

Article

Single Neuron Stochastic Predictive PID Control Algorithm for Nonlinear and Non-Gaussian Systems Using the Survival Information Potential Criterion

Mifeng Ren ¹, Ting Cheng ¹, Junghui Chen ^{2,*}, Xinying Xu ¹ and Lan Cheng ¹

¹ College of Information Engineering, Taiyuan University of Technology, Taiyuan 030024, China; renmifeng@126.com (M.R.); chengtingtyut@126.com (T.C.); xuxinyingtut@sina.com (X.X.); taolan_1983@126.com (L.C.)

² Department of Chemical Engineering, Chung-Yuan Christian University, Chung-Li, Taoyuan 32023, Taiwan, Republic of China

* Correspondence: jason@wavenet.cycu.edu.tw; Tel.: +886-3-265-4107

Academic Editor: Raúl Alcaraz Martínez

Received: 17 March 2016; Accepted: 26 May 2016; Published: 3 June 2016

Abstract: This paper presents a novel stochastic predictive tracking control strategy for nonlinear and non-Gaussian stochastic systems based on the single neuron controller structure in the framework of information theory. Firstly, in order to characterize the randomness of the control system, survival information potential (SIP), instead of entropy, is adopted to formulate the performance index, which is not shift-invariant, *i.e.*, its value varies with the change of the distribution location. Then, the optimal weights of the single neuron controller can be obtained by minimizing the presented SIP based predictive control criterion. Furthermore, mean-square convergence of the proposed control algorithm is also analyzed from the energy conservation perspective. Finally, a numerical example is given to show the effectiveness of the proposed method.

Keywords: nonlinear and non-Gaussian systems; single neuron controller; stochastic predictive control; survival information potential criterion; mean square convergence

1. Introduction

Since almost all the control systems are subject to random signals (such as those originating from system parameter variations and sensor noise, *etc.*), stochastic systems are widely encountered in control engineering design. Minimizing the randomness in the closed-loop system is one of the important practical issues in controller design. Therefore, minimum variance control [1] has been obtained significant attentions. Its purpose is to minimize variations in the controlled system outputs or the tracking errors. Indeed, even today, most stochastic control design methods have only focused on control of the output mean and the variance of stochastic systems. In general, these developments are done mostly based on the assumptions that the system variables are of Gaussian types. Such assumptions, although strict, allow control engineers to make use of the well-established stochastic theory to perform controller design and closed-loop system analysis.

In industrial processes, product quality data can be approximated by the Gaussian probability density function (PDF) when the system operates normally. However, when abnormality occurs along the production line, the variabilities of these quality variables would not follow Gaussian distributions. In this regard, actions need to be taken so that the manipulated variables can be tuned to bring these quality variables back to certain desired ones. In fact, most industrial processes have difficulty meeting the Gaussian assumption because of the mixture of different courses with Gaussian disturbances or other factors. Moreover, the nonlinearity in stochastic systems could lead to non-Gaussian randomness

even if the disturbances follow a Gaussian distribution. Thus, controlling the mean and variance of system variables may be far from sufficiently characterizing the statistical property of the stochastic processes. It is known that, in many cases, the behavior of a stochastic process can be completely characterized by the shape of its statistical distribution represented by PDF. Therefore, for analysis and design purposes, it is important to consider the entire PDF. PDF-shaping control design accounts for the issues mentioned above by selecting a certain shape for the process PDF as the goal of the control design procedure would provide an accurate and flexible control strategy that can accommodate a wide class of objectives.

In order to solve the problems existing in paper-making processes, the stochastic distribution control (SDC) theory was proposed by Wang (1996) [2]. This theory aims at controlling the shape of the output PDF instead of the mean and variance for stochastic systems [3–5]. After that, SDC has been used to handle the stochastic systems with non-Gaussian disturbances. Then, a linear-matrix-inequality-based convex optimization algorithm was developed for control, filter design and fault detection in non-Gaussian systems [6–8]. Nevertheless, the PDFs of the output are not necessarily measurable; instead, a more general measure of uncertainty, namely the entropy, has been used to characterize the uncertainty of the output tracking error. Compared with SDC (PDF shaping strategy) proposed earlier, it is more straightforward to us a minimum error entropy (MEE) based stochastic control algorithm to design a controller for tracking errors. It is more appropriate than the traditional minimum mean square error (MSE) criterion when dealing with nonlinearities and non-Gaussian disturbances. Shannon entropy is the most important and commonly used method in MEE based stochastic control [9–11]. A well-known generalization of Shannon entropy is Renyi entropy. When the order of Renyi entropy approaches 1, Renyi entropy will reduce and become Shannon entropy. The argument of the log in Renyi entropy is named the information potential (IP) and used as an alternative entropy criterion because of its monotonic property. Puya *et al.* [12] applied the minimum Renyi entropy control scheme to decrease the closed-loop randomness of the output under an iterative learning control (ILC) basis for general nonlinear and unknown non-Gaussian stochastic systems. Besides, information potential (IP) based dynamic neural networks were used to perform the modeling and control of the plant. In [13], the quadratic IP of tracking errors was employed to design controllers for nonlinear multivariate and non-Gaussian systems. (h, ϕ) -entropy is the most generalized definition of entropy [14]. (h, ϕ) -entropy has been employed in stochastic control systems [15,16]. Ren *et al.* [15] proposed a new tracking control algorithm for a class of networked control systems (NCSs) with non-Gaussian random disturbances and delays. Zhang *et al.* [16] presented an improved single neuron controller for multivariable stochastic systems with non-Gaussianities and unmodeled dynamics by minimizing (h, ϕ) -entropy of tracking errors.

The Shannon and order- α Renyi entropies of a continuous random variable are both defined based on the probability density function. This kind of entropy has several drawbacks. (1) The definition will be ill-suited for the case in which PDF does not exist; (2) the value can be negative; (3) the approximation using empirical distribution is impossible in general. And the IP criterion is conservative. It should be maximized to achieve smaller errors only when $\alpha > 1$. Some new definitions of entropy have been made to solve these problems. Rao *et al.* [17] proposed the cumulative residual entropy (CRE), which is defined based on the cumulative distribution function. In [18], Zografos and Nadarajah proposed the survival exponential and the generalized survival exponential entropies, both of which are broad entropy definitions based on the survival function and include CRE as a special case.

In this paper, survival information potential (SIP), proposed by Chen *et al.* [19], will be utilized to construct the performance index of stochastic control systems. Compared with the MEE criterion, adding a bias term to the tracking error [6,9,10,13] would not be necessary because of the shift-variance property of the SIP. In the previous work of MEE control [9–16], although the randomness of control input exists in practical conditions, the control input was considered as a deterministic variable for

simplicity, which is unsuitable and conservative. Therefore, in this paper, the performance index with the integration of SIPs of the control input and the tracking error is proposed.

The predictive control idea was brought up as an industrial approach to process control in the 1970s. Today this technique is the most frequently applied advanced process control method in the industry. The stochastic distribution control algorithms have been extended by the advanced control algorithm to increase the control performance. So far mainly two classes of algorithms have been developed for systems affected by stochastic noise and subject to probabilistic state and/or input constraints. (1) The randomized, or scenario-based approach [20–22]: It is a very general methodology that can consider linear or nonlinear systems affected by noise with general distributions characterized by possibly unbounded and nonconvex support. (2) The probabilistic approximation approach [23–28]: It is based on the point-wise reformulation of probabilistic or expectation constraints in deterministic terms to be included in the MPC formulation. Reference [29] gave an overview of the main developments in the area of stochastic model predictive control (SMPC) in the past decade. It described different SMPC algorithms and the key theoretical challenges in stochastic predictive control without undue mathematical complexity. However, the above results were obtained under the assumption that the system variables obey Gaussian distribution, and only mean value and variance were considered. As presented above, PDF contains the whole characteristics of random variables. From this perspective, SMPC approaches to a class of nonlinear systems with unbounded stochastic uncertainties were proposed in [30,31]. In [30], the Fokker–Planck equation was used for describing the dynamic evolution of the states' PDFs and the closed-loop stability was ensured by designing a stability constraint in terms of a stochastic control Lyapunov function. Polynomial chaos expansions were utilized to propagate the probabilistic parametric uncertainties through the system model in [31]. In the framework of SDC, SMPC was used to control the molecular weight distribution (MWD) [32,33] with the existence of non-Gaussian noises. As mentioned previously, considering the non-Gaussian SMPC in the statistical information framework may be an alternative and effective method.

Based on the preliminary work [19,34–37], in this paper, a single neuron stochastic predictive control method for nonlinear stochastic discrete systems affected by non-Gaussian noise is proposed. The proposed algorithm will be detailed in the rest of the sections. In Section 2, the models of the nonlinear stochastic system and single neuron controller are firstly presented. Then a new SIP-based predictive criterion, which contains both the randomness of tracking errors and the control input, is formulated. Based on the established models and the new criterion, the single neuron stochastic predictive control (SNSPC) algorithm is derived and the online computation procedure is also summarized. To analyze the convergence of the proposed control algorithm, the energy conversion principle is used in Section 3. A numerical simulation example is introduced to illustrate the efficiency of the proposed control strategy in Section 4. The last section concludes this paper.

2. Single Neuron Stochastic Predictive Control (SNSPC)

2.1. System Model and Single Neuron Controller

A general discrete, nonlinear dynamic system with process disturbances and measurement noises is considered here. It is described in the state-space form as:

$$\begin{cases} \mathbf{x}_{k+1} = f(\mathbf{x}_k, u_k, \zeta_k) \\ y_k = g(\mathbf{x}_k, u_k, v_k) \end{cases} \quad (1)$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state, $y_k \in \mathbb{R}$ is the measured output, and $u_k \in \mathbb{R}$ is the control input. ζ_k and v_k are external disturbances in the input and the measurement channels, respectively and they may not follow Gaussian distributions. $f(\cdot)$ and $g(\cdot)$ are known nonlinear functions that represent the system dynamics.

The goal of the control design is to find the optimal control input to make the output track the desired set-point r_k as soon as possible. Due to the nonlinearity and disturbances involved in the

system (1), the traditional PID control algorithm cannot have excellent performance here. In the past work, a single neuron adaptive PID controller has the advantages of neural network intelligent control and PID control. Single neuron adaptive PID controller achieves control system’s self-organization, adaptive function by adjusting weighting coefficients with the changes of the control parameters [35]. The controller is a multiple input and single output nonlinear processing unit with self-learning ability. The output of the single neuron adaptive PID controller is shown as Equation (2):

$$u_k = u_{k-1} + K \sum_{l=1}^3 w_{lk} x_{lk} / \|\Sigma_k\| \tag{2}$$

where $\Sigma_k = \sum_{l=1}^3 w_{lk}$, $K > 0$ is the proportional coefficient of the neuron. w_{lk} ($l = 1, 2, 3$) stands for the weight corresponding to each input. x_{lk} ($l = 1, 2, 3$) is the input of the neuron from the tracking error e_k and it can be defined as:

$$\begin{cases} x_{1k} = e_k \\ x_{2k} = e_k - e_{k-1} \\ x_{3k} = e_k - 2e_{k-1} + e_{k-2} \end{cases} \tag{3}$$

Remark 1. x_{lk} ($l = 1, 2, 3$) is the input of the neuron from the tracking error e_k . The weights w_{lk} ($l = 1, 2, 3$) reflect the dynamic characteristics of the controlled object and the process response. The neuron weights can be updated through self-learning strategy, and the tracking error can be correspondingly eliminated and approach to the steady state effected by three kinds of control parts together.

2.2. SIP-Based Predictive Criterion

If noises ζ_k and v_k are not Gaussian, the variance of the output tracking error $e_k = r_k - y_k$ would not be sufficient to characterize the randomness of the tracking error dynamic systems. Therefore, an alternative measure of uncertainty, survival information potential (SIP), is given to construct the performance index for measuring the dispersion of the stochastic systems.

Firstly, the definition and properties of SIP are presented as follows.

Definition 1. Definition of SIP [19]: For a random vector \mathbf{X} in \mathbb{R}^m , SIP of order α ($\alpha > 0$) is defined by

$$S_\alpha(\mathbf{X}) = \int_{\mathbb{R}_+^m} \bar{F}_{|\mathbf{X}|}^\alpha(\mathbf{x}) d\mathbf{x} \tag{4}$$

where $\bar{F}_{|\mathbf{X}|}(\mathbf{x}) = P(|\mathbf{X}| > \mathbf{x}) = E[I(|\mathbf{X}| > \mathbf{x})]$ is the multivariate survival function (or equivalently, the distribution function) of the random vector $|\mathbf{X}|$, and $\mathbb{R}_+^m = \{\mathbf{x} \in \mathbb{R}^m : \mathbf{x} = (x_1, \dots, x_m), x_i \geq 0, i = 1, \dots, m\}$. Note that $|\mathbf{X}| > \mathbf{x}$ means $|X_i| > x_i, i = 1, \dots, m$ and $I(\cdot)$ is the indicator function. Similarly, when $\alpha = 2$, SIP is called the quadratic survival information potential (QSIP).

Property 1. Property 1 of SIP [19]: $S_\alpha(\mathbf{X}) \geq 0$, with equality if and only if $P(\mathbf{X} = \mathbf{0}) = 1$.

Property 2. Property 2 of SIP [19]: Let \mathbf{X} be an m -dimensional random vector, and let another m -dimensional random vector $\mathbf{Y} = (Y_1, Y_2, \dots, Y_m)^T$ with $Y_i = c_i X_i, c_i \in \mathbb{R}, i = 1, 2, \dots, m$, then $S_\alpha(\mathbf{Y}) = (\prod_{i=1}^m |c_i|) S_\alpha(\mathbf{X})$.

Remark 2. SIP has some advantages over the IP: (1) It has consistent definition in the continuous and discrete domains; (2) it is not shift-invariant (i.e., its value would vary with the location of distribution); (3) it can be computed more easily from the sample data (without kernel computation and the choice of the kernel width), and the estimation asymptotically converges to the true value; (4) it is a more

robust measure since the distribution function is more regular than the density function. (Note that the density is computed as the derivative of the distribution.)

Because of the nonlinearity and noises involved in the system, the tracking error is a non-Gaussian stochastic variable at a typical instant. Therefore, the control input is also non-Gaussian according to the single neuron controller Equations (2) and (3). However, in the previous work on stochastic distribution control [9–13,15,16], the control input is considered as a deterministic variable, which is conservative. And moreover, the multi-step predictive control strategy which can achieve better performance in industrial processes is adopted to design the optimal control input here.

The cost function of a SIP-based predictive controller for set point control is similar to that of a conventional predictive control with a continuous actor, but the SIP criterion is used in the control design instead of the MSE criterion

$$J = \sum_{i=1}^P S_{\alpha}(e_{k+i-1}) + \lambda \sum_{j=1}^M S_{\alpha}(u_{k+j-1}) \tag{5}$$

where e_{k+i} is the i -step ahead prediction of the system tracking error. P and M are the prediction horizon and the control horizon, respectively, and $M \leq P$. It is assumed that the control variable would not change after M steps, *i.e.*, $u_{k+j-1} = u_{k+M-1}$ ($j > M$).

2.3. SNSPC Algorithm

Based on the structure of the single neuron adaptive controller Equations (2) and (3) and the performance index Equation (5), the optimal weights of the single neuron can be solved by the stochastic gradient method

$$\mathbf{w}_k^* = \underset{\mathbf{w}_k \in \mathbb{R}^{3M}}{\operatorname{argmin}} J = \underset{\mathbf{w}_k \in \mathbb{R}^{3M}}{\operatorname{argmin}} \left\{ \sum_{i=1}^P S_{\alpha}(e_{k+i-1}) + \lambda \sum_{j=1}^M S_{\alpha}(u_{k+j-1}) \right\} \tag{6}$$

where $\mathbf{w}_k = [w_{1k} \ w_{2k} \ w_{3k} \ w_{1,k+1} \ w_{2,k+1} \ w_{3,k+1} \ \cdots \ w_{1,k+M-1} \ w_{2,k+M-1} \ w_{3,k+M-1}]^T \in \mathbb{R}^{3M}$.

According to the definition of SIP, the optimal weights \mathbf{w}_k^* can be calculated from

$$\frac{\partial J}{\partial \mathbf{w}_k} = \frac{\partial}{\partial \mathbf{w}_k} \left\{ \sum_{i=1}^P \int_{-\infty}^{+\infty} \bar{F}_{|e_{k+i-1}|}(x) dx + \lambda \sum_{j=1}^M \int_{-\infty}^{+\infty} \bar{F}_{|u_{k+j-1}|}(z) dz \right\} = 0 \tag{7}$$

where $\bar{F}_{|e_{k+i-1}|}$ and $\bar{F}_{|u_{k+j-1}|}$ are the survival functions of the tracking error e_{k+i-1} and the control input u_{k+j-1} , respectively.

In the practical application, it is generally difficult to develop a comprehensive first-principle model to describe the distribution of the error e_k . Instead of the theoretical SIP, the data-driven empirical SIP used as the cost function would be useful alternative to timely online prediction of the errors. Given a sequence of error samples $(e_{1k}, e_{2k}, \dots, e_{\bar{N}k})$, assuming, without the loss of generality, that $|e_{1k}| \leq |e_{2k}| \leq \dots \leq |e_{\bar{N}k}|$, the empirical SIP would be [20]

$$\hat{S}_{\alpha}(e_k) = \sum_{j=1}^{\bar{N}} \mu_j |e_{jk}| \tag{8}$$

where $\mu_j = \left(\frac{\bar{N}-j+1}{\bar{N}}\right)^{\alpha} - \left(\frac{\bar{N}-j}{\bar{N}}\right)^{\alpha}$.

Similarly, the empirical SIP of u_{k+j-1} , $\hat{S}_{\alpha}(u_{k+j-1})$, can also be formulated. The empirical cost Equation (8) is the weighted sum of the ordered absolute errors. One drawback of this empirical SIP

criterion is that it is not smooth at $e_{jk} = 0$ and $u_{jk} = 0$. To address this problem, a more general cost can be defined using the empirical SIP of any mapped errors and inputs; that is,

$$\bar{J}(\mathbf{e}_{k:k+P-1}) = \sum_{i=1}^P \hat{S}_\alpha(\phi(e_{k+i-1})) + \lambda \sum_{j=1}^M \hat{S}_\alpha(\phi(u_{k+j-1})) = \sum_{i=1}^P \sum_{l=1}^{\bar{N}} \mu_l \phi(e_{l,k+i-1}) + \lambda \sum_{j=1}^M \sum_{l=1}^{\bar{N}} \sigma_l \phi(u_{l,k+j-1}) \quad (9)$$

where $\mathbf{e}_{k:k+P-1} = [e_k \ e_{k+1} \ \dots \ e_{k+P-1}]^T$, $\mu_l = \left(\frac{\bar{N}-l+1}{\bar{N}}\right)^\alpha - \left(\frac{\bar{N}-l}{\bar{N}}\right)^\alpha$. The mapping function $\phi(\cdot)$ usually satisfies

$$\begin{cases} (i) \text{ positivity : } \phi(\cdot) \geq 0 \\ (ii) \text{ symmetry : } \phi(x) = \phi(-x) \\ (iii) \text{ monotonicity : } |x_1| < |x_2| \Rightarrow \phi(x_1) \leq \phi(x_2) \end{cases} \quad (10)$$

Remark 3. As the data-driven empirical SIP-based performance index is not smooth at $e_{jk} = 0$ and $u_{jk} = 0$, a proper function transformation method can be used to establish an equivalent SIP-based criterion, which can be formulated as Equation (9) and (10). In the simulation, the empirical SIP of the square errors $(e_{1k}^2, e_{2k}^2, \dots, e_{Nk}^2)$ and the square control inputs $(u_{1k}^2, u_{2k}^2, \dots, u_{Nk}^2)$ as an alternative adaptation cost is used here, given by

$$\bar{J} = \sum_{i=1}^P \hat{S}_\alpha(e_{k+i-1}^2) + \lambda \sum_{j=1}^M \hat{S}_\alpha(u_{k+j-1}^2) = \sum_{i=1}^P \sum_{l=1}^{\bar{N}} \mu_l e_{l,k+i-1}^2 + \lambda \sum_{j=1}^M \sum_{l=1}^{\bar{N}} \sigma_l u_{l,k+j-1}^2$$

The above cost is the weighted sum of the ordered square errors with the popular minimum mean square error (MSE) criterion as a special case (when $\alpha = 1$).

Based on the general cost Equation (9), the weight update equation for adaptive system training is

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta \frac{\partial \bar{J}}{\partial \mathbf{w}_k} = \mathbf{w}_k - \eta \frac{\partial \mathbf{e}_{k:k+P-1}^T}{\partial \mathbf{w}_k} \frac{\partial \bar{J}}{\partial \mathbf{e}_{k:k+P-1}} \quad (11)$$

where $\eta > 0$ denotes the step-size (or adaptation gain).

According to the above presentation, the optimal input can be computed. The procedure below describes the steps of implementing SNSPC specifically.

Step 1: Initialize the weight vector of the single neuron \mathbf{w}_0 . Choose the α value, step-size η , and the sliding window length \bar{N} .

Step 2: Estimate the performance index using empirical SIPs of the output tracking error and the control input in (9).

Step 3: Solve the optimal weight vector \mathbf{w}_k^* by (11).

Step 4: Compute the next control input

$$u_k = u_{k-1} + \frac{K}{\|\Sigma_k\|} \begin{bmatrix} x_{1k} & x_{2k} & x_{3k} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times (3P-3)} \end{bmatrix} \mathbf{w}_k^* \quad (12)$$

and implement the input on the process.

Step 5: Collect the process outputs to update the SIPs of the output tracking error and the control input. Then repeat the procedure from *Step 2* to *Step 5* for the next time step, $k = k + 1$.

Remark 4. The computations of the SIP-based performance index in Equation (9) are easier than the minimum error entropy (MEE) criterion, because the index is directly computed from the sample data without kernel computation and the choice of the kernel width, but the major computations of the partial derivative, $\frac{\partial \mathbf{e}_{k:k+P-1}^T}{\partial \mathbf{w}_k}$, in the control algorithm Equation (12) would be done.

3. Mean-Square Stability

In order to analyze the convergence of the proposed control algorithm, the nonlinear stochastic system Equation (1) is firstly linearized as

$$\begin{cases} \Delta \mathbf{x}_{k+1} = \frac{\partial f}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{k-1}} \Delta \mathbf{x}_k + \frac{\partial f}{\partial u} \Big|_{u_{k-1}} \Delta u_k + \frac{\partial f}{\partial \zeta} \Big|_{\zeta_{k-1}} \Delta \zeta_k \\ \Delta y_k = \frac{\partial g}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{k-1}} \Delta \mathbf{x}_k + \frac{\partial g}{\partial u} \Big|_{u_{k-1}} \Delta u_k + \frac{\partial g}{\partial v} \Big|_{v_{k-1}} \Delta v_k \end{cases} \quad (13)$$

where $\Delta u_k = u_k - u_{k-1}$, $\Delta \mathbf{x}_k = \mathbf{x}_k - \mathbf{x}_{k-1} = \tilde{\mathbf{x}}_k$, $\Delta \zeta_k = \zeta_k - \zeta_{k-1} = \tilde{\zeta}_k$, $\Delta y_k = y_k - y_{k-1} = \tilde{y}_k$, and $\Delta v_k = v_k - v_{k-1} = \tilde{v}_k$. It can be simply denoted as

$$\begin{cases} \tilde{\mathbf{x}}_{k+1} = \mathbf{A}_k \tilde{\mathbf{x}}_k + \mathbf{B}_k \Delta u_k + \mathbf{N}_{1k} \tilde{\zeta}_k \\ \tilde{y}_k = \mathbf{C}_k \tilde{\mathbf{x}}_k + \mathbf{D}_k \Delta u_k + \mathbf{N}_{2k} \tilde{v}_k \end{cases} \quad (14)$$

where $\mathbf{A}_k = \frac{\partial f}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{k-1}}$, $\mathbf{B}_k = \frac{\partial f}{\partial u} \Big|_{u_{k-1}}$, $\mathbf{C}_k = \frac{\partial g}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{k-1}}$, $\mathbf{D}_k = \frac{\partial g}{\partial u} \Big|_{u_{k-1}}$, $\mathbf{N}_{1k} = \frac{\partial f}{\partial \zeta} \Big|_{\zeta_{k-1}}$, $\mathbf{N}_{2k} = \frac{\partial g}{\partial v} \Big|_{v_{k-1}}$.

Based on the state space representation, an extended state space model can be formulated as [36]

$$\mathbf{y}_k^f = \mathbf{L}_k^v \mathbf{m}_k^p + \mathbf{L}_k^u \mathbf{u}_k^f + \mathbf{L}_k^\zeta \boldsymbol{\zeta}_k^f \quad (15)$$

where $\mathbf{y}_k^f = [\tilde{y}_k \ \cdots \ \tilde{y}_{k+M-1}]^T$, $\mathbf{u}_k^f = [\Delta u_k \ \cdots \ \Delta u_{k+M-1}]^T$, $\mathbf{m}_k^p = \left[\begin{matrix} (\mathbf{v}_{1k}^p)^T & (\mathbf{v}_{2k}^p)^T \end{matrix} \right]^T$, $\mathbf{v}_{1k}^p = [\tilde{y}_{k-M} \ \cdots \ \tilde{y}_{k-1}]^T$, $\mathbf{v}_{2k}^p = [\Delta u_{k-M} \ \cdots \ \Delta u_{k-1}]^T$. \mathbf{L}_k^v , \mathbf{L}_k^u and \mathbf{L}_k^ζ can be obtained from Equation (14); $\boldsymbol{\zeta}_k^f$ is the composite noise, including external noises (ζ_k and v_k) and the model mismatch randomness, and it may be non-Gaussian noises.

From Equation (2), Equation (15) can be reformulated as

$$\mathbf{y}_k^f = \mathbf{L}_k^v \mathbf{m}_k^p + \mathbf{L}_k^u \boldsymbol{\chi}_k \mathbf{w}_k + \mathbf{L}_k^\zeta \boldsymbol{\zeta}_k^f \quad (16)$$

where $\boldsymbol{\chi}_k$ is the input matrices of the single neural network, which consists of tracking errors:

$$\boldsymbol{\chi}_k = \begin{bmatrix} \mathbf{z}_k & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{z}_k & \mathbf{z}_{k+1} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{z}_k & \mathbf{z}_{k+1} & \cdots & \mathbf{z}_{k+M} \end{bmatrix}$$

where $\mathbf{z}_k = \left[\begin{matrix} \frac{K}{\|\Sigma_k\|} x_{1k} & \frac{K}{\|\Sigma_k\|} x_{2k} & \frac{K}{\|\Sigma_k\|} x_{3k} \end{matrix} \right]$ and $\mathbf{0} = [0 \ 0 \ 0]$.

Thus, the difference between the future predictions and the set-point trajectory is

$$\mathbf{e}_k^f = \mathbf{L}_k^v \mathbf{m}_k^p + \mathbf{L}_k^u \boldsymbol{\chi}_k \mathbf{w}_k + \mathbf{L}_k^\zeta \boldsymbol{\zeta}_k^f - \Delta \mathbf{r}_k^f \quad (17)$$

where $\Delta \mathbf{r}_k^f = \mathbf{r}_k^f - \mathbf{r}_{k-1}^f$ is the change of the set-point at time point k and $\boldsymbol{\theta}_k = \mathbf{L}_k^v \mathbf{m}_k^p + \mathbf{L}_k^\zeta \boldsymbol{\zeta}_k^f - \Delta \mathbf{r}_k^f$, which is a new random vector; then Equation (17) can be rewritten as

$$\mathbf{e}_k^f = \mathbf{L}_k^u \boldsymbol{\chi}_k \mathbf{w}_k + \boldsymbol{\theta}_k \quad (18)$$

Now a priori error vector and a posteriori error vector, \mathbf{e}_k^{af} and \mathbf{e}_k^{pf} , are defined as:

$$\begin{cases} \mathbf{e}_k^{af} = \mathbf{L}_k^u \boldsymbol{\chi}_k \mathbf{w}_k \\ \mathbf{e}_k^{pf} = \mathbf{L}_k^u \boldsymbol{\chi}_k \mathbf{w}_{k+1} \end{cases} \quad (19)$$

Obviously, \mathbf{e}_k^{af} and \mathbf{e}_k^{pf} have the following relationship

$$\mathbf{e}_k^{pf} = \mathbf{e}_k^{af} + \mathbf{L}_k^u \chi_k (\mathbf{w}_{k+1} - \mathbf{w}_k) \tag{20}$$

By incorporating Equation (11), the following equation is gotten,

$$\mathbf{e}_k^{pf} = \mathbf{e}_k^{af} - \mathbf{L}_k^u \chi_k \eta \frac{\partial (\mathbf{e}_k^f)^\top}{\partial \mathbf{w}_k} \frac{\partial \bar{J}}{\partial \mathbf{e}_k^f} = \mathbf{e}_k^{af} + \eta \mathbf{L}_k^u \chi_k \chi_k^\top \mathbf{L}_k^u \bar{\mathbf{h}} (\mathbf{e}_k^f) \tag{21}$$

where $\bar{\mathbf{h}} (\mathbf{e}_k^f) = \frac{\partial \bar{J}}{\partial \mathbf{e}_k^f} = \left[\frac{\partial \bar{J}}{\partial e_k} \quad \frac{\partial \bar{J}}{\partial e_{k+1}} \quad \dots \quad \frac{\partial \bar{J}}{\partial e_{k+M-1}} \right]^\top$. $\mathbf{R}_k = \mathbf{L}_k^u \chi_k (\mathbf{L}_k^u \chi_k)^\top$ is an $M \times M$ -dimensional symmetric matrix. Assume \mathbf{R}_k is invertible (i.e., $\det \mathbf{R}_k \neq 0$),

$$\begin{aligned} \mathbf{e}_k^{pf} &= \mathbf{e}_k^{af} + \eta \mathbf{R}_k \bar{\mathbf{h}} (\mathbf{e}_k^f) \\ \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) &= \eta \bar{\mathbf{h}} (\mathbf{e}_k^f) \\ \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) &= \eta \chi_k^\top (\mathbf{L}_k^u)^\top \bar{\mathbf{h}} (\mathbf{e}_k^f) \\ \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) &= (\mathbf{w}_{k+1} - \mathbf{w}_k) \end{aligned} \tag{22}$$

And hence

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) \tag{23}$$

Both sides of Equation (23) should have the same energy; that is

$$\begin{aligned} \mathbf{w}_{k+1}^\top \mathbf{w}_{k+1} &= \left[\mathbf{w}_k + \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) \right]^\top \times \left[\mathbf{w}_k + \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) \right] \\ &= \mathbf{w}_k^\top \mathbf{w}_k + 2 \mathbf{w}_k^\top \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) \\ &\quad + (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af})^\top \mathbf{R}_k^{-1} (\mathbf{L}_k^u)^\top \chi_k \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) \end{aligned} \tag{24}$$

From $\mathbf{R}_k = \mathbf{L}_k^u \chi_k (\mathbf{L}_k^u \chi_k)^\top$, we have

$$\begin{aligned} \mathbf{w}_{k+1}^\top \mathbf{w}_{k+1} &= \mathbf{w}_k^\top \mathbf{w}_k + 2 \mathbf{w}_k^\top \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) + (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af})^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) \\ &= \mathbf{w}_k^\top \mathbf{w}_k + 2 \mathbf{w}_k^\top \chi_k^\top (\mathbf{L}_k^u)^\top \mathbf{R}_k^{-1} (\mathbf{e}_k^{pf} - \mathbf{e}_k^{af}) + (\mathbf{e}_k^{pf})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{pf} \\ &\quad - 2 (\mathbf{e}_k^{af})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{pf} + (\mathbf{e}_k^{af})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{af} \end{aligned} \tag{25}$$

Since $\mathbf{R}_k^{-1} = (\mathbf{L}_k^u)^{-\top} \chi_k^{-\top} \chi_k^{-1} (\mathbf{L}_k^u)^{-1}$, $\mathbf{e}_k^{af} = \mathbf{L}_k^u \chi_k \mathbf{w}_k$, then

$$\begin{aligned} \mathbf{w}_{k+1}^\top \mathbf{w}_{k+1} &= \mathbf{w}_k^\top \mathbf{w}_k + 2 \mathbf{w}_k^\top \chi_k^\top (\mathbf{L}_k^u)^\top (\mathbf{L}_k^u)^{-\top} \chi_k^{-\top} \chi_k^{-1} (\mathbf{L}_k^u)^{-1} (\mathbf{e}_k^{pf} - \mathbf{L}_k^u \chi_k \mathbf{w}_k) \\ &\quad + (\mathbf{e}_k^{pf})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{pf} - 2 (\mathbf{L}_k^u \chi_k \mathbf{w}_k)^\top (\mathbf{L}_k^u)^{-\top} \chi_k^{-\top} \chi_k^{-1} (\mathbf{L}_k^u)^{-1} \mathbf{e}_k^{pf} + (\mathbf{e}_k^{af})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{af} \\ &= \mathbf{w}_k^\top \mathbf{w}_k + 2 \mathbf{w}_k^\top \chi_k^{-1} (\mathbf{L}_k^u)^{-1} \mathbf{e}_k^{pf} - 2 \mathbf{w}_k^\top \chi_k^{-1} (\mathbf{L}_k^u)^{-1} \mathbf{L}_k^u \chi_k \mathbf{w}_k \\ &\quad + (\mathbf{e}_k^{pf})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{pf} - 2 \mathbf{w}_k^\top \chi_k^{-1} (\mathbf{L}_k^u)^{-1} \mathbf{e}_k^{pf} + (\mathbf{e}_k^{af})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{af} \\ &= -\mathbf{w}_k^\top \mathbf{w}_k + (\mathbf{e}_k^{pf})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{pf} + (\mathbf{e}_k^{af})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{af} \end{aligned} \tag{26}$$

Adding $(\mathbf{e}_k^{af})^\top \mathbf{R}_k^{-1} \mathbf{e}_k^{af}$ to both sides of the above equation and substituting $\mathbf{e}_k^{af} = \mathbf{L}_k^u \chi_k \mathbf{w}_k$ and $\mathbf{R}_k^{-1} = (\mathbf{L}_k^u)^{-\top} \chi_k^{-\top} \chi_k^{-1} (\mathbf{L}_k^u)^{-1}$ into it, one can calculate the energy conservation relation of $\mathbf{w}_{k+1}^\top \mathbf{w}_{k+1}$ and $\mathbf{w}_k^\top \mathbf{w}_k$:

$$\begin{aligned}
 & \mathbf{w}_{k+1}^T \mathbf{w}_{k+1} + \left(\mathbf{e}_k^{af}\right)^T \mathbf{R}_k^{-1} \mathbf{e}_k^{af} \\
 &= -\mathbf{w}_k^T \mathbf{w}_k + \left(\mathbf{e}_k^{pf}\right)^T \mathbf{R}_k^{-1} \mathbf{e}_k^{pf} + 2 \left(\mathbf{L}_k^u \boldsymbol{\chi}_k \mathbf{w}_k\right)^T \left[\left(\mathbf{L}_k^u\right)^{-T} \boldsymbol{\chi}_k^{-T} \boldsymbol{\chi}_k^{-1} \left(\mathbf{L}_k^u\right)^{-1} \right] \left(\mathbf{L}_k^u \boldsymbol{\chi}_k \mathbf{w}_k\right) \\
 & \mathbf{w}_{k+1}^T \mathbf{w}_{k+1} + \left(\mathbf{e}_k^{af}\right)^T \mathbf{R}_k^{-1} \mathbf{e}_k^{af} = \mathbf{w}_k^T \mathbf{w}_k + \left(\mathbf{e}_k^{pf}\right)^T \mathbf{R}_k^{-1} \mathbf{e}_k^{pf} \\
 & \|\mathbf{w}_{k+1}\|^2 + \|\mathbf{e}_k^{af}\|_{\mathbf{R}_k^{-1}}^2 = \|\mathbf{w}_k\|^2 + \|\mathbf{e}_k^{pf}\|_{\mathbf{R}_k^{-1}}^2
 \end{aligned} \tag{27}$$

where $\|\mathbf{w}_k\|^2 = \mathbf{w}_k^T \mathbf{w}_k$, $\|\mathbf{e}_k^{af}\|_{\mathbf{R}_k^{-1}}^2 \triangleq \left[\mathbf{e}_k^{af}\right]^T \mathbf{R}_k^{-1} \mathbf{e}_k^{af}$, and $\|\mathbf{e}_k^{pf}\|_{\mathbf{R}_k^{-1}}^2 \triangleq \left[\mathbf{e}_k^{pf}\right]^T \mathbf{R}_k^{-1} \mathbf{e}_k^{pf}$. To study the mean-square behavior of the algorithm, one takes expectations of both sides of (27) and write

$$E \left[\|\mathbf{w}_{k+1}\|^2 \right] + E \left[\|\mathbf{e}_k^{af}\|_{\mathbf{R}_k^{-1}}^2 \right] = E \left[\|\mathbf{w}_k\|^2 \right] + E \left[\|\mathbf{e}_k^{pf}\|_{\mathbf{R}_k^{-1}}^2 \right] \tag{28}$$

By substituting $\mathbf{e}_k^{pf} = \mathbf{e}_k^{af} + \eta \mathbf{R}_k \bar{\mathbf{h}} \left(\mathbf{e}_k^f\right)$ into Equation (28), the following equation can be obtained.

$$E \left[\|\mathbf{w}_{k+1}\|^2 \right] = E \left[\|\mathbf{w}_k\|^2 \right] + 2\eta E \left[\left(\mathbf{e}_k^{af}\right)^T \bar{\mathbf{h}} \left(\mathbf{e}_k^f\right) \right] + \eta^2 E \left[\bar{\mathbf{h}}^T \left(\mathbf{e}_k^f\right) \mathbf{R}_k \bar{\mathbf{h}} \left(\mathbf{e}_k^f\right) \right] \tag{29}$$

In order to evaluate the expectations $E \left[\left(\mathbf{e}_k^{af}\right)^T \bar{\mathbf{h}} \left(\mathbf{e}_k^f\right) \right]$ and $E \left[\bar{\mathbf{h}}^T \left(\mathbf{e}_k^f\right) \mathbf{R}_k \bar{\mathbf{h}} \left(\mathbf{e}_k^f\right) \right]$, the following assumptions [37] are used in this paper:

Assumption 1. The noise $\boldsymbol{\theta}_k$ is independent, identically distributed (i.i.d.), and independent of the input $\mathbf{L}_k^u \boldsymbol{\chi}_k$.

Assumption 2. The a priori error vector \mathbf{e}_k^{af} is jointly Gaussian distributed.

Assumption 3. The input vectors $\mathbf{L}_k^u \boldsymbol{\chi}_k$ are zero-mean independent, identically distributed (i.i.d.).

Assumption 4. $\forall ij \in \{1, \dots, M\}$, $\mathbf{R}_k(ij)$ is independent of $\{e_i, e_j\}$.

Based on the above assumptions, we have

$$\begin{aligned}
 E \left[\left(\mathbf{e}_k^{af}\right)^T \bar{\mathbf{h}} \left(\mathbf{e}_k^f\right) \right] &= E \left[\left(\mathbf{e}_k^{af}\right)^T \bar{\mathbf{h}} \left(\mathbf{e}_k^{af} + \boldsymbol{\theta}_k\right) \right] \\
 &= \left(\frac{1}{\sqrt{2\pi\sigma_k^2}} \right)^N \int \gamma_{\boldsymbol{\theta}_k} \left(\boldsymbol{\theta}_k\right) d\boldsymbol{\theta}_k \times \\
 & \int \left(\mathbf{e}_k^{af}\right)^T \bar{\mathbf{h}} \left(\mathbf{e}_k^{af} - \mathbf{L}_k^u \boldsymbol{\zeta}_k\right) \prod_{i=1}^N \exp \left(-\frac{\left(e_{k+i-1}^a\right)^2}{2\sigma_k^2} \right) d\mathbf{e}_k^{af}
 \end{aligned} \tag{30}$$

where $\gamma_{\boldsymbol{\theta}_k} \left(\boldsymbol{\theta}_k\right)$ is the density distribution of $\boldsymbol{\theta}_k$ and $\sigma_k = \sqrt{E \left[\left(e_{k+i-1}^a\right)^2 \right]}$.

Therefore, the expectation $E \left[\left(\mathbf{e}_k^{af}\right)^T \bar{\mathbf{h}} \left(\mathbf{e}_k^f\right) \right]$ can be expressed as a function of σ_k^2 . Define the following function

$$h_G \left(\sigma_k^2\right) \triangleq \frac{E \left[\left(\mathbf{e}_k^{af}\right)^T \bar{\mathbf{h}} \left(\mathbf{e}_k^f\right) \right]}{\sigma_k^2} \tag{31}$$

and then yields

$$E \left[\left(\mathbf{e}_k^{af} \right)^T \bar{\mathbf{h}} \left(\mathbf{e}_k^f \right) \right] = h_G \left(\sigma_k^2 \right) \sigma_k^2 \tag{32}$$

We now evaluate the expectation $E \left[\bar{\mathbf{h}}^T \left(\mathbf{e}_k^f \right) \mathbf{R}_k \bar{\mathbf{h}} \left(\mathbf{e}_k^f \right) \right]$.

$$\begin{aligned} E \left[\bar{\mathbf{h}}^T \left(\mathbf{e}_k^f \right) \mathbf{R}_k \bar{\mathbf{h}} \left(\mathbf{e}_k^f \right) \right] &= \sum_{i=1}^N \sum_{j=1}^N E \left[\bar{h}_i \left(\mathbf{e}_k^f \right) \bar{h}_j \left(\mathbf{e}_k^f \right) \mathbf{R}_k \left(ij \right) \right] \\ &= \sum_{i=1}^N E \left[\bar{h}_i^2 \left(\mathbf{e}_k^f \right) \right] E \left[\mathbf{R}_k \left(ii \right) \right] \end{aligned} \tag{33}$$

where:

$$h_I \left(\sigma_k^2 \right) \triangleq \sum_{i=1}^N E \left[\bar{h}_i^2 \left(\mathbf{e}_k^f \right) \right] \tag{34}$$

Then it is gotten:

$$E \left[\bar{\mathbf{h}}^T \left(\mathbf{e}_k^f \right) \mathbf{R}_k \bar{\mathbf{h}} \left(\mathbf{e}_k^f \right) \right] = h_I \left(\sigma_k^2 \right) E \left[\mathbf{R}_k \left(ii \right) \right] \tag{35}$$

By substituting Equations (32) and (35) into Equation (29), the following equation is gotten.

$$E \left[\left\| \mathbf{w}_{k+1} \right\|^2 \right] = E \left[\left\| \mathbf{w}_k \right\|^2 \right] + 2\eta h_G \left(\sigma_k^2 \right) \sigma_k^2 + \eta^2 h_I \left(\sigma_k^2 \right) E \left[\mathbf{R}_k \left(ii \right) \right] \tag{36}$$

Then the convergent condition for the sequence $E \left[\left\| \mathbf{w}_k \right\|^2 \right]$ (i.e., $E \left[\left\| \mathbf{w}_{k+1} \right\|^2 \right] \leq E \left[\left\| \mathbf{w}_k \right\|^2 \right]$) would be:

$$\eta \leq \frac{-2}{E \left[\mathbf{R}_k \left(ii \right) \right]} \sup_{k \geq 0} \left\{ \frac{\sigma_k^2 h_G \left(\sigma_k^2 \right)}{h_I \left(\sigma_k^2 \right)} \right\} \tag{37}$$

Remark 5. Since the nonlinear system can be approximated by a linear system at the equilibrium, the convergence of the nonlinear system Equation (1) using the linearizing method is studied here.

4. Simulation Results

In order to illustrate the efficiency of the presented SNSPC algorithm, consider a nonlinear stochastic system described by

$$\begin{cases} x_{k+1} = \frac{1}{(1+x_k)^2} + 1.2u_k + \sqrt{\zeta_k} \\ y_k = x_k + 0.1\sin x_k + v_k \end{cases} \tag{38}$$

At each time point k , PDFs of ζ_k and v_k are given by

$$\begin{aligned} \gamma_\zeta(x) &= \begin{cases} \beta^{-1} (\alpha_1 + 1, \lambda_1 + 1) x^{\alpha_1} (1-x)^{\lambda_1} & , x \in (0, 1) \\ 0, & \text{otherwise} \end{cases} \\ \gamma_v(x) &= \begin{cases} [2^{\alpha_2 + \lambda_2 + 1} \beta (\alpha_1 + 1, \lambda_1 + 1)]^{-1} x^{\alpha_2} (2-x)^{\lambda_2} & , x \in (0, 2) \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{39}$$

where $\beta(\alpha + 1, \lambda + 1) = \int_0^1 x^\alpha (1-x)^\lambda dx$, $\alpha_1 = 1$, $\lambda_1 = 2$, $\alpha_2 = 2$, $\lambda_2 = 1$.

In this example, the set point of the system Equation (38) is set to be $r_k = \begin{cases} 1, & 0s \leq k < 50s \\ 2, & 50s \leq k < 100s \\ 3, & k \geq 100s \end{cases}$.

The prediction horizon and control horizon, $P = M = 5$, are chosen. The sampling period is $T = 1$. The initial control input ($u_0 = 0$) and the state variable ($x_0 = 0$) are selected. The step size in Equation (12) is $\eta = 0.0001$. In the performance function Equation (9), we choose $\phi(x) = x^2$, $\alpha = 2$ and $\lambda = 0.5$.

In this simulation, the corresponding control inputs based on SIP and entropy criteria respectively are implemented on the system Equation (38) at each time point. Some comparative results are given to illustrate the superiority of the proposed SIP based stochastic predictive tracking control algorithm. In Figure 1, it is clear that the proposed control algorithm based on the SIP criterion has better performance, the fluctuation of the output response is smaller, and the response is quicker using the SIP method. In Figure 2, the trend of the performance index Equation (9) is presented. It is found that the performance index is overall decreasing with the progress of the time although some small variations can be recorded. The variances of the single neuron weights are presented in Figure 3. Figures 4 and 5 demonstrate the PDFs of the output tracking errors using the entropy-based controller, while Figures 6 and 7 illustrate the PDFs of the tracking errors using the proposed SNSPC method. Compared with Figure 4, the shape of the PDF of the tracking error in Figure 6 turns to be narrower and sharper over the control process, which indicates that the proposed SNSPC control system has achieved a smaller uncertainty in the tracking error distribution. In addition, it can also be seen that the peak of the tracking error PDF locates in the vicinity of zero using the proposed method. In Figure 8, the final PDF of the tracking error under the proposed control law is sharper and narrower than that under the entropy-based control law.

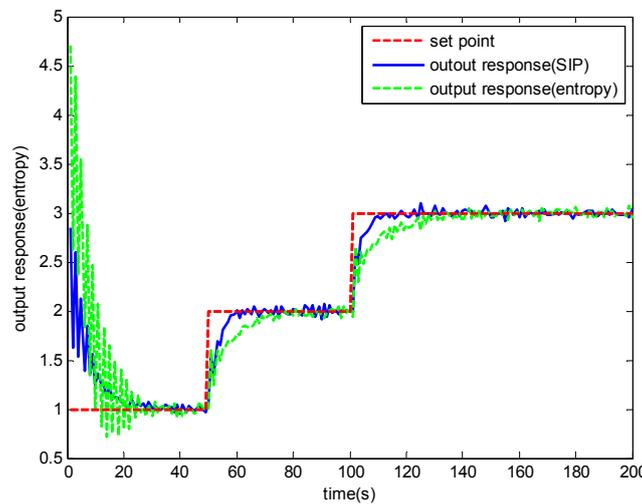


Figure 1. Output responses.

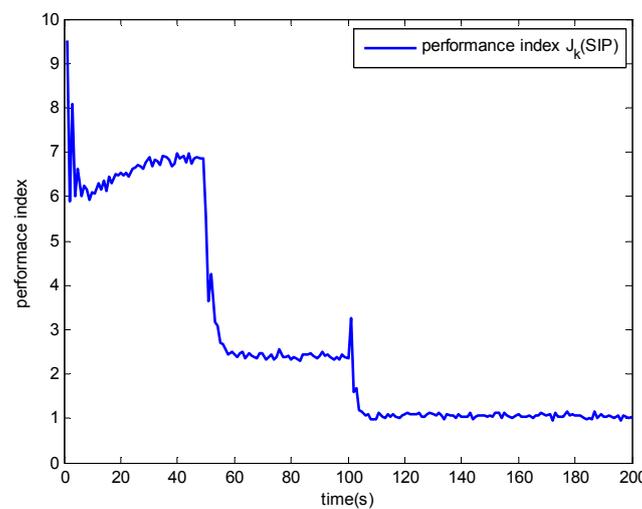


Figure 2. The performance index.

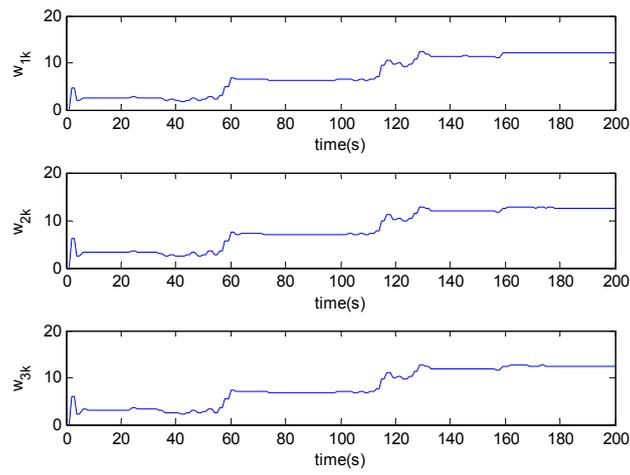


Figure 3. Weights of a single neuron.

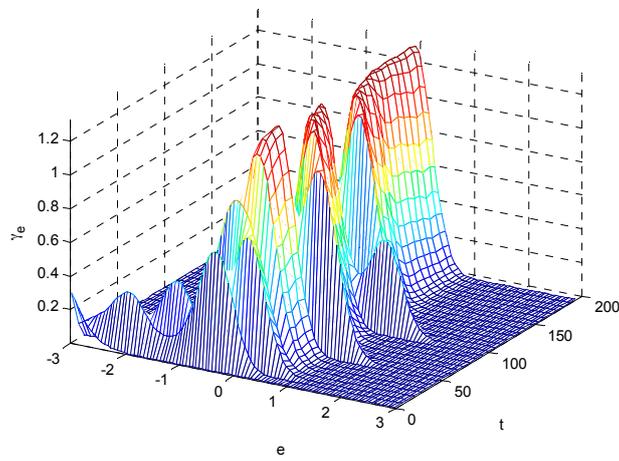


Figure 4. 3D mesh PDF using the entropy based controller.

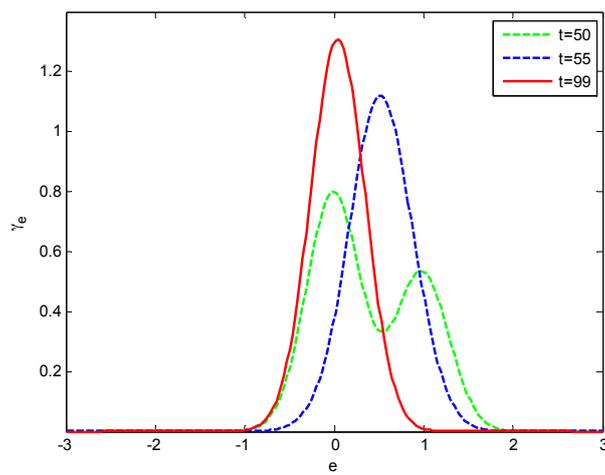


Figure 5. Selected tracking error PDFs using the entropy based controller.

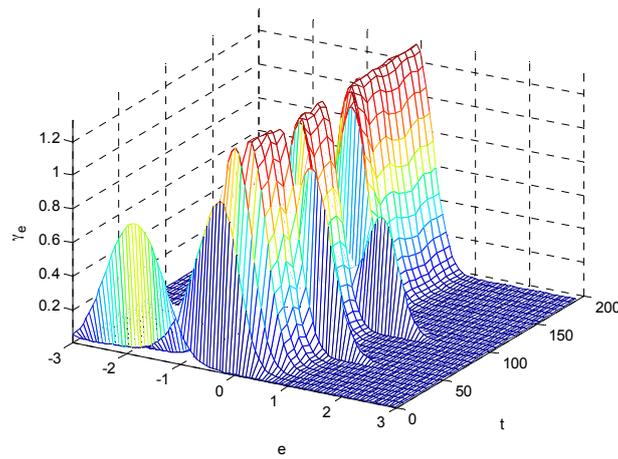


Figure 6. 3D mesh PDF using the proposed SIP based controller.

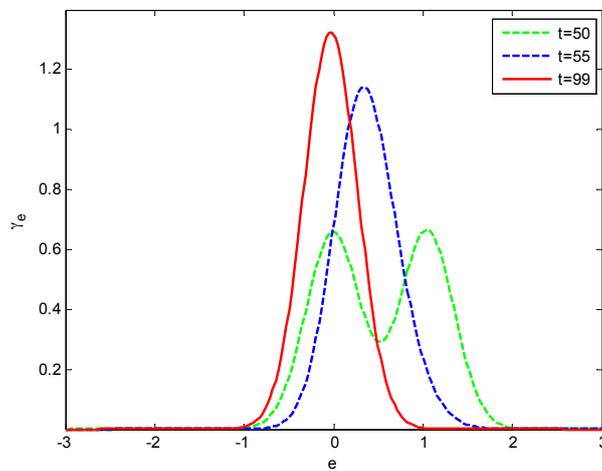


Figure 7. Selected tracking error PDFs using the SIP based controller.

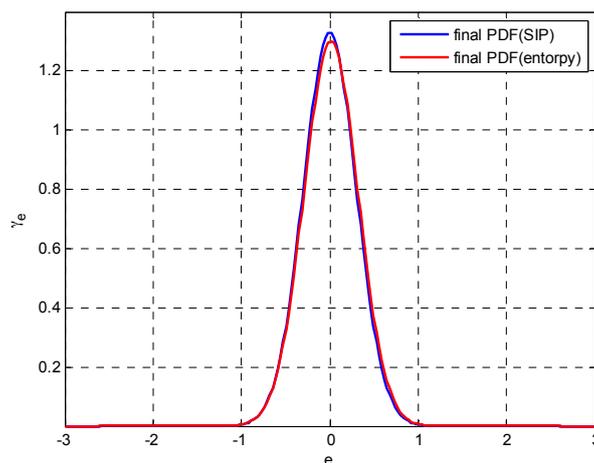


Figure 8. Final PDFs of the tracking error using the SIP based and entropy based controllers.

5. Conclusions

In the past, several different approaches, such as PDF shaping control, minimum entropy control, generalized minimum entropy control, *etc.* have been proposed to solve the control problem for

non-Gaussian stochastic systems. These control strategies can achieve good performance, but there are still two main issues to be further improved: (1) The entropy value can be negative and it is shift-invariant. More suitable statistical information that describes objective functions is necessary; (2) Control input in a stochastic system is also a random variable. The randomness of control input should be considered.

In this work, a convergent SNSPC algorithm is presented for the controlled system with non-Gaussian disturbance. The proposed SNSPC is obtained by minimizing a SIP-based predictive criterion, in which the randomness of the control input is also considered besides randomness of the tracking error. Compared with the entropy or IP, the randomness measure SIP has some advantages, such as validity in a wide range of distributions, robustness, and the simplicity in computation. Moreover, the multistep predictive control strategy, rather than single step control is developed in this paper as it is more robust to disturbances and nonlinearities involved in the systems. Also, the convergent condition of the proposed SNSPC based on the energy conservation principle is proposed. The proposed control strategy is applied in a nonlinear and non-Gaussian stochastic numerical example. The simulation results confirm that this new SIP based predictive control method can achieve a good tracking performance.

Compared with the previous work in the field of stochastic distribution control, the contributions of this paper are three folds: (1) the randomness of control inputs is considered for the first time; (2) instead of the instantaneous performance index, a novel SIP-based cumulative criterion is formulated; (3) a single neuron multi-step predictive control algorithm is obtained, and it is much better than the single-step control method. However, most of the practical industrial processes have multi-inputs and multi-outputs, and there are also many constraints when the controller is designed. Future research should be focused on such problems.

Acknowledgments: This work was supported by the Natural Science Foundation of China (61503271, 61374052) and Shanxi (2014011018-2, 20140210022-7). They are gratefully acknowledged.

Author Contributions: Mifeng Ren, Xinying Xu and Lan Cheng conceived the project. Mifeng Ren, Junghui Chen and Xinying Xu searched relevant literatures. Mifeng Ren and Junghui Chen carried out the theoretical derivation. Ting Cheng and Lan Cheng provided the simulation results. Mifeng Ren, Ting Cheng and Junghui Chen analysed the simulation results and wrote the paper. Correspondence and requests for materials should be addressed to Junghui Chen. All authors have read and approved the final manuscript.

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

References

1. Åström, K.J. *Introduction to Stochastic Control Theory*; Academic Press: New York, NY, USA, 1970.
2. Wang, H. *Bounded Dynamic Stochastic Systems: Modeling and Control*; Springer: London, UK, 2000.
3. Wang, H. Robust Control of the Output Probability Density Functions for Multivariable Stochastic Systems with Guaranteed Stability. *IEEE Trans. Autom. Control* **1999**, *44*, 2103–2107. [[CrossRef](#)]
4. Wang, H.; Zhang, J. Bounded Stochastic Distributions Control for Pseudo-ARMAX Stochastic Systems. *IEEE Trans. Autom. Control* **2001**, *46*, 486–490. [[CrossRef](#)]
5. Wang, H.; Yue, H. A rational spline model approximation and control of output probability density functions for dynamic stochastic systems. *Trans. Inst. Meas. Control* **2003**, *25*, 93–105. [[CrossRef](#)]
6. Guo, L.; Wang, H. *Stochastic Distribution Control System Design: A Convex Optimization Approach*; Springer: Berlin/Heidelberg, Germany, 2010.
7. Guo, L.; Wang, H. PID Controller Design for Output PDFs of Stochastic Systems Using Linear Matrix Inequalities. *IEEE Trans. Syst. Man Cybern. Part B Cybern.* **2005**, *35*, 65–71. [[CrossRef](#)]
8. Yi, Y.; Li, T.; Guo, L.; Wang, H. Statistic tracking strategy for non-Gaussian systems based on PID controller structure and LMI approach. *Dyn. Contin. Discret. Impuls. Syst. B* **2008**, *15*, 859–872.
9. Yue, H.; Wang, H. Minimum entropy control of closed-loop tracking errors for dynamic stochastic systems. *IEEE Trans. Autom. Control* **2003**, *48*, 118–122.
10. Zhang, J.; Ren, M.; Wang, H. Minimum entropy control for non-linear and non-Gaussian two-input and two-output dynamic stochastic systems. *IET Control Theory Appl.* **2012**, *6*, 2434–2441. [[CrossRef](#)]

11. Zhang, J.; Wang, H. Iterative learning-based minimum tracking error entropy controller for robotic manipulators with random communication time delays. *IET Control Theory Appl.* **2008**, *2*, 682–692. [[CrossRef](#)]
12. Afshar, P.; Wang, H.; Chai, T. An ILC-based adaptive control for general stochastic systems with strictly decreasing entropy. *IEEE Trans. Neural Netw.* **2009**, *20*, 471–482. [[CrossRef](#)] [[PubMed](#)]
13. Ren, M.; Zhang, J.; Wang, H. Minimized Tracking Error Randomness Control for Nonlinear Multivariate and Non-Gaussian Systems Using the Generalized Density Evolution Equation. *IEEE Trans. Autom. Control* **2014**, *9*, 2486–2490. [[CrossRef](#)]
14. Chen, B.; Zhu, Y.; Hu, J.; Principe, J.C. *System Parameter Identification: Information Criteria and Algorithms*; Elsevier: Amsterdam, The Netherlands, 2013.
15. Ren, M.; Zhang, J.; Jiang, M.; Yu, M.; Xu, J. Minimum (h, ϕ) - Entropy Control for Non-Gaussian Stochastic Networked Control Systems and Its Application to a Networked DC Motor Control System. *IEEE Trans. Control Syst. Technol.* **2015**, *23*, 406–411. [[CrossRef](#)]
16. Zhang, J.; Jiang, M.; Ren, M.; Hou, G.; Xu, J. Improved single neuron controller for multivariable stochastic systems with non-gaussianities and unmodeled dynamics. *ISA Trans.* **2013**, *52*, 752–758. [[CrossRef](#)] [[PubMed](#)]
17. Rao, M.; Chen, Y.; Vemuri, B.C.; Wang, F. Cumulative residual entropy: a new measure of information. *IEEE Trans. Inf. Theory* **2004**, *50*, 1220–1228. [[CrossRef](#)]
18. Zografos, K.; Nadarajah, S. Survival exponential entropies. *IEEE Trans. Inf. Theory* **2005**, *51*, 1239–1246. [[CrossRef](#)]
19. Chen, B.; Zhu, P.; Principe, J.C. Survival information potential: A new criterion for adaptive system training. *IEEE Trans. Signal Process.* **2012**, *60*, 1184–1194. [[CrossRef](#)]
20. Marcello, F.; Giulioni, L.; Magni, L.; Scattolini, R. An approach to output-feedback MPC of stochastic linear discrete-time systems. *Automatica* **2015**, *55*, 140–149.
21. Blackmore, L.; Ono, M.; Bektassov, A.; Williams, B.C. A probabilistic particle-control approximation of chance-constrained stochastic predictive control. *IEEE Trans. Robot.* **2010**, *26*, 502–517. [[CrossRef](#)]
22. Calafiore, G.C.; Fagiano, L. Robust model predictive control via scenario optimization. *IEEE Trans. Autom. Control* **2013**, *58*, 219–224. [[CrossRef](#)]
23. Zhou, Z.; Cogill, R. Reliable approximations of probability-constrained stochastic linear-quadratic control. *Automatica* **2013**, *49*, 2435–2439. [[CrossRef](#)]
24. Bernardini, D.; Bemporad, A. Stabilizing model predictive control of stochastic constrained linear systems. *IEEE Trans. Autom. Control* **2012**, *57*, 1468–1480. [[CrossRef](#)]
25. Korda, M.; Gondhalekar, R.; Oldewurtel, F.; Jones, C.N. Stochastic MPC framework for controlling the average constraint violation. *IEEE Trans. Autom. Control* **2014**, *59*, 1706–1721. [[CrossRef](#)]
26. Cannon, M.; Cheng, Q.; Kouvaritakis, B.; Raković, S.V. Stochastic tube MPC with state estimation. *Automatica* **2012**, *48*, 536–541. [[CrossRef](#)]
27. Kouvaritakis, B.; Cannon, M.; Rakovic, S.V.; Cheng, Q. Explicit use of probabilistic distributions in linear predictive control. *Automatica* **2010**, *46*, 1719–1724. [[CrossRef](#)]
28. Cannon, M.; Kouvaritakis, B.; Wu, X. Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints. *Automatica* **2009**, *45*, 167–172. [[CrossRef](#)]
29. Ali, M. Stochastic Model Predictive Control: An Overview and Perspectives for Future Research. *IEEE Control Syst. Mag.* **2016**. in press.
30. Edward, A.B.; Joel, A.P.; Ali, A.; Ali, M. Lyapunov-Based Stochastic Nonlinear Model Predictive Control: Shaping the State Probability Distribution Functions. In Proceedings of the 2016 American Control Conference, Boston, MA, USA, 6–8 July 2016.
31. Mesbah, A.; Streif, S.; Findeisen, R.; Braatz, R.D. Stochastic Nonlinear Model Predictive Control with Probabilistic Constraints. In Proceedings of the 2014 American Control Conference (ACC), Portland, OR, USA, 4–6 June 2014; pp. 2413–2419.
32. Zhang, J.; Yue, H.; Zhou, J. A new pdf modelling algorithm and predictive controller design. In Proceedings of the 10th IFAC International Symposium on Dynamics and Control of Process Systems, Mumbai, India, 18–20 December 2013; pp. 271–276.
33. Zhang, J.; Yue, H.; Zhou, J. Predictive PDF control in shaping of molecular weight distribution based on a new modeling algorithm. *J. Process Control* **2015**, *30*, 80–89. [[CrossRef](#)]
34. Ren, M.; Zhang, J.; Jiang, M.; Tian, Y.; Hou, G. Statistical Information Based Single Neuron Adaptive Control for Non-Gaussian Stochastic Systems. *Entropy* **2012**, *14*, 1154–1164. [[CrossRef](#)]

35. Zan, X.; Xie, F. Switched Reluctance Generator System Based on Single Neuron adaptive PID Controller. In Proceedings of the 2011 International Conference on Advanced Mechatronic Systems, Zhengzhou, China, 11–13 August 2011; pp. 123–127.
36. Xia, Y.; Xie, W.; Liu, B.; Wang, X. Data-driven predictive control for networked control systems. *Inf. Sci.* **2013**, *235*, 45–54. [[CrossRef](#)]
37. Chen, B.; Zhu, Y.; Hu, J. Mean-Square Convergence Analysis of ADALINE Training With Minimum Error Entropy Criterion. *IEEE Trans. Neural Netw.* **2010**, *21*, 1168–1179. [[CrossRef](#)] [[PubMed](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).