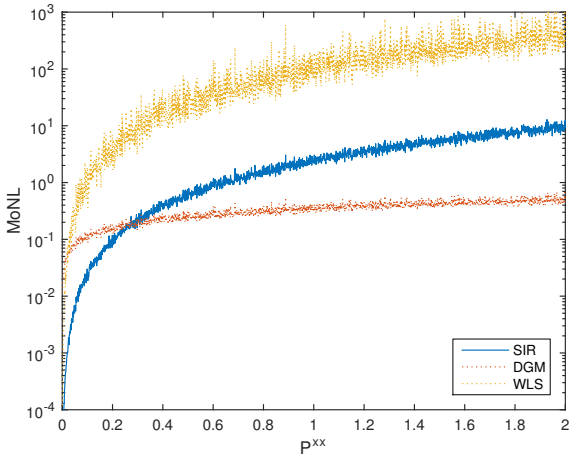

Figure 1: Dependency of MoNLs on growing variance P^{xx} for $\gamma(x)$.

As a consequence of (35) the SIR algorithm provides a quantity for an estimation of the variance of calculated SRs

$$\mathrm{var}(\mathrm{SR}(\mathbf{Q},\rho)) \approx N \cdot \mathbf{V}_N. \tag{38}$$

4.2 Relation of Square Error of Estimated Integral with MoNL

The variance of the calculated SRs, $\mathrm{var}(\mathrm{SR}(\mathbf{Q},\rho))$ approximately given by $N \cdot \mathbf{V}_N$, can be viewed as a measure of nonlinearity, because it is zero for a linear function and there is a reason to believe that for a growing strength of nonlinearity of γ, the variance grows as well. The quantity $N \cdot \mathbf{V}_N$ will be denoted as $\mathrm{MoNL}^{\mathrm{SIR}}$. An illustration follows to support this belief.

To facilitate analyzing of behavior of $\mathrm{MoNL}^{\mathrm{SIR}}$ for different nonlinearities, the following approach will be adopted: Instead of changing the function γ for a fixed $\bar{\mathbf{x}}$ and $\mathbf{P}^{\mathbf{xx}}$, the nonlinear function γ will be fixed and the variance $\mathbf{P}^{\mathbf{xx}}$ will change. The variance $\mathbf{P}^{\mathbf{xx}}$ governs the region over which the nonlinearity is analyzed and hence for growing $\mathbf{P}^{\mathbf{xx}}$, i.e., enlarging the region, the nonlinearity should become stronger.

The following static example has been chosen: Assume the function of interest $y = \gamma(x) = x^2$, the random variable x is given by mean value $\bar{x} = 0.1$ and variance $P^{xx} \in (0,2)$. The SIR was used to calculate $\mathrm{MoNL}^{\mathrm{SIR}}$ with the parameter $N_{max} = 100$. The $\mathrm{MoNL}^{\mathrm{DGM}}$ and $\mathrm{MoNL}^{\mathrm{WLS}}$ used the same ζ-points to calculate the value of the measure. In case of the $\mathrm{MoNL}^{\mathrm{WLS}}$, the weighting matrix was chosen as the eye matrix. The Figure 1 depicts the results of the experiment. All three measures inform the user of rising influence of nonlinearity for growing variance P^{xx}.

4.3 Prediction of Number of Iterations in SIR

In some cases, the user may specify the stopping condition in **Algorithm 2** by the error tolerance ε only, i.e., without specification of the maximum number of iterations $N_{\max}$. Throughout the iterations, it is possible to predict the number of steps required to achieve all elements of $\mathbf{V}_N$ to be lower than ε^2. The procedure is as follows: At each iteration N the variance $\mathrm{var}(\mathrm{SR}(\mathbf{Q},\rho))$ can be estimated by $N \cdot \mathbf{V}_N$. Consider N_{TOT} being a number of iterations such that after N_{TOT} steps the variance $\mathbf{V}_{N_{\mathrm{TOT}}}$ of $\hat{\mathbf{I}}_{N_{\mathrm{TOT}}}$ error will be approximately

$$\frac{1}{N_{\mathrm{TOT}}}\mathrm{var}(\mathrm{SR}(\mathbf{Q},\rho)) \approx \frac{1}{N_{\mathrm{TOT}}} \cdot N \cdot \mathbf{V}_N \triangleq \hat{\mathbf{V}}_{N_{\mathrm{TOT}}} \tag{39}$$

and the maximum element of $\hat{\mathbf{V}}_{N_{\mathrm{TOT}}}$ will be smaller than ε^2, i.e., $\max(\hat{\mathbf{V}}_{N_{\mathrm{TOT}}}) < \varepsilon^2$. Hence

$$\begin{aligned} \max \hat{\mathbf{V}}_{N_{\mathrm{TOT}}} = \max\left(\frac{1}{N_{\mathrm{TOT}}} \cdot N \cdot \mathbf{V}_N\right) &< \varepsilon^2 \\ \frac{N}{N_{\mathrm{TOT}}} \max \mathbf{V}_N &< \varepsilon^2 \\ \frac{\varepsilon^2}{N \cdot \max \mathbf{V}_N} &< N_{\mathrm{TOT}}. \end{aligned} \tag{40}$$

Thus the total number of steps can be estimated by

$$N_{\mathrm{TOT}} = \left\lceil \frac{\varepsilon^2}{N \cdot \max \mathbf{V}_N} \right\rceil, \tag{41}$$

where $\lceil \cdot \rceil$ is the ceil function rounding to the nearest larger integer.

The total number of steps N_{TOT} gives the user the information about how long it will take to obtain the integral value $\hat{\mathbf{I}}_N$ with the desired accuracy ε.

5 MSE ESTIMATED BY SIF

In this section the SIF will be analyzed in the terms of the MSE and based on the findings from the previous section and calculation of the MSE estimate will be proposed.

5.1 Theoretical values of MSE

The GFs compute the CM $\mathbf{P}^{\mathbf{xx}}_{k|k}$, which can be seen as an estimate of the MSE, i.e.,

$$\mathbf{P}^{\mathbf{xx}}_{k|k} \approx \mathrm{MSE} = \mathrm{E}\left[(\mathbf{x}_k - \bar{\mathbf{x}}_{k|k})(\cdot)^T\right]. \tag{42}$$

Now, the true MSE will be calculated and compared with the CM $\hat{\mathbf{P}}^{\mathbf{xx}}_{k|k}$ calculated by the SIF.

For convenience, consider the estimate $\bar{\mathbf{x}}_{k|k}$ (10) as an affine function of $\mathbf{z}_k$, i.e.,

$$\bar{\mathbf{x}}_{k|k} = \mathbf{A}\mathbf{z}_k + \mathbf{b}, \tag{43}$$

where

$$\mathbf{A} = \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}, \tag{44}$$
$$\mathbf{b} = \bar{\mathbf{x}} - \mathbf{A}\hat{\bar{\mathbf{z}}}. \tag{45}$$

Note that for clarity purposes the subscript $k|k-1$ will be omitted for the predictive moments.

Substituting (43) into (42) yields

$$\begin{aligned}\mathrm{E}\left[(\mathbf{x}_k - \bar{\mathbf{x}}_{k|k})(\cdot)^T\right] &= \underbrace{\mathrm{E}[(\mathbf{x}_k - \mathbf{A}\mathbf{z}_k - \mathbf{b})(\cdot)^T]}_{\mathrm{E}[\alpha\alpha^T]} = \\ &= \underbrace{\mathrm{E}\left[(\mathbf{x}_k - \bar{\mathbf{x}} - \mathbf{A}(\mathbf{z}_k - \bar{\mathbf{z}}))(\cdot)^T\right]}_{\mathrm{E}[(\alpha-\bar{\alpha})(\alpha-\bar{\alpha})^T]} + \underbrace{[\bar{\mathbf{x}} - \mathbf{A}\bar{\mathbf{z}} - \mathbf{b}][\cdot]^T}_{\bar{\alpha}\bar{\alpha}^T}.\end{aligned} \tag{46}$$

After applying the expectation operator the MSE is

$$\begin{aligned}\mathrm{E}\left[(\mathbf{x}_k - \bar{\mathbf{x}}_{k|k})(\cdot)^T\right] &= \mathbf{P}^{\mathbf{xx}} \underbrace{-\mathbf{P}^{xz}\mathbf{A}^T}_{\mathcal{A}} \underbrace{-\mathbf{A}\mathbf{P}^{zx}}_{\mathcal{B}} \\ &+ \underbrace{\mathbf{A}\mathbf{P}^{zz}\mathbf{A}^T}_{\mathcal{C}} + \underbrace{[\bar{\mathbf{x}} - \mathbf{A}\bar{\mathbf{z}} - \mathbf{b}][\cdot]^T}_{\mathcal{D}},\end{aligned} \tag{47}$$

where the estimator-related expressions were labeled as $\mathcal{A}$,$\mathcal{B}$,$\mathcal{C}$, and $\mathcal{D}$. In the following paragraphs these expressions will be analyzed separately by substituting (44) and (45) into the terms:

$$\begin{aligned}\mathcal{A} &= -\mathbf{P}^{\mathbf{xz}}\mathbf{A}^T = -\mathbf{P}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}} \\ &= -(\tilde{\mathbf{P}}^{\mathbf{xz}} + \hat{\mathbf{P}}^{\mathbf{xz}})(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}} \\ &= -\tilde{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}(\mathbf{P}^{\mathbf{zx}} - \tilde{\mathbf{P}}^{\mathbf{zx}}) - \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}} \\ &= \tilde{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{P}}^{\mathbf{zx}} - \tilde{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\mathbf{P}^{\mathbf{zx}} - \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}},\end{aligned} \tag{48}$$

$$\mathcal{B} = \mathcal{A}^T \tag{49}$$

$$\begin{aligned}\mathcal{C} &= \mathbf{A}\mathbf{P}^{zz}\mathbf{A}^T = \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}(\tilde{\mathbf{P}}^{\mathbf{zz}} + \hat{\mathbf{P}}^{\mathbf{zz}})(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}} \\ &= \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{P}}^{\mathbf{zz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}} + \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}},\end{aligned} \tag{50}$$

$$\begin{aligned}\mathcal{D} &= (\bar{\mathbf{x}} - \mathbf{A}\bar{\mathbf{z}} - \mathbf{b})(\cdot)^T = (\mathbf{A}\hat{\bar{\mathbf{z}}} - \mathbf{A}\bar{\mathbf{z}})(\mathbf{A}\hat{\bar{\mathbf{z}}} - \mathbf{A}\bar{\mathbf{z}})^T \\ &= \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}}.\end{aligned} \tag{51}$$

Hence, (47) can be rewritten as

$$\begin{aligned}\mathrm{E}\left[(\mathbf{x}_k - \bar{\mathbf{x}}_{k|k})(\cdot)^T\right] &= \mathbf{P}^{\mathbf{xx}} + \mathcal{A} + \mathcal{B} + \mathcal{C} + \mathcal{D} \\ &= \underbrace{\mathbf{P}^{\mathbf{xx}} - \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}}}_{\hat{\mathbf{P}}^{\mathbf{xx}}_{k|k}} + 2\underbrace{\tilde{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{P}}^{\mathbf{zx}}}_{\mathcal{E}} \\ &\quad \underbrace{-\tilde{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\mathbf{P}^{\mathbf{zx}} - \mathbf{P}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{P}}^{\mathbf{zx}}}_{\mathcal{F}} \\ &\quad + \underbrace{\hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}}}_{\mathcal{G}} \\ &\quad + \underbrace{\hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{P}}^{\mathbf{zz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}}}_{\mathcal{H}} \\ &= \hat{\mathbf{P}}^{\mathbf{xx}}_{k|k} + \underbrace{2\mathcal{E} + \mathcal{F} + \mathcal{G} + \mathcal{H}}_{\mathrm{MSE}^{\mathrm{err}}_{\mathrm{k|k}}},\end{aligned} \tag{52}$$

which can be understood as the sum of the SIF estimate $\hat{\mathbf{P}}^{\mathbf{xx}}_{k|k}$ of $\mathbf{P}^{\mathbf{xx}}_{k|k}$ and a sum of numerical integration caused error terms $2\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$ (denoted as the $\mathrm{MSE}^{\mathrm{err}}_{\mathrm{k|k}}$), which provides the error in approximation of the MSE. By using the estimate $\mathbf{V}_N$ of the integral mean squared error related to calculation of $\hat{\bar{\mathbf{z}}}$, $\tilde{\mathbf{P}}^{\mathbf{xz}}$, and $\hat{\mathbf{P}}^{\mathbf{zz}}$, which will be denoted as $\mathbf{V}^{\hat{\bar{\mathbf{z}}}}_N$, $\mathbf{V}^{\hat{\mathbf{P}}^{\mathbf{xz}}}_N$ and $\mathbf{V}^{\hat{\mathbf{P}}^{\mathbf{zz}}}_N$, respectively, the terms $\mathcal{E}$, $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$ of $\mathrm{MSE}^{\mathrm{err}}_{\mathrm{k|k}}$ can be estimated as:

The term $\mathcal{E}$ can be calculated during computation of $\mathbf{V}^{\mathbf{P}^{\mathbf{xz}}}_N$ using (36) as

$$\begin{aligned}&\tilde{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{P}}^{\mathbf{zx}} \approx \\ &\approx \frac{1}{N}\frac{1}{N-1}\sum_{i=1}^{N}\left(\mathrm{SR}(\mathbf{Q}_i, \rho_i) - \hat{\mathbf{P}}^{\mathbf{xz}}\right)(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1} \\ &\quad \left(\mathrm{SR}(\mathbf{Q}_i, \rho_i) - \hat{\mathbf{P}}^{\mathbf{xz}}\right)^T,\end{aligned} \tag{53}$$

where the SR approximate $\mathbf{P}^{\mathbf{xz}}$ at i-th iteration.

The term $\mathcal{F}$ depends on the true CM $\mathbf{P}^{\mathbf{xz}}$. Thus the expectation of $\mathcal{F}$ w.r.t. the Q and ρ used for $\hat{\mathbf{P}}^{\mathbf{xz}}$ computation is zero, i.e.,

$$\mathrm{E}[\mathcal{F}] = \mathrm{E}\left[-\tilde{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\mathbf{P}^{\mathbf{zx}} - \mathbf{P}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbf{P}}^{\mathbf{zx}}\right] = 0. \tag{54}$$

The term $\mathcal{G}$ can be approximated after applying the expectation operator due to mutual independence of $\hat{\mathbf{P}}^{\mathbf{xz}}$, $\hat{\mathbf{P}}^{\mathbf{zz}}$ and $\tilde{\mathbf{z}}$. Then, $\mathrm{E}[\tilde{\mathbf{z}}\tilde{\mathbf{z}}^T]$ is the variance of $\bar{\mathbf{z}}$, which can be approximated by $\mathbf{V}^{\hat{\bar{\mathbf{z}}}}_N$, i.e.,

$$\mathcal{G} \approx \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\mathbf{V}^{\hat{\bar{\mathbf{z}}}}_N(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}}. \tag{55}$$

The term $\mathcal{H}$ cannot be replaced by its estimate as $\hat{\mathbf{P}}^{\mathbf{zz}}$ and $\tilde{\mathbf{P}}^{\mathbf{zz}}$ are clearly dependent. Here, the approach replacing $\tilde{\mathbf{P}}^{\mathbf{zz}}$ by its interval estimate will

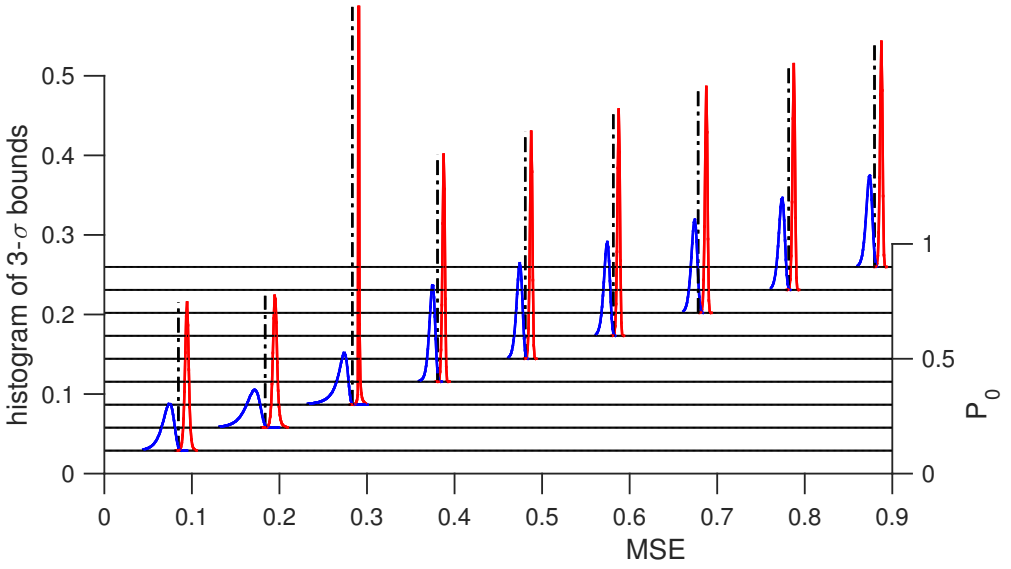

Figure 2: MSE interval estimates (blue solid line - low limit histogram, red solid line - high limit histogram) and true MSE (black vertical dashed line) for several values of covariance P_0.

be adopted. As $\tilde{\mathbf{P}}^{\mathbf{zz}}$ has multivariate Gaussian distribution with zero mean and matrix of variances given by $\mathbf{V}_N^{\hat{\mathbf{P}}^{\mathbf{zz}}}$, $\tilde{\mathbf{P}}^{\mathbf{zz}}$ in $\mathcal{H}$ can be replaced by a matrix where its each element is a 3-σ interval estimate of the corresponding element of $\tilde{\mathbf{P}}^{\mathbf{zz}}$. This matrix of 3-σ interval estimates will be denoted as $\tilde{\mathbb{P}}^{\mathbf{zz}}$. Then the term $\mathcal{H}$ can be approximated as

$$\mathcal{H} \approx \hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbb{P}}^{\mathbf{zz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}}. \tag{56}$$

Hence, the true MSE can be estimated by a 3-σ interval estimate taking into account all errors of numerical integration being neglected so far. This interval is given by substituting estimates (53) - (56) into (52), i.e.,

$$\begin{aligned}
&\mathrm{E}\left[(\mathbf{x}_k - \bar{\mathbf{x}}_{k|k})(\cdot)^T\right] = \hat{\mathbf{P}}^{\mathbf{xx}}_{k|k} + \mathrm{MSE}^{\mathrm{err}}_{\mathrm{k|k}} \\
&\approx \hat{\mathbf{P}}^{\mathbf{xx}}_{k|k} + \frac{2}{N(N-1)}\sum_{i=1}^{N}\Big(\mathrm{SR}(\mathbf{Q}_i,\rho_i) - \hat{\mathbf{I}}_N(\gamma)\Big) \\
&\quad\times(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\Big(\mathrm{SR}(\mathbf{Q}_i,\rho_i) - \hat{\mathbf{I}}_N(\gamma)\Big)^T \\
&\quad+\hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\mathbf{V}_N^{\hat{\hat{\mathbf{z}}}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}} \\
&\quad+\hat{\mathbf{P}}^{\mathbf{xz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\tilde{\mathbb{P}}^{\mathbf{zz}}(\hat{\mathbf{P}}^{\mathbf{zz}})^{-1}\hat{\mathbf{P}}^{\mathbf{zx}}.
\end{aligned} \tag{57}$$

The relation (57) represents the improved estimate of the MSE provided by the SIF. This improvement is facilitated by the SIR, which in addition to the approximate integration also calculates the MSE of the calculation error $\mathbf{V}_N$. This quantity is not provided by any other GF.

5.2 Example

In the following example $\mathrm{MSE}^{\mathrm{err}}_{\mathrm{k|k}}$ from the previous section is illustrated. Again the same static scalar example as in Section 4.2 has been chosen, i.e.,

$$z_0 = x_0^2 + v_0, \tag{58}$$

where $v_0 \sim \mathcal{N}(0,R)$ is the measurement noise with zero mean Gaussian distribution and its variance is $R = 10^{-4}$, the random variable x_0 is Gaussian with $x_0 \sim \mathcal{N}(x_0; 0.1, P_0)$, where P_0 had been tested for values $0.1, 0.2, \ldots, 0.9$. The test scenario was performed in the scope of $N_{\mathrm{MC}} = 10^6$ independent Monte Carlo simulations of state x_0 estimation. The filtering estimate $\bar{x}_{0|0}$ and most importantly the filtering CM $P^{xx}_{0|0}$ were obtained using the SIF with parameters $N_{max} = 100$ and $\varepsilon = 10^{-4}$. The relatively high simulation count secured high precision computation of the true MSE value denoted as

$$\mathrm{MSE}^{\mathrm{true}} = \frac{1}{N_{\mathrm{MC}}}\sum_{i=1}^{N_{\mathrm{MC}}}(\bar{x}^i_{0|0} - x^i_0)^2, \tag{59}$$

where superscript i denotes i-th Monte Carlo simulation, x^i_0 is the true state at the i-th simulation and $\bar{x}^i_{0|0}$ is its filtering estimate. The $\mathrm{MSE}^{\mathrm{true}}$ was compared to the MSE estimate (57) calculated by the SIF.

The results are depicted in Figure 2, where histograms of low (blue) and high (red) limits of the 3-σ intervals are plotted. The true values of the MSE are depicted as dashed vertical lines. The results confirm for several values of P_0, that the SIF is capable to provide quality interval estimate of the MSE. The SIF is

thus very versatile estimator capable of high quality self-assessment

6 CONCLUSIONS

The paper dealt with the Bayesian state estimation of nonlinear stochastic dynamic systems and specifically with the stochastic integration filter, which is based on a stochastic integration rule. Within the iterative algorithm of the SIF an instrument for measuring the nonlinearity was discovered. The instrument uses a quantity, which is a byproduct of the SIR, which has been used in the stopping condition so far. Such information provides an information of how much the nonlinearity in the system violates the Gaussian assumption of the SIF. The paper also provided a relation that can be used to predict the number of SIR iterations required to achieve the accuracy requested by the user.

Further, in the paper a method to calculate improved estimates of the MSE was developed. It was shown that the MSE contains additional terms besides the conditional variance matrix provided by the standard GFs. By utilizing the SIR byproducts the additional terms can be estimated by interval estimate. By including the additional term estimates the SIF is able to provide more accurate information about its state estimates. Both the new nonlinearity measure and the improved MSE were illustrated using simple numerical examples.

The interesting part of the discoveries made in this paper is that it opens wide area for a future work. The SIF can be further improved in terms of saving computational demands, when the measurement function is almost linear, or when the function is strongly nonlinear an execution of the algorithm can be paused and for example Gaussian sum approach can be adopted to improve estimate quality.

ACKNOWLEDGEMENTS

This work was supported by the Czech Science Foundation, project no. GA 16-19999J.

REFERENCES

Arasaratnam, I. and Haykin, S. (2009). Cubature Kalman filters. *IEEE Transactions on Automatic Control*, 54(6):1254–1269.

Arasaratnam, I., Haykin, S., and Elliott, R. J. (2007). Discrete-time nonlinear filtering algorithms using Gauss–Hermite quadrature. *Proceedings of the IEEE*, 95(5):953–977.

Bates, D. M. and Watts, D. G. (1988). *Nonlinear Regression Analysis and Its Applications*. John Wiley & Sons.

Doucet, A., De Freitas, N., and Gordon, N. (2001). *Sequential Monte Carlo Methods in Practice*, chapter An Introduction to Sequential Monte Carlo Methods. Springer. (Ed. Doucet A., de Freitas N., and Gordon N.).

Duník, J., Straka, O., and Šimandl, M. (2013a). Nonlinearity and non-Gaussianity measures for stochastic dynamic systems. In *Proceedings of the 16th International Conference on Information Fusion*, Istanbul.

Duník, J., Straka, O., and Šimandl, M. (2013b). Stochastic integration filter. *IEEE Transactions on Automatic Control*, 58(6):1561–1566.

Genz, A. and Monahan, J. (1998). Stochastic integration rules for infinite regions. *SIAM Journal on Scientific Computing*, 19(2):426–439.

Ito, K. and Xiong, K. (2000). Gaussian filters for nonlinear filtering problems. *IEEE Transactions on Automatic Control*, 45(5):910–927.

Julier, S. J. and Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *IEEE Review*, 92(3):401–421.

Kramer, S. C. and Sorenson, H. W. (1988). Recursive Bayesian estimation using piece-wise constant approximations. *Automatica*, 24(6):789–801.

Li, X. R. (2012). Measure of nonlinearity for stochastic systems. In *Proceedings of the 15th International Conference on Information Fusion*, Singapore.

Mallick, M. (2004). Differential geometry measures of nonlinearity with applications to ground target tracking. In *Proceedings of the 7th International Conference on Information Fusion*, Stockholm, Sweden.

Nørgaard, M., Poulsen, N. K., and Ravn, O. (2000). New developments in state estimation for nonlinear systems. *Automatica*, 36(11):1627–1638.

Ristic, B., Arulampalam, S., and Gordon, N. (2004). *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House.

Sorenson, H. W. (1974). On the development of practical nonlinear filters. *Information Sciences*, 7:230–270.