e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

OPTIMAL INPUT DESIGN FOR NONLINEAR DYNAMICAL SYSTEMS: A GRAPH-THEORY APPROACH

SUMALATHA
Research Scholar
M.Phil Mathematics
Bharath Institute Of Higher Education And Research
Mail Id : kumarasrinila@gmail.com
Guide Name: **Dr. N. RAMYA**
Head of the Department, Department Of Mathematics
Bharath Institute Of Higher Education And Research

Address for Correspondence

SUMALATHA
Research Scholar
M.Phil Mathematics
Bharath Institute Of Higher Education And Research
Mail Id : kumarasrinila@gmail.com
Guide Name: **Dr. N. RAMYA**
Head of the Department, Department Of Mathematics
Bharath Institute Of Higher Education And Research

Abstract:

In many applications, modelling is a crucial step. The process of developing a mathematical description for a particular natural event is referred to as modelling (also called system). Price prediction in banking [85], channel estimation in communication systems [88], and controller design in industrial processes [80] are some examples.Modeling may be divided into two categories. The first method considers modelling a natural event using physical rules (for example, modelling an electrical circuit using Kirchhoff's current and voltage laws). The second method is to describe the natural phenomena as a black box model (or a grey box model if some physical understanding is taken into consideration) and identify the model using input-output data from a system experiment (e.g., provide a voltage excitation to an electrical circuit, and measure the current to obtain a model for the equivalent impedance). The second method, often referred to as system identification, is the focus of this thesis.The input sequence we give to excite the system is

one of the most important aspects of system identification. Because system identification requires input-output data, it's important to create an input excitation that, in certain ways, optimises the information in the experiment (e.g., optimising a cost function related to the intended model application). Input design refers to the process of creating an input sequence for system identification, which is the subject of this thesis.The words input design and experiment design are differentiated. The specification of input and output signals, measurement instants, modified signals, and how to modify them are all options in experiment design (which is the focus of input design). Signal conditioning is also included (e.g., the choice of presampling filters [59, Paper 13]).

INTRODUCTION

The link between stationary processes with finite alphabet and memory and de Bruijn graphs was explored in Paper 2. The findings will now be used to create techniques for creating nonlinear model input sequences based on the findings.

The primary problem in constructing input sequences to detect nonlinear models, as stated in Paper 1, is that frequency domain methods cannot be used. As a consequence, most input design findings for linear models cannot be applied to nonlinear models. Furthermore, most current input design findings are incapable of dealing with amplitude limitations, which may occur owing to physical or safety concerns.

This paper presents a method for identifying nonlinear output-error (NOE) models by constructing input sequences. The approach treats the creation of an input sequence as the realisation of a stationary process with limited memory and alphabet. Maximizing a scalar cost function of the Fisher information matrix yields the best stationary process. We use de Bruijn graphs to derive the probability measures associated with the extreme points after parameterizing the set of stationary processes in terms of its extreme points, and then estimate the related Fisher information matrix for each extreme point. As a result, the issue becomes convex for nonlinear models as well. Numerical examples demonstrate that the approach described in this paper is compatible with previous findings in the literature and that it is a viable option for constructing nonlinear model input sequences.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

The results in [54] and [6, 7] may be viewed as an extension of the findings in this paper. The major difference between [6, 7] and [6, 7] is that instead of simply optimising over the transition probabilities, we optimise over the stationary pmf associated with the Markov chain. This method produces a convex problem (which cannot be solved in [6, 7]), which necessitates the use of optimization methods that guarantee local optima.

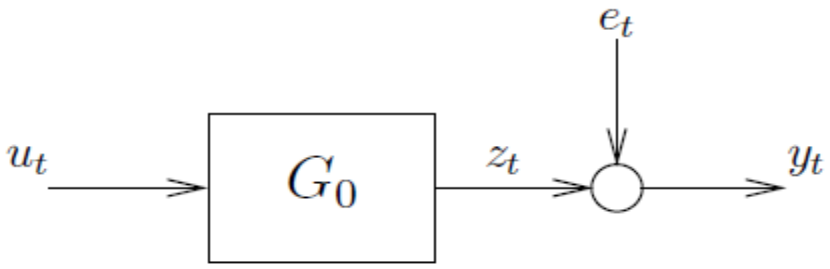

Figure 3.1: Block diagram of a (possibly nonlinear) system.

to find work). [54] discusses a similar method to the one given in this paper, although it is limited to nonlinear FIR systems. The input design issue in [54] is addressed in terms of a limited-length input realisation by using the finite memory feature of nonlinear FIR models. However, since the models will usually rely on the whole previous input sequence, the findings in [54] cannot be used to construct input sequences for identification of more generic nonlinear output-error models. In this vein, the current paper broadens the study to encompass nonlinear FIR systems as well as more generic nonlinear model structures.

3.1 Formulation of the Problem

Consider the time invariant single input, single output (SISO) system shown in Figure 3.1. G0 is a dynamical system (potentially nonlinear) in this case.

), defined for $t \geq 1$

$$G_0(u_t) := \begin{cases} x_{t+1} = f_0(x_t, u_t), \\ z_t = h_0(x_t, u_t), \\ x_1 = \mu, \end{cases} \tag{3.1}$$

where $\{e_t\}$ is white noise sequence with zero mean and finite variance λ_e, $u_t \in \mathrm{R}$ is the input, $x_t \in$

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

R^{nx} are the internal states of G_0 with initial condition $\mu \in R^{nx}$, and $y_t \in R$ is the measured output. We assume that we have a model structure G, defined for any $\theta \in \Theta \subset R^{n\theta}$ as

$$G(u_t;\,\theta) := \begin{cases} x_{t+1} = f(x_t,\,u_t;\,\theta)\,, \\ z_t = h(x_t,\,u_t;\,\theta)\,, \\ x_1 = \mu\,. \end{cases} \tag{3.2}$$

We assume that there exists a $\theta_0 \in \Theta$ such that $G(u_t; \theta_0) = G_0(u_t)$ [59], i.e., there is no undermodelling. Notice that the noise, e_t, is assumed to enter only at the output. To continue, we introduce the following definition:

Definition 3.1 *Consider a bounded signal* $\{u_t\}$, $|u_t| \leq K$ ($K > 0$), *and a non linear system* $y_t = G_0(u_t)$. *We say that* G_0 *is* exponentially stable *if and only if there are constants* $C > 0$ *(depending possibly on K),* $0 < \delta < 1$, *such that for all* $t, s \in Z$,

$$|G_0(u_t) - G_0(u^s_t)| < C\delta^{t-s}, \tag{3.3}$$

where

$$u_k^s = \begin{cases} u_k, & k > s, \\ 0, & otherwise. \end{cases} \tag{3.4}$$

We notice that Definition 3.1 differs from that given by [58] since it considers deterministic systems, and it is not defined in terms of moments of order 4.

The objective in this paper is to design an input signal $u_{1:nseq} = (u_1, \ldots, u_{nseq})$ as a realization of a stationary process, such that the system (3.1) can be estimated with maximum accuracy as defined by a scalar function of the Fisher information matrix I^e_F[59]. I^e_Fcan be computed using Lemma 1.1 as[1]

$$\mathcal{I}_F^e = \frac{1}{\lambda_e}\mathbf{E}_e\left\{\sum_{t=1}^{n_{seq}} \psi_t^{\theta_0}(u_t)\psi_t^{\theta_0}(u_t)^\top \,\middle|\, u_{1:n_{seq}}\right\}, \tag{3.5}$$

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

where

$$\psi_t^{\theta_0}(u_t) := \left.\frac{d\,\hat{y}_t(u_t,\,\theta)}{d\theta}\right|_{\theta=\theta_0}, \tag{3.6a}$$

$$\hat{y}_t(u_t,\,\theta) := G(u_t;\,\theta)\,, \tag{3.6b}$$

and $\theta,\ \theta_0 \in \Theta$. Since we are interested in computing $u_{1:n\text{seq}}$ as a realization from a stationary process, we need to quantify the Fisher information matrix in terms of its expected value with respect to $u_{1:n\text{seq}}$. We define the result of this expected value as the *per-sample* Fisher information matrix, which is computed as

$$\begin{aligned}\mathcal{I}_F &:= \mathbf{E}_{u_{1:n_{\text{seq}}}}\left\{\mathcal{I}_F^e\right\} \\ &= \frac{1}{\lambda_e}\mathbf{E}_{u_{1:n_{\text{seq}}},\,e}\left\{\sum_{t=1}^{n_{\text{seq}}}\psi_t^{\theta_0}(u_t)\psi_t^{\theta_0}(u_t)^\top\right\}\,.\end{aligned} \tag{3.7}$$

Equation (3.6b) does not depend on the noise realization. Therefore, we can rewrite (3.7)as

$$\mathcal{I}_F = \frac{1}{\lambda_e}\int_{u_{1:n_{\text{seq}}}\in\mathbb{R}^{n_{\text{seq}}}}\sum_{t=1}^{n_{\text{seq}}}\psi_t^{\theta_0}(u_t)\psi_t^{\theta_0}(u_t)^\top\,dP(u_{1:n_{\text{seq}}})\,, \tag{3.8}$$

We note that (3.8) depends on $P(u_{1:n\text{seq}})$. Therefore, the input design problem we will consider is to find a cdf, $P^{\text{opt}}(u_{n\text{seq}})$, which maximizes a scalar function of (3.8). We define this function as h: $\mathrm{R}^{n_\theta \times n^\theta} \rightarrow \mathrm{R}$. As it is customary in input design [33, 45, 59], h is assumed to be a concave, nondecreasing, matrix function [5, pp. 108]. The assumption of h being nondecreasing function is to guarantee that, for any two matrices X and Y in the positive semidefinite cone, we have $h(X) \geq h(Y)$ when $X \quad Y$. Several choices of h have been proposed in the literature [74]. In this paper, we leave the selection of h open to the user.

Since $P(u_{1:n\text{seq}})$ has to be a stationary cdf, the optimization must be constrained to the set

$$\mathcal{P}_{n_{\text{seq}}} := \Big\{ P : \mathbb{R}^{n_{\text{seq}}} \to \mathbb{R} \,\Big|\, P(\mathrm{x}) \geq 0,\ \forall \mathrm{x} \in \mathbb{R}^{n_{\text{seq}}};$$
$$P \text{ is monotone nondecreasing};$$
$$\lim_{x_i \to \infty} P(x_1, \ldots, x_{n_{\text{seq}}}) = 1, \quad \text{for } i = 1, \ldots, n_{\text{seq}}$$
$$\int_{v \in \mathbb{R}} dP([v, \mathrm{z}]) = \int_{v \in \mathbb{R}} dP([\mathrm{z}, v]),\ \forall \mathrm{z} \in \mathbb{R}^{n_{\text{seq}}-1} \Big\} \tag{3.9}$$

As mentioned in Paper 2, the last condition in (3.9) guarantees that $P \in \mathrm{P}_{n\text{seq}}$ is the projection of the cdf of a stationary sequence [100].

Based on Definition 3.1, we impose the following assumption over $\psi^{\theta}0_{\,t}(u_t)\psi^{\theta}0_t(u_t)^{\top}$.

Input design via graph theory

We must parameterize the set PC in a computationally tractable way to solve Problem 3.2. To this purpose, we'll characterise each element in PC as a convex combination of its extreme points, as described in Section 2.2. In this part, we use the notation established in Section 2.2 to represent the set of PC's extreme points, which we refer to as VPC.

We know how to calculate the probability measures associated with the elements thanks to Theorem 2.1 and Lemma 2.2. In the VPC. Indeed, Theorem 2.1 says that the prime cycles in the de Bruijn graph GCnm correspond one-to-one to the elements in VP (the set of Cnm's extreme points), with a uniform distribution whose support equals the prime cycle's elements. Furthermore, Lemma 2.2 says that the de Bruijn graph GCnm's prime cycles are generated from the de Bruijn graph GCnm1's basic cycles. As a result, the Fisher information matrix corresponding to the i-th prime cycle and element wi VPC may be defined as

$$\mathcal{I}_F^{(i)} := \frac{1}{\lambda_e} \sum_{u_{1:n_m} \in \mathcal{C}^{n_m}} \sum_{t=1}^{n_m} \psi_t^{\theta_0}(u_t^i)\psi_t^{\theta_0}(u_t^i)^{\top}\, w_i(u_{1:n_m}) \tag{3.16}$$

for all $i \in \{1, \ldots, n_V\}$. Notice that each $I^{(i)}_F$is associated with the i-th prime cycle. since it requires the computation of expected values of nonlinear functions of u^i_t. Therefore, a numerical approximation of (3.16) is needed. To this end, instead of approximating $I_F^{(i)}$ as an average over different realizations of the input sequence $u_{1:n\text{seq}}$, we consider an approximation of $I_F^{(i)}$an average over different realization of the input sequence

$$\mathcal{I}_F^{(i)} \approx \frac{1}{\lambda_e N_{\text{sim}}} \sum_{t=1}^{N_{\text{sim}}} \psi_t^{\theta_0}(u_t^i)\psi_t^{\theta_0}(u_t^i)^\top, \tag{3.17}$$

where N_{sim} is sufficiently large. The approximation (3.17) converges to $I_F^{(i)}$ as $N_{\text{sim}} \to \infty$ since $\psi_t{}^{\theta}0\,(u_t)^{\mathrm{T}}$ satisfies Assumption 3.1 (see Appendix B for a proof of this statement). The approximation error incurred when (3.17) is employed to compute $I_F^{(i)}$ is of order δ_{φ}^{Nsim}

Reducible Markov chains

When the optimization (3.19) is solved, it might occur that the resulting p^{opt} is associated to a reducible Markov chain. The last means that there exists at least two sets of vertices in the Markov chain, such that each set of vertices cannot be accessed from the others. This is an issue of the proposed approach, since Algorithm 2.1 cannot be employed to generate samples from p^{opt} if it is the pmf of a reducible Markov chain.

One possibility to overcome this issue is to perturb the optimal pmf p^{opt} in order to achieve an irreducible Markov chain, and then use Algorithm 2.1 to generate samples from the perturbed pmf. In this case, the samples will be distributed according to a suboptimal pmf. The problem of optimal pmfs p^{opt} resulting in reducible Markov chains is a topic for future research.

Numerical examples

In this section, we will introduce numerical examples to illustrate the proposed input design method.

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

Example 3.1 *(Input design for nonlinear FIR models) In this example we will solve the input design problem for the system in Figure 3.1 in page 38, with*

$$G_0(u_t) = G_1(q, \theta)\, u_t + G_2(q, \theta)\, u^2_t \qquad (3.22)$$

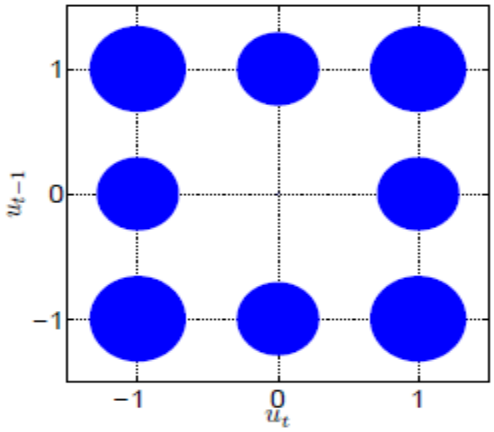

Figure 3.2: Plot with the stationary probabilities for the optimal input signal of Example 3.1. The radius of each disc is proportional to the probability of the state (u_t, u_{t-1}).

where

$$G_1(q, \theta) = \theta_1 + \theta_2\, q^{-1}, \qquad (3.23)$$

$$G_2(q, \theta) = \theta_3 + \theta_4\, q^{-1}. \qquad (3.24)$$

We assume that $\{e_t\}$ *is Gaussian white noise with variance* $\lambda_e = 1$. *This system has been introduced as an example in [54].*

We will solve Problem 3.2 by considering $h(\cdot) = \log\det(\cdot)$, *and a ternary se quence (*$n_C = 3$*) of length* $n_m = 2$. *For this example, we take* $\mathrm{C} = \{-1, 0, 1\}$.

To solve (3.19)-(3.20) *we consider* $N_{\text{sim}} = 5 \cdot 10^3$ *in* (3.17). *The implementation of* (3.19)-(3.20) *was made in* Matlab *using the* cvx *toolbox [35].*

e-ISSN 2320 –7876 www.ijfans.org
Research Paper
Vol.11, Iss.9, Dec 2022

The simulation results give an optimal cost[2] $\log\det(I^{app}{}_F) = -1.717$. *Figure 3.2 shows the optimal stationary probabilities for each state* (u_t, u_{t-1}) *(cf. Figure 4(a) in [54]). The results presented here show that the proposed method is consistent with the results in [54].*

Example 3.1 shows that this method is equivalent to the method introduced in [54] when G_0 has a nonlinear FIR-type structure.

The results in this paper can also be employed for linear systems when ampli tude constraints are considered in the input sequence by forcing u_t to belong to a finite alphabet. The next example shows an application in that direction.

Example 3.2 *(Input design with amplitude constraints) In this example we con sider the mass-spring-damper system introduced in [7]. The input, u, is the force applied to the mass and the output, y, is the mass position. The continuous-time system is described by the transfer function*

$$G_0(s) = \frac{\frac{1}{m}}{s^2 + \frac{c}{m}s + \frac{k}{m}}, \tag{3.25}$$

with $m = 100$ *[Kg],* $k = 10$ *[N/m], and* $c = 6.3246$ *[Ns/m]. This choice results in the natural frequency* $\omega_n = 0.3162$ *[rad/s] and damping* $\xi = 0.1$. *The noise* $\{e_t\}$ *is white with zero mean and variance* $\lambda_e = 10^{-4}$. *The system* (3.25) *is sampled by using a zero-order-hold with sampling period* $T_s = 1$ *[s]. This gives the discrete-time system*

$$G_0(u_t) = \frac{4.86 \cdot 10^{-3} q^{-1} + 4.75 \cdot 10^{-3} q^{-2}}{1 - 1.84\, q^{-1} + 0.94 q^{-2}}\, u_t\,. \tag{3.26}$$

As a model, we define

$$G(u_t;\, \theta) = \frac{\theta_1\, q^{-1} + \theta_2 q^{-2}}{1 + \theta_3\, q^{-1} + \theta_4 q^{-2}}\, u_t\,, \tag{3.27}$$

where

$$\theta = [\theta_1\ \theta_2\ \theta_3\ \theta_4]^{\top}. \qquad (3.28)$$

We will solve Problem 3.2 for two cost functions: $h(\cdot) = -\operatorname{tr}\{(\cdot)^{-1}\}$ *and* $h(\cdot) = \log\det(\cdot)$, *subject to a binary sequence (*$n_C = 2$*) of length* $n_m = 2$. *In this example, we define* $C = \{-1, 1\}$, *and* $N_{sim} = 5 \cdot 10^3$.

The solution of Problem 3.2 for this example gives $\operatorname{tr}\{(I_F^{app})^{-1}\} = 0.1108$ and $\log\det(I_F^{app}) = 28.22$. *Figure 3.3 and 3.4 present the stationary probabilities of the optimal input signal for both cost functions. We can see that the stationary probabilities depend on the cost function h. However, both cost functions assign higher stationary probabilities to the states* $(-1, -1)$ *and* $(1, 1)$.

We can compare the performance of our approach with the method introduced in [7]. For this purpose, we generate an input sequence of length N_{sim} *by running the Markov chain associated to the stationary distribution in Figure 3.3, and the 4-states Markov chain presented in [7]. To guarantee that the input is a realization of a stationary process, we discard the first* 10^6 *realizations of the Markov chain. The results for the sampled information matrix are* $\operatorname{tr}\{I_F^{-1}\} = 1.8233.10^{-4}$ *for thefour states markov chain presented in [7], and* $tr\{L_F^{-1}\} = 1.6525 \cdot 10^{-4}$ *for our method (we note that our results are consistent with those reported in [7], since the scaling factor* N_{sim} *is not considered here). Therefore, based on the variance of the parameter estimates, we conclude that the approach in this paper gives better results for the example introduced in [7].*

To have an idea of the computation time required for this example, the opti mization was solved in a laptop Dell Latitude E6430, equipped with Intel Core i7

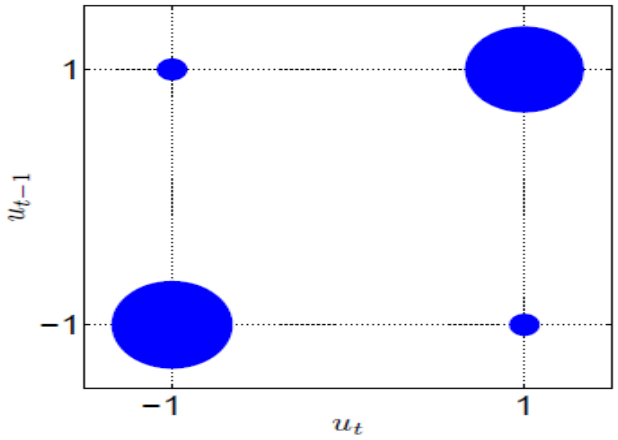

Figure 3.3: Plot with the stationary probabilities for the optimal input signal in Example 3.2 as in Figure 3.2. $h(\cdot) = -\operatorname{tr}\{(\cdot)^{-1}\}$.

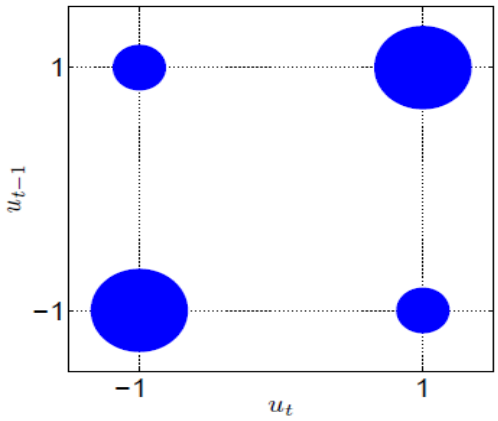

Figure 3.4: Plot with the stationary probabilities for the optimal input signal in Example 3.2 as in Figure 3.2. $h(\cdot) = \log\det(\cdot)$.

2.6 *[GHz] processor, and* 8 *[Gb] of RAM memory. The time required for the com putation of elementary cycles to the computation of stationary probabilities is* 1.9 *seconds.*

The results presented in the previous examples show that the method intro duced in this paper retrieves (or improves) the results in the literature. The next examples show an application of the input design method to a model structure not covered by existing techniques.

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.9, Dec 2022

Research Paper

OPTIMAL INPUT DESIGN FOR NONLINEAR STATE SPACE MODELS

A novel technique for input design for nonlinear model identification was presented in Paper 3. The topic in the paper was limited to nonlinear output error models. The model structure assumption was included to make the Fisher information matrix approximation, which is based on numerical approximations, easier. The approach in Paper 3 does not, however, address the situation when the noise process also influences the model's internal states. The approximation of the Fisher information matrix provided in Paper 3 is no longer applicable in these models. To apply the findings of Paper 3 to more broad nonlinear model structures, additional approximation techniques must be explored.

We extend the input design technique presented in Paper 3 to nonlinear state space models in this paper. We utilise a graph theoretical method to construct an input sequence as a realisation of a stationary process that maximises the Fisher information matrix's scalar cost function, as we did in Paper 3. We utilise particle techniques to get the necessary estimates as the sample covariance matrix of the scoring function to estimate the Fisher information matrices for the extreme points in the set of stationary processes. The optimization is then performed for several realisations of the score function's sample covariance matrix, and the optimum pdf is derived as the sample mean of the solutions found for the various data realisations. The technique is shown to be an appealing strategy for designing input sequences for the identification of nonlinear state space models using numerical examples.

Formulation of the Problem

This section 1 introduces a nonlinear state space model extension of the optimum input design formulation given in Paper 3. The discussion in this section is similar to that in Section 3.1, with the exception that we construct the issue in this section to include nonlinear state space models.

The goal, as in Paper 3, is to create an input signal u1:nseq = (u1,..., unseq) that realises a stationary process. This is done in order to identify a state space model (SSM) as accurately as

possible, as described by a scalar function of the Fisher information matrix IF [59]. The states x0:T = (x0,..., xT), inputs u1:T, and measurements y1:T of an SSM are provided by

$$x_0 \sim \mu\,(x_0), \tag{4.1a}$$

$$x_t|x_{t-1} \sim f_\theta\,(x_t \mid x_{t-1}, u_{t-1}), \tag{4.1b}$$

$$y_t|x_t \sim g_\theta\,(y_t \mid x_t, u_t). \tag{4.1c}$$

Here, $f_\theta\,(\cdot)$ and $g_\theta\,(\cdot)$ denote known probability distributions parametrized by $\theta \in \Theta \subset \mathrm{R}^{n\theta}$. In the sequel, we make the rather restrictive albeit standard assumption that there exists $\theta_0 \in \Theta$ such that the model (4.1) describes the system to be identified when $\theta = \theta_0$, i.e., there is no undermodeling. This assumption is necessary in order to quantify the information retrieved from the experiment as a function of the Fisher information matrix.

We notice that we can write the joint distribution of states and measurements for (4.1) as

$$p_\theta(x_{1:T}, y_{1:T}|u_{1:T}) = \mu(x_0)\prod_{t=1}^{T} f_\theta(x_t|x_{t-1}, u_{t-1})g_\theta(y_t|x_t, u_t). \tag{4.2}$$

This quantity is used in the sequel for estimating I_F by

$$\mathrm{I}_{F\,=\,\mathbf{E}}\,\mathrm{S}(\theta_0)\mathrm{S}^{\mathsf{T}}(\theta_0)\,, \tag{4.3a}$$

$$\mathrm{S}(\theta_0) = \nabla_\theta \log p_\theta(y_{1:n\mathrm{seq}})_{\;\theta=\theta 0}, \tag{4.3b}$$

where $p_\theta(y_{1:n\mathrm{seq}})$ and $\mathrm{S}(\theta)$ denote the likelihood function and the score function, respectively. Note that the expected value in (4.3a) is with respect to the probability distribution $p_\theta(x_{0:T}|y_{1:T})$ and the realizations of $u_{1:n\mathrm{seq}}$.

We note that (4.3a) depends on the cumulative distribution function of $u_{1:n\mathrm{seq}}$, $P_u(u_{1:n\mathrm{seq}})$. Therefore, the input design problem is to find a cdf $P_u^{opt}(u_{1:n\mathrm{seq}})$which maximizes a scalar function

of (4.3a). We define this scalar function as $h : \mathrm{R}^{n_\theta \times n^\theta} \to \mathrm{R}$. To obtain the desired results, h must be a concave, nondecreas ing, matrix function [5, pp. 108] (cf. Section 3.1). In this paper we leave the selection of h to the user.

Algorithm 4.1 Sequential importance sampling

Algorithm 4.1 Sequential importance sampling

INPUTS: N (number of samples), $p_{\mathrm{I}}(x_{1:T})$ (importance distribution), and $\gamma_T(x_{1:T})$ (unnormalized pdf).
OUTPUT: $\{x_{1:T}^{(i)}\}_{i=1}^N$ (realizations), and $\{w^{(i)}\}_{i=1}^N$ (normalized weights).

1: Sample $x_1^{(i)}$ from $p_{\mathrm{I}}(x_1)$ for $i = 1$ to N.
2: Compute the weights $w_1(x_1^{(i)}) = p_{\mathrm{I}}(x_1)$, and

$$w_1^{(i)} = \frac{w_1(x_1^{(i)})}{\sum_{j=1}^N w_1(x_1^{(i)})}, \tag{4.28}$$

for $i = 1$ to N.
3: **for** $t = 2$ to T **do**
4: Sample $x_t^{(i)}$ from $p_{\mathrm{I}}(x_t|x_{1:t-1}^{(i)})$ for $i = 1$ to N.
5: Compute the weights

$$w_t(x_{1:t}^{(i)}) = w_{t-1}(x_{1:t-1}^{(i)}) \cdot \alpha(x_{1:t}^{(i)}), \tag{4.29}$$

where α given in (4.32), and

$$w_t^{(i)} = \frac{w_t(x_{1:t}^{(i)})}{\sum_{j=1}^N w_t(x_{1:t}^{(i)})}, \tag{4.30}$$

for $i = 1$ to N.
6: **end for**

e-ISSN 2320 –7876 www.ijfans.org
Vol.11, Iss.9, Dec 2022

Research Paper

Resampling

We have seen that IS (and therefore SIS) provides estimates whose variance in creases with *t*. Resampling techniques are a key ingredient of SMC methods which (partially) solve this problem in some important scenarios.

Resampling is an intuitive idea with major practical and theoretical benefits. We consider first an IS approximation $\hat{p}(x_{1:T})$ of the target distribution $p(x_{1:T})$. This approximation is based on weighted samples from $p_I(x_{1:T})$, and does not pro vide samples approximately distributed according to $p(x_{1:T})$. To obtain approxi mate samples from $p(x_{1:T})$, we can simply draw samples from its IS approximation $\hat{p}(x_{1:T})$, where $\hat{p}(x_{1:T})$ is defined in (4.17), with normalized weights

This operation is called *resampling* as it corresponds to sampling from an approximation $\hat{p}(x_{1:T})$ which was itself obtained by sampling. If we are interested in obtaining N samples from $\hat{p}(x_{1:T})$, then we can resample N times from $\hat{p}(x_{1:T})$: This is equivalent to associating a number of off spring $N_T^{(i)}$with each sample $x_{1:T}^{(i)}$ in such a way that $N_T^{(1:N)}:=\left(N_T^{(1)},\ldots\ldots N_T^{(N)}\right)$ follow a multinomial distribution with parameter vector (N, $w_T^{(1:N)}$and associating a weight of 1/N with each offspring. Thus, we approximate $\hat{p}(x_{1:T})$ by the

CONCLUSION

The concepts of graph theory and stationary processes used in the thesis were presented in this paper 2. We showed how to use de Bruijn graphs to represent a collection of stable processes with limited memory. It was demonstrated that the convex hull of the probability measures associated with the prime cycles of the analogous de Bruijn graph may be represented as the set of stationary processes with finite memory and finite alphabet. We're looking for realisations from a certain pmf in the set of stationary processes. We treat the samples as the output of a Markov chain with stationary distribution provided by the desired pmf to get a realisation with the desired distribution. This paper presented a method for constructing the related transition probability matrix, resulting in a Markov chain with the required pmf as stationary distribution. Nonetheless, the method does not ensure that the resultant transition probability matrix has the optimum mixing

time; however, numerical optimization techniques may be used to enhance this feature, which will be addressed in future work.

A novel input design method for identifying nonlinear outputerror models is presented in this paper 3. The approach treats the creation of an input sequence as a realisation of a stationary process achieved by maximising the information matrix's scalar cost function. We parameterize the set of stationary processes in terms of convex combinations of the set's extreme points to produce a computationally tractable issue. The probability measures associated with the extreme points are calculated as the set of prime cycles associated with the corresponding de Bruijn graph, provided the stationary process has finite memory and alphabet. As a result, each member in the set of stationary processes' information matrix may be calculated as a convex combination of the information matrices acquired for each extreme point. The information matrices for each extreme point are generated using numerical approximations due to the complexity of nonlinear model architectures. The primary benefit of this method is that, even for nonlinear model structures, the input design issue becomes convex.

We addressed an expansion of the input design technique described in Paper 3 in this paper 4. The expansion incorporates a nonlinear state space description as a more generic model structure. We construct the input sequence by maximising a scalar cost function of the Fisher information matrix over the set of stationary processes with finite memory and finite alphabet, as described in Paper 3. The calculation of the Fisher information matrix for the collection of basis inputs is the most challenging part of this paper. The Monte Carlo approximation presented in Paper 3 cannot be used here because we assume a more broad nonlinear model structure. We utilise particle techniques to estimate the Fisher information matrices for the collection of basis inputs to solve this problem. The resultant optimization is repeated many times across various realisations, and the ultimate result is the sample mean over the different realisations, since the technique is based on numerical approximations. The suggested method may be used to construct input sequences to discover nonlinear state space models, as shown by numerical examples.

The computational cost of this method is, as one would assume, linked not only to the calculation of the basic inputs, but also to the number of particles and the length of the input used

to estimate the score function. The decrease of the computational complexity associated with the estimate of the score function will be the focus of future research in this area.

We offer an application of the input design approach described in previous papers in this fifth paper. The aim is to create an input sequence that is a realisation of a stationary process, with the goal of getting the most information out of the experiment by applying the input sequence to a system with quantized output. Because the output measurements are quantized, the calculation of the Fisher information matrix becomes problematic when applying the suggested method on these systems. To address this problem, this paper presents a Fisher information matrix estimate based on Fisher's identification. Numerical approximations are used to estimate the Fisher information matrices associated with the probability measures characterising the set of stationary processes, which are then used to calculate the estimate. The numerical example shows how the suggested technique may be used to create input sequences for quantized output systems. More complicated model structures will be considered in future research, necessitating the development of new expressions to approximate the Fisher information matrix.

REFERENCES

[1] J.C. Agüero, W. Tang, J.I. Yuz, R. Delgado, and G.C. Goodwin. Dual timefrequency domain system identification. *Automatica*, 48(12):3031–3041, 2012.

[2] B.D.O. Anderson and J.B.Moore. *Optimal filtering*. Prentice Hall, Englewood Cliffs, N.J., 1979.

[3] X. Bombois, G. Scorletti, M. Gevers, P.M.J. Van den Hof, and R. Hildebrand. Least costly identification experiment for control. *Automatica*, 42(10):1651–1662, October 2006.

[4] S. Boyd, P. Diaconis, and L. Xiao. Fastest mixing Markov chain on a graph. *SIAM Review*, 46(4):667–689, October 2004.

[5] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[6] C. Brighenti. On input design for system identification: input design using Markov chains, M.Sc. Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2009.

[7] C. Brighenti, B.Wahlberg, and C.R. Rojas. Input design using Markov chains for system identification. In *Joint 48th Conference on Decision and Control and 28th Chinese Conference*, pages 1557–1562, Shangai, P.R. China, 2009.

[8] O. Cappé, E. Moulines, and T. Rydén. *Inference in Hidden Markov Models*. Springer, 2005.

[9] G. Casella and R.L. Berger. *Statistical inference*. Duxbury Pacific Grove, CA, 2002.

[10] M. Casini, A. Garulli, and A. Vicino. Input design for worst-case system identification with uniformly quantized measurements. In *15th IFAC Symposium on System Identification (SYSID)*, Saint-Malo, France, 2009.

[11] M. Casini, A. Garulli, and A. Vicino. Input design in worst-case system identification using binary sensors. *IEEE Transactions on Automatic Control*, 56(5):1186–1191, 2011.

[12] M. Casini, A. Garulli, and A. Vicino. Input design in worst-case system identification with quantized measurements. *Automatica*, 48(5):2297–3007, 2012.

[13] B. Chen and P.K. Willett. Channel optimized binary quantizers for distributed sensor networks. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pages (iii) 845–848. IEEE, 2004.

[14] D.R. Cox. *Planning of experiments*. New York: Wiley, 1958.

[15] H. Cramér. *Mathematical methods of statistics*. Princeton University Press, Princeton, 1946.

[16] J. Dahlin, F. Lindsten, and T.B. Schön. Second-order particle MCMC for Bayesian parameter inference. In *Proceedings of the 19th IFAC World Congress*, Cape Town, South Africa, August 2014.

[17] N.G. de Bruijn and P. Erdos. A combinatorial problem. *Koninklijke Nederlands Akademie v. Wetenschappen*, 49:758–764, 1946.

[18] A. De Cock, M. Gevers, and J. Schoukens. A preliminary study on optimal input design for nonlinear systems. In *Proceedings of the IEEE Conference on Decision and Control (CDC)*, Florence, Italy, 2013.

[19] P. Del Moral, A. Doucet, and A. Jasra. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B*, 68(3):411–436, 2006.

[20] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B*, 39(1).

[21] J.L. Doob. *Stochastic processes*. New York Wiley, 1953.

[22] A. Doucet and A. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. In D. Crisan and B. Rozovsky, editors, *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press, 2011.

[23] V.V. Fedorov. *Theory of optimal experiments*. New York: Academic Press, 1972.

[24] R.A. Fisher. On an absolute criterion for fitting frequency curves. *Messenger of Mathematics*, 41(8):155–160, 1912.

[25] R.A. Fisher. Theory of statistical estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 22(5):700–725, 1925.