

УДК 519.87

## Identification of Multidimensional Technological Processes with Dependent Input Variables

Alexander V. Medvedev\*

Siberian State University of science and technologies Reshetnev  
Krasnoyarsky Rabochy, 31, Krasnoyarsk, 660037  
Russia

Eugene D. Mihov†

Nikolay D. Ivanov‡

Siberian Federal University  
Svobodny, 79, Krasnoyarsk, 660041  
Russia

Received 26.05.2017, received in revised form 13.06.2018, accepted 26.07.2018

*In the article the problem of identification of multidimensional technological processes is studied. The question of modelling H-processes with several output variables is considered. A method of analysis of the training sample that determines the possibility of modelling an H-process with several output variables is proposed.*

*Keywords: apriori information, identification, nonparametric model, nonparametric algorithms, H-model, fractional dimension space.*

DOI: 10.17516/1997-1397-2018-11-5-649-658.

## Introduction

Identification of stochastic objects is often reduced to identification of static systems with delay. This is due to the fact that some output variables of the object are controlled by much larger intervals than the input variables. For example, some variables are measured electronically (in this case, discrete control can be sufficiently small), while the other variables are controlled by chemical analysis or physical-mechanical tests (in this case the discrete control  $\Delta T$  is large, i.e.  $\Delta T \gg \Delta t$ ).

The most common scheme a discrete-continuous process under study can be represented as follows on Fig. 1.

The following notation is used on Fig. 1:  $A$  is the investigated object (the process);  $\overrightarrow{x(t)}$ ,  $\overrightarrow{q(t)}$ , and  $\overrightarrow{z(t)}$  are the output vectors of the process;  $\overrightarrow{u(t)}$  is the vector of control actions;  $\overrightarrow{\mu(t)}$  is an uncontrolled but measured input vector of the process;  $\overrightarrow{\lambda(t)}$  is the input vector of uncontrolled and not measured process variables;  $\overrightarrow{\xi(t)}$  is the casual influence;  $\overrightarrow{\omega^i(t)}$  :  $i = \overline{1, k}$  are the process variables controlled in the object;  $t$  is continuous time  $H^\mu, H^u, H^x, H^z, H^q, H^\omega$  are the communication channels corresponding to various variables, including control devices, devices for measurement of observed variables;  $\mu_t, u_t, x_t, \omega_t$  are measurement of  $\overrightarrow{\mu(t)}, \overrightarrow{u(t)}, \overrightarrow{x(t)}, \overrightarrow{\omega(t)}$  in discrete time;  $h^\mu(t), h^u(t), h^x(t); h^{\omega_1}(t) \dots h^{\omega_k}(t)$  are measurement errors of the corresponding variables of the process.

\*Saor\_medvedev@sibsau.ru

†edmihov@mail.ru

‡blackhawk009@yandex.ru

© Siberian Federal University. All rights reserved

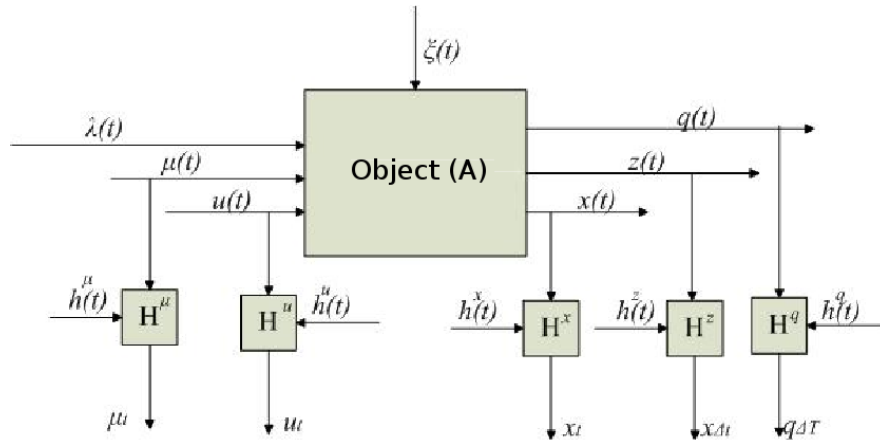


Fig. 1. The general scheme of a process

It should be noted that processes related to cybernetics and the automatic control theory, rather than processes from probability theory (like Wiener or Gauss processes) [1] are considered here.

## 1. Levels of a priori information

There are systems with different levels of a priori information [2]:

- systems with parametric uncertainty. The parametric level of a priori information assumes presence of a parametric structure of the model and some characteristics of random interference such as zero mathematical expectation and bounded variance. Iterative probabilistic procedures are commonly used to evaluate parameters.
- systems with nonparametric uncertainty. This level does not assume presence of a model, however, some knowledge about qualitative characteristics of the process such as ambiguity or unambiguity of its characteristics, linearity for dynamic processes or the nature of their nonlinearity is required. Methods of nonparametric statistics are used to solve problems of identification.
- systems with parametric and nonparametric uncertainty. The task of identifying multiply connected systems, when the amount of initial information does not correspond to the types described above, is important in practice. The problem of identification is formulated under the conditions of both parametric and nonparametric a priori information. Models represent an interconnected system of parametric and nonparametric relations in this case.

## 2. General description of $H$ -processes

The situation when the input actions are stochastically dependent often appears in the modelling and control of inertia-free objects with delay. The nature of dependence between the input variables is often unknown. The described situation is typical for discrete-continuous processes dominating in metallurgy, coal industry, construction industry, etc.

The case when the input variables have a stochastic dependence is the special case, which may be classified as a new class of identification tasks. Such processes are called  $H^\lambda$ -processes [3]. In other words, an  $H^\lambda$  is a process with input variables stochastically dependent.

The space of input and output variables of the process is the domain of the process. When input variables have stochastic dependence, the process is localized in a subdomain. Any process with proportional input variables is an  $H$ -process: steel smelting (dependent variables are carbon and iron), dough kneading (dependent variables are flour and water).

In three-dimensional space an  $H^\lambda$ -process may look like a tube, a line or a part of a plane. In the simplest case an  $H^\lambda$ -process is a line (Fig. 2).

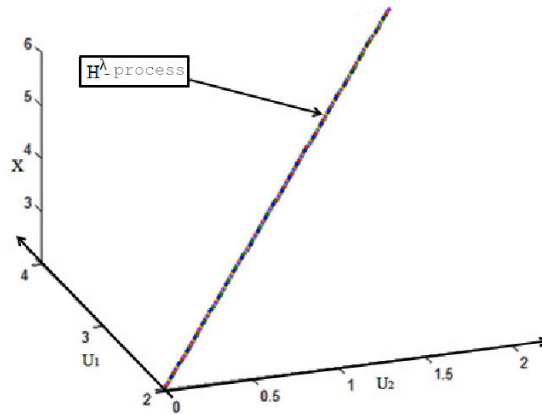


Fig. 2. Example of an  $H^\lambda$ -process. The case of linear dependence between input variables

In this example the relation between  $u_2$  and  $u_1$  is linear; noise does not affect this process.

In real production noise is always present, but the linear dependence between input variables may exist. Of course, complex processes may have a non-linear dependence between input variables (Fig. 3).

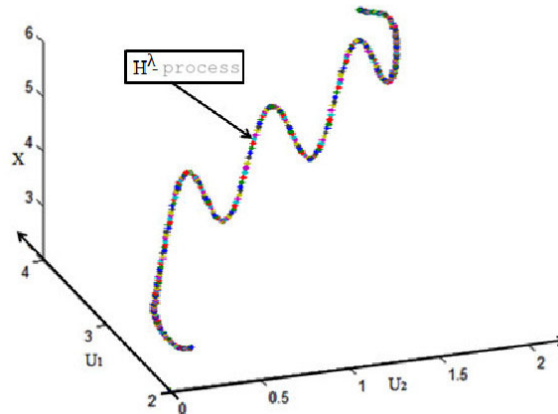


Fig. 3. Example of an  $H^\lambda$ -process. The case of non-linear dependence between input variables

Such processes are more complicated than the processes depicted in Fig. 2. The researcher must know the non-linear dependence between variables for successful modelling of the process. Similar processes are more often encountered in industry than those considered earlier.

Nethertheless, this process is represented by a curve. This means that if the input variables are given even with a slight deviation from the dependency, the process does not exist.

An  $H^\lambda$ -process close to a real process is shown in Fig. 4. The faceted structure indicates the domain of existence of an  $H$ -process.

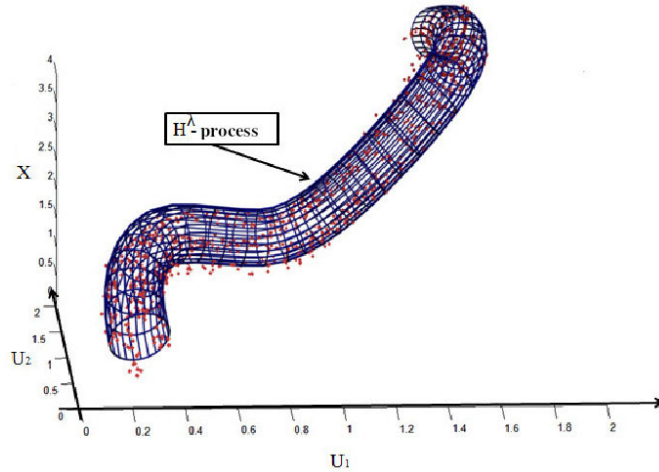


Fig. 4. Example of an  $H^\lambda$ -process close to a real process

### 3. Indicator function

If we try to recover a parametric model of an  $H^\lambda$ -process without taking into account its tubular form, then only points of  $\Omega_\lambda^H(u, x)$  will be used. In particular, even in the simplest case there are infinitely many models of a line in the for of a plane (see Fig. 5).

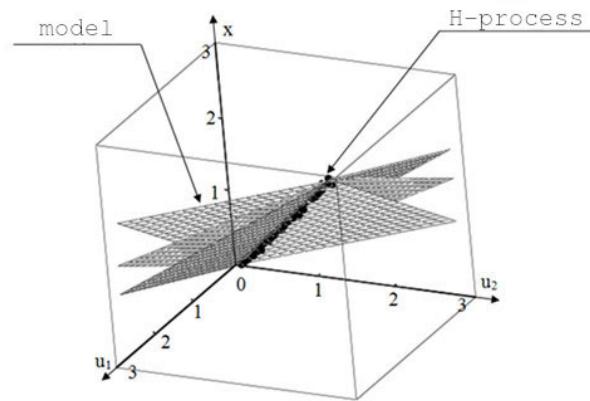


Fig. 5. Parametric models of an  $H$ -process

In the case of a non-linear dependence a recovered surface may also be not unique. To rectify the situation we consider the indicator function. Using the indicator function we transform the standard model

$$\hat{x}_s = \hat{f}(u, \alpha_s), \tag{1}$$

into

$$\hat{x}_s = \hat{f}(u, \alpha_s)I(u). \tag{2}$$

The indicator function  $I(u)$  is defined as follows

$$I(\vec{u}) = \begin{cases} 1, & u \in \Omega_\lambda^H(\vec{u}), \\ 0, & u \notin \Omega_\lambda^H(\vec{u}). \end{cases} \quad (3)$$

Note that the domain  $\Omega^H(\vec{u})$  is not known, only the sample  $x_i, u_i, i = \overline{1, s}$  is given. The estimate  $\hat{x}_s$  can not be calculated if the indicator function is zero, i.e. for such values of the vector components  $\vec{u} \in \Omega(\vec{u})$  the process does not exist. If the indicator function  $I(\vec{u})$  is equal to one for any value,  $\vec{u} \in \Omega(u)$ , then the model converts to the standard one.

As an estimate of the indicator function  $I_s(u)$  we can take

$$I_s(u) = \text{sgn} \sum_{i=1}^s \prod_{j=1}^k \Phi(c_s^{-1}(u^j - u_i^j)), \quad (4)$$

where  $\Phi$  is a core function.

A core function  $\Phi(*)$  is a function satisfying the following conditions

$$\frac{1}{c_s} \int_{-\infty}^{+\infty} \Phi\left(\frac{u - u_i}{c_s}\right) du = 1, \quad (5)$$

$$\lim_{c_s \rightarrow 0} \frac{1}{c_s} \int_{-\infty}^{+\infty} \varphi(u) \Phi\left(\frac{u - u_i}{c_s}\right) dt = \varphi(u_i). \quad (6)$$

The triangular core

$$\Phi(*) = \begin{cases} (1 - |*|), & |*| \leq 1, \\ 0, & |*| > 1 \end{cases} \quad (7)$$

is used for the indicator function estimate.

### $H^\lambda$ -processes with several output variables

$H^\lambda$ -processes described above have one output variable. Certainly, this is unrealistic, as a rule real technological processes are characterized by sets of output variables  $(x_j, j = \overline{1, m})$ , where  $m$  is the number of output variables. It is logical to assume that some of the output variables vary not in  $\Omega_{j\lambda}(\vec{u})$  but in its subdomain  $\Omega_{j\lambda}^H(\vec{u})$ .

An example of such process with two input and two output variables is shown in Fig. 6.

Existence of such processes implies appearance of qualitatively new tasks in the field of modeling and control.

### $H^\lambda$ -processes with intersecting outputs

Consider a special case of  $H^\lambda$  processes with several output variables. Assume that it has two input and two output variables; the training sample has size  $n: (x_{1i}, x_{2i}, \vec{u}_i), i = \overline{1, n}$ . The structure of the process is shown in Fig. 7.

Here we use the notation:  $\Omega_\lambda(\vec{u})$  is the domain of input variables;  $\Omega_\lambda^{H1}(\vec{u})$  is the domain of input variables for the output variable  $x_1$ ;  $\Omega_\lambda^{H2}(\vec{u})$  is the domain of input variables for the output variable  $x_2$ ;  $\Omega_\lambda^{H12}(\vec{u})$  is the intersection of  $\Omega_\lambda^{H1}(\vec{u})$  and  $\Omega_\lambda^{H2}(\vec{u})$ ;  $u_1, u_2$  are the input variables;  $x_1, x_2$  are the output variables.

The output variable  $x_1$  is defined only in  $\Omega_\lambda^{H1}(\vec{u})$ . Outside this domain a subprocess giving  $x_1$  does not exist. The same holds for the output variable  $x_2$  and the domain  $\Omega_\lambda^{H2}(\vec{u})$ . It follows that a process with two output variables  $x_1$  and  $x_2$  exists only in  $\Omega_\lambda^{H12}(\vec{u})$ .

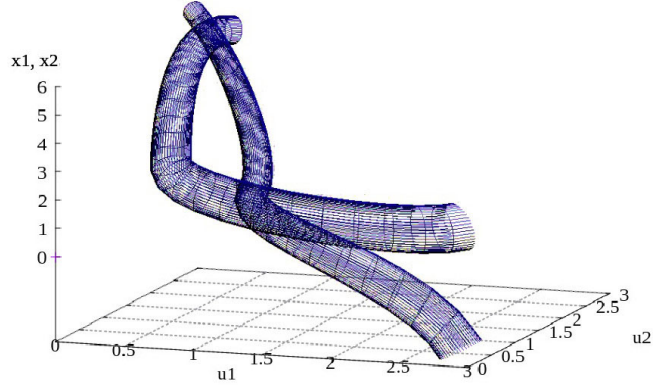


Fig. 6. Example  $H^\lambda$ -processes with several output variables

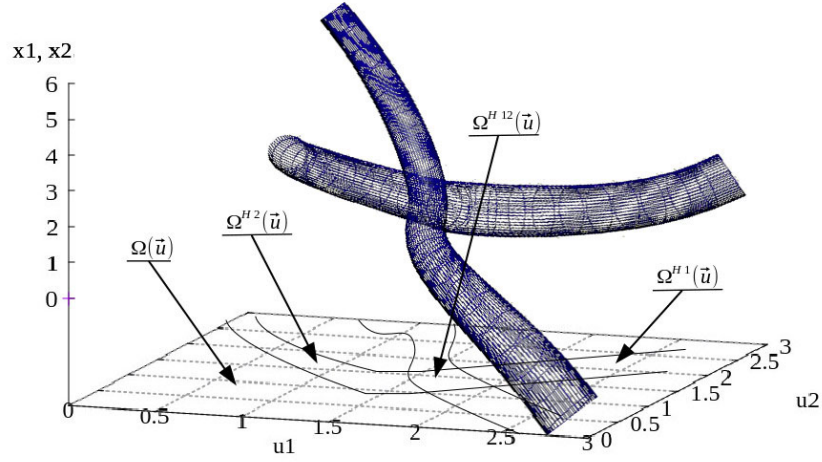


Fig. 7. A process with two outputs having common domain

A researcher does not know about existence of  $\Omega_\lambda^{H^1}(\vec{u})$ ,  $\Omega_\lambda^{H^2}(\vec{u})$ , and  $\Omega_\lambda^{H^{12}}(\vec{u})$ . This can adversely affect the accuracy of the design model (it may forecast even in the case when the process does not exist, which means that the forecasts are incorrect). A logical solution in this case is to use two indicators for each output variable.

With indicators the model looks as follows:

$$\begin{aligned} x_1 &= I_1(\vec{u}_i)I_2(\vec{u}_i)\hat{f}_1(u_i, \alpha_s), \\ x_2 &= I_1(\vec{u}_i)I_2(\vec{u}_i)\hat{f}_2(u_i, \alpha_s), \end{aligned} \tag{8}$$

where  $I_j(\vec{u})$  is the indicator function that determines if the the vector of input variables  $\vec{u}$  belongs to the domain  $\Omega_\lambda^{H^j}(\vec{u})$ ,  $x_1$  and  $x_2$  are the output variables of the process.

The indicator function s have the form

$$\begin{aligned} I_1(\vec{u}) &= \begin{cases} 1, & \vec{u} \in \Omega^{H^1}(\vec{u}), \\ 0, & \vec{u} \notin \Omega^{H^1}(\vec{u}), \end{cases} \\ I_2(\vec{u}) &= \begin{cases} 1, & \vec{u} \in \Omega^{H^2}(\vec{u}), \\ 0, & \vec{u} \notin \Omega^{H^2}(\vec{u}). \end{cases} \end{aligned} \tag{9}$$

In general, the model of an  $H$ -process with  $k$  output variables looks as follows:

$$\begin{aligned}
 x_1 &= \hat{f}_1(u_i, \alpha_s) \prod_{v=0}^k I_v(\vec{u}_i), \\
 &\dots\dots\dots \\
 x_k &= \hat{f}_k(u_i, \alpha_s) \prod_{v=0}^k I_v(\vec{u}_i).
 \end{aligned}
 \tag{10}$$

If one of the output variables  $x$  is defined for any value of the input variables from  $\Omega(\vec{u})$ , then the indicator function is identically 1 and does not affect the expression  $\prod_{v=0}^k I_v(\vec{u})$ .

If a process is not an  $H$ -process, then all indicator functions are identically 1, and  $\prod_{v=0}^k I_v(\vec{u}) = 1$ . In this case the model becomes

$$\begin{aligned}
 x_1 &= \hat{f}_1(u_i, \alpha_s), \\
 &\dots\dots\dots \\
 x_k &= \hat{f}_k(u_i, \alpha_s),
 \end{aligned}
 \tag{11}$$

which proves that model (6) is a more general form of a process model with several output variables.

Let us consider an example of a process model with indicator functions. Take a simple system of equations:

$$\begin{aligned}
 x_1 &= u_1 + u_2, \\
 x_2 &= u_1 + 0.6u_2,
 \end{aligned}
 \tag{12}$$

where  $\vec{u} \in (\overline{0, 3})$ ; the noise that affects the process is 5%. This is an  $H^\lambda$  process. For the first output variable we have the following model:

$$u_2 = \sin(u_1/1.3) * 2 + \xi_1, \quad \xi_1 \in (-0.2; 0.2).
 \tag{13}$$

For the second output variable we have

$$u_2 = 3 - u_1 + \xi_2, \quad \xi_2 \in (-0.2; 0.2).
 \tag{14}$$

The training sample is given in Fig. 8.

Fig. 8 demonstrates that some part of the training sample does not belong to  $\Omega_\lambda^{H12}$ . We process the elements of the training sample by multiplying them by the indicator functions.

Fig. 9 demonstrates that only those elements that belong to  $\Omega_\lambda^{H12}$  remain after applying the indicator functions. To find the intersection of  $\Omega_\lambda^{H1}$  and  $\Omega_\lambda^{H2}$  a sliding test can be employed. The rule for determining whether an element of the training sample belongs to the intersection is:

$$M(\vec{u}_i) = \prod_{v=0}^k I_v(\vec{u}), \quad i = \overline{1, n}.$$

The function  $M(\vec{u}_i)$  can assume one of the following values:

$$M(\vec{u}) = \begin{cases} 1, & \vec{u} \in \Omega_\lambda^{H12}, \\ 0, & \vec{u} \notin \Omega_\lambda^{H12}. \end{cases}
 \tag{15}$$

This analysis allows to judge about characteristics of this intersection, as well as whether it is empty or not.

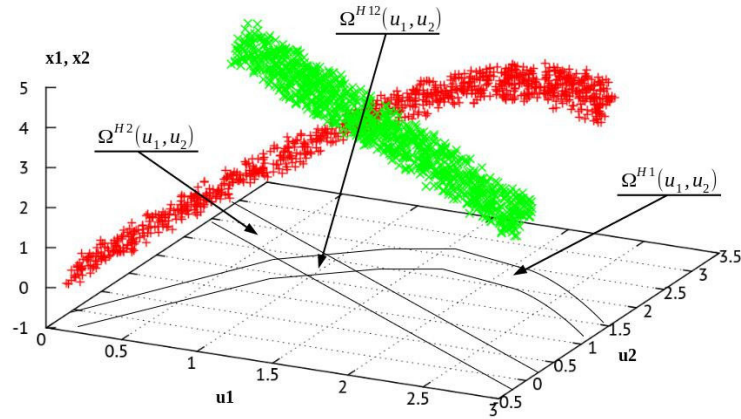


Fig. 8. The training sample before processing

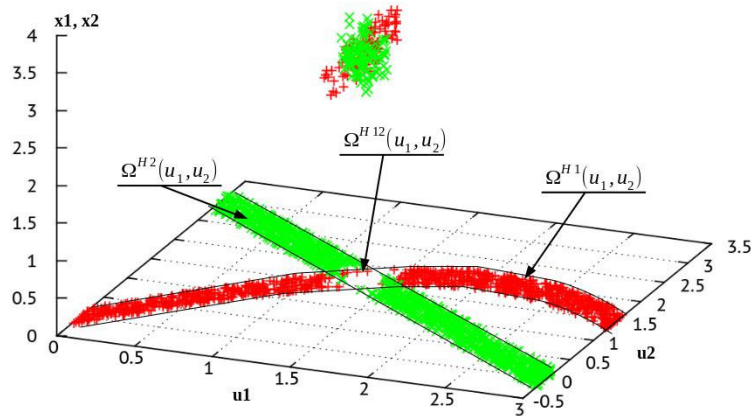


Fig. 9. The training sample after processing

### $H^\lambda$ -processes without intersecting outputs

If  $M(\vec{u}_i) = 0, i = \overline{1, n}$ , the output variables of the process under investigation do not have a common domain. This case is depicted on Fig. 10.

This means that the process cannot exist with such set of output variables, for example if they exclude each other. Such a situation is typical for active processes (processes involving people), organization processes, rarely for technological processes. In this case an expert has to change the set of output variables so that they have non empty intersection domain. It should be noted that this clearly implies an action of an expert.

The following algorithm can be used to find sets of output variables with intersection to present them to an expert: One checks the values of  $M(\vec{u}, \vec{x})$  of the process using only two output variables  $x_i$  and  $x_j, i = \overline{1, k-1}, j = \overline{i, k}$ , where  $k$  is the number of output variables. The pair of output variables for which  $M(\vec{u}, \vec{x}) = 1$  should be marked as having non empty intersection; thus, we find  $\vec{x}_p, p = \overline{1, q}$ , where  $q$  is the number of found pairs. To each found pair we add one of the remaining variables and proceed as above to find appropriate triples of output variables. This operation is repeated until addition of a new variable leads to  $M(\vec{u}, \vec{x}) = 0$ .



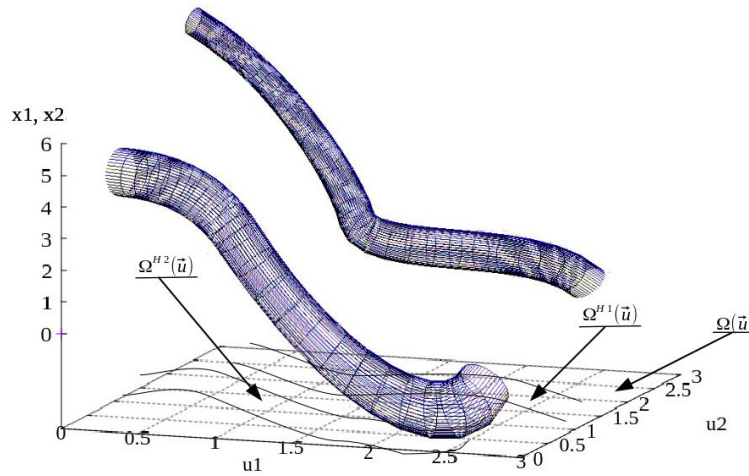


Fig. 10. An  $H$ -process without intersecting outputs

The expert should choose one set of output variables from the obtained ones.

The proposed algorithm is a ‘greedy’ one, but it does not miss the global minimum in this problem. The condition  $M(\vec{u}, \vec{x}_i) = 0$ ,  $M(\vec{u}, \vec{x}_j) = 1$  must be met to skip the global minimum, where  $\vec{x}_j$  is a vector of output variables that includes all output variables from  $\vec{x}_i$  and one more output variable. This situation is excluded, because if  $M(\vec{u}_i) = \prod_{v=0}^{q_i} I_v(\vec{u}) = 0$ ,  $M(\vec{u}_j)$  will also be equal to 0.

The described case is a new one for identification tasks. It should be added also that a lot of active processes fall into the situation shown above, therefore the study of this particular situation is very relevant.

## Conclusion

The problem of modelling  $H$ -processes was considered. A method of modelling  $H$ -processes based on the modification of the parametric model using the indicator function is proposed.

A new kind of  $H$ -processes, namely multidimensional  $H$ -processes, is considered. The case when the intersection of domains of variables is not empty is described, and a method of modelling such processes is proposed.

Also the case when the intersection of domains of variables is empty is considered. A method for selecting a new set of output parameters that has non empty intersection is proposed.

*The research was supported by a grant from the Russian Science Foundation (project no. 16-19-10089).*

## References

- [1] J.L. Doob, Stochastic Processes, Wiley Classics Library, 1990.
- [2] N.G.Zagoruiko, Cognitive data analysis, Geo. Novosibirsk, 2013 (in Russian).
- [3] A.V.Medvedev, Analysis of the data in the problem identification, *Computer analysis of modelling data. Minsk, BSU*, **2**(1995), 201–206 (in Russian).

## Идентификация многомерных технологических процессов с зависимыми входными переменными

**Александр В. Медведев**

Сибирский государственный университет науки и технологий им. М. Ф. Решетнева  
Красноярский рабочий, 31, Красноярск, 660037  
Россия

**Евгений Д. Михов**

**Николай Д. Иванов**

Сибирский федеральный университет  
Киренского, 26, Красноярск, 660041  
Россия

---

*В статье рассматривается задача идентификации многомерных технологических процессов. Изучен вопрос моделирования  $H$ -процессов с несколькими выходными переменными. Предложен метод анализа обучающей выборки, при помощи которого определяется, возможно ли создать модель исследуемого  $H$ -процесса с несколькими выходными переменными.*

*Ключевые слова: априорная информация, идентификация, непараметрическая модель, непараметрические алгоритмы,  $H$ -модель, пространство дробной размерности.*