Diagnostic Goal-driven Reduction of Multiscale Process Models

Erzsébet Németh, Rozália Lakner, Katalin M. Hangos

Research Report SCL-001/2005

Diagnostic Goal-driven Reduction of Multiscale Process Models

Erzsébet Németh[1] Rozália Lakner[2]
E-mail: nemethe@scl.sztaki.hu E-mail: lakner@almos.vein.hu

Katalin M. Hangos[1]
E-mail: hangos@scl.sztaki.hu

[1]Process Control Research Group, Systems and Control Research Laboratory,
Computer and Automation Institute HAS
H-1518 Budapest, P.O. Box 63, Hungary

[2]Department of Computer Science, University of Veszprém
H-8201 Veszprém, P.O. Box 158, Hungary

Abstract

A systematic model reduction method is proposed in this paper that is applicable for multiscale process models with steady-state operating points for prediction-based fault detection, isolation and loss prevention in the initial phase of a fault. Based on the time-scale separation of the dynamics a refined scale-map, a model structure map is constructed for each root cause and symptom and the relevant target time scale is determined. Then the scale-reduced nonlinear model of the system is constructed that only contains the dynamic balance equations on the target time scale with all the other dynamic equations reduced by steady-state assumptions. Thereafter local linearization is applied to obtain the final reduced model.

The resulted set of reduced models preserves the engineering meaning of the remaining state variables, while the number of state variables and parameters is decreased significantly.

Contents

List of Figures

List of Tables

1 Introduction

Fault detection and diagnosis in large-scale process systems is a of great practical importance and present various challenging research problems at the same time. One of them is the computational complexity of the algorithms, that causes an exponential growth of the computational resources (time and memory) with increasing system sizes [Venkatasubramanian *et al.* (2003)]. One remedy of this problem is to decompose the system model and effectively focus on its relevant sub-model when doing the fault detection, isolation and loss prevention.

The multi-scale approach [Ingram *et al.* (2004)] to process modelling is an effective way both of decomposing and handling the available information and of performing the on-line fault detection, diagnosis and loss prevention in these large complex systems. The multi-scale approach provides us with a natural mechanism-driven hierarchical decomposition of the underlying process model with any related information and an integration framework to organize the information exchange between the partial models.

Lumped process models are the most important and widespread class of process models for control and diagnostic applications. Therefore we also restrict ourselves to this case.

The most effective way of focusing to a part of a dynamic system relevant to our purposes is to apply model reduction or model simplification techniques. Many of such is reported in the literature both for linear and nonlinear models. There is a well-known model reduction technique for linear time-invariant state-space models that is based on controllability and observability indices, called gramians and on linear state transformations to construct a balanced realization. This method is purely black-box in its nature because the physical meaning of the state variables in the reduced model is completely lost. [Hahn *et al.* (2003)] generalize this method for stable nonlinear systems for nonlinear model-based predictive control purposes.

Model structure simplification methods [Leitold *et al.* (2002)] offer a grey-box alternative to model reduction where the number of state variables is reduced using steady-state and/or variable lumping transformations and the physical meaning of the remaining state variables remains unchanged. The model reduction approach proposed in this paper can be regarded as an extension to this method.

Almost all of the reported model reduction and simplification methods apply analytical or combinatorial techniques that are difficult to apply in a multiscale context where naturally both analytical and qualitative methods are used. Therefore, our aim was to propose a mixed method for model reduction of multiscale process models for diagnosis that effectively utilizes the model structure inherent in the multiscale nature of the model.

2 The Model Reduction Problem

The basic notions for formulating and solving the model reduction problem are briefly described in this section.

2.1 Process system models and their modelling goals

Process systems are a sub-class of systems that *obey the law of thermodynamics*. This implies that dynamic process models based on first engineering principles are constructed by using the first and second law of thermodynamics.

The construction of process models starts by formulating a *modelling problem statement* in the following general form.

Given : a process system together with a ***modelling goal***
that can be process design, process control or diagnosis to mention just a few,
Construct : a model ***for the modelling goal***.

Construction principle. Process models for control and diagnostic purposes are constructed from the dynamic conservation balances of mass, component masses and energy (see e.g. [Hangos and Cameron (2001)]). The terms in the conservation balances correspond to the various *mechanisms* that are taken into account:

- in-convection (inlet term),
- out-convection (outlet term),
- (interphase) transfer,
- sources (and sinks including chemical reactions)

Modelling goal and its effect on the model. It is important to emphasize that the modelling goal has a determining effect on the model to be constructed: it determines the model performance variables, the mechanisms to be taken into account and the desired accuracy of the model in terms of its performance variables, as well as the type (static – dynamic, deterministic – stochastic, lumped parameter – distributed parameter) of the model. There is no matured systematic way of how to take the effect of the modelling goal on the model into account, just a few early steps about the "goal-directed modelling" [Hangos *et al.* (2005)] is available.

2.2 The structure of process models

Lumped process models that are constructed on first engineering principles possess a well defined structure. The *model equations* form a set of DAEs, where the

- *conservation balances are the differential equations* that are equipped with
- constitutive equations being algebraic equations.

The structure of the model equations gives rise to the classification of the *variables of a state-space model* that has been derived from a process model based on first engineering principles.

- The *state variables* are the intensive pairs of conserved extensive quantities (mass, concentrations, temperature) [Hangos and Cameron (2001)] for which conservation balances are constructed, while
- the *inputs, outputs and disturbance* variables are problem formulation dependent.

2.3 Functionally equivalent process models

Two process models of the same process system are called functionally equivalent, if they fulfill the same modelling goal [Hangos *et al.* (2005)]. If the modelling goal is formulated in terms of performance indices, that are predicates defined on the performance variables, than the functional equivalence can be determined algorithmically.

Suitably defined size indices that give the size of a model in a generalized sense can be used to relate functionally equivalent models and to determine which one is of less size, i.e. "reduced" compared to the other. Figure 1 shows how the performance and size indices relate to each other.

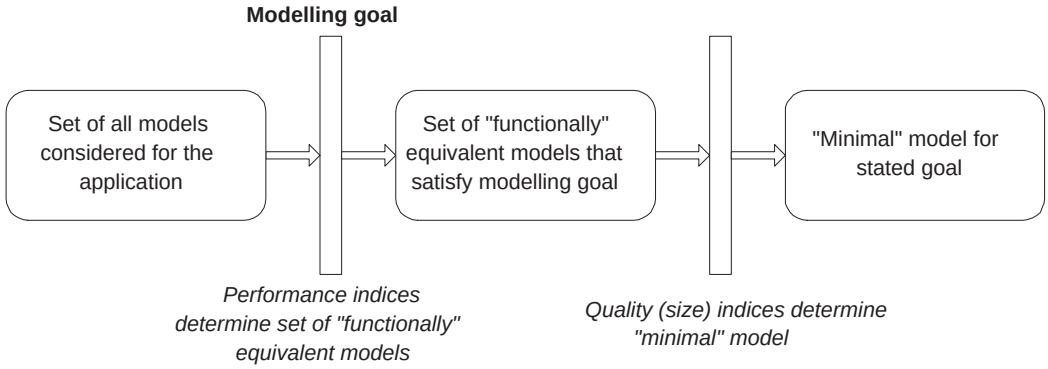

Figure 1: Functionally equivalent models and their reduction

2.4 The problem statement of model reduction

The aim of model reduction is to find a "simple" model that is functionally equivalent with a given detailed "original" one but it is more easy to handle computationally, i.e. it is of smaller size than the original one.

If one wants to obtain reduced models for diagnostic purposes, then the following assumptions can be made about the process system and its model:

$\mathcal{A}1$ Lumped dynamic process models are considered for both the original and the reduced model.

$\mathcal{A}2$ The process system has smooth nonlinearities and we want to describe its behaviour in the neighbourhood of a steady-state operating point.

Under the above assumptions it is possible to formulate the problem statement of model reduction to this special case as follows.

Given:

- a *detailed process model* that originates from first engineering principles
- a set of *prescribed input-output scenarios* that define the required performance of the models,
- a set of *simplicity indices* as generalized size indices being the number of state variables (dynamic conservation balances) and the linearity of the model

Compute: the *"most simple" reduced model* that is

- functionally equivalent with the detailed process model in terms of the prescribed input-output scenarios, and it is
- of the smallest size with respect to the simplicity indices.

The reduced model with the above two properties can be called *minimal model* [Hangos *et al.* (2005)] with respect to the specified simplicity indices. Note that minimal models are not necessarily unique.

2.5 Elementary model reduction steps

Under the assumptions $\mathcal{A}1$ and $\mathcal{A}2$ one can consider two basic elementary model reduction steps that are applicable to nonlinear state-space models: the reduction of the number of state equations by separating the different time-modes (fast, medium and slow modes, for example) and the linearization of the state equations around a steady-state point.

Reduction of the number of state equations. State equations originate from dynamic conservation balances in a process model derived from first engineering principles with the constitutive equations substituted. If balances for different conserved quantities, such as component masses and energy, or balances over balance volumes with largely different total holdup are present, then one observes the *separation of the time-constants*. There are also formal methods based on singular perturbation analysis (see e.g. [Kumar and Daoutidis (2002)], [Baldea and Daoutidis (2005)] and the references therein) to find the separated characteristic times and to obtain dynamic models from the original detailed model separately for each time level by model reduction.

If one considers the elementary steps of model reduction as model transformations, specifically projections acting on the set of lumped dynamic models of a process system, then the transformation applicable in the case of time-scale separation can be characterized by a ***steady-state assumption for a "fast" or "slowly changing" variable*** χ. Formally one applies the transformation

$$\frac{d\chi}{dt} \approx 0 \tag{1}$$

to a set of model equations. As a result of the transformation, the differential equation that originates from the dynamic conservation balance for χ becomes algebraic, and thus it should be substituted to the remaining differential equations. Thus the number of state equations (and the number of state variables) decreases by one, and χ formally disappears from the equations (see [Leitold *et al.* (2002)] for the details).

Linearization. Linearization around a steady-state point is a basic and well-known model simplification (or reduction) operation [Hangos and Cameron (2001)]. The applicability condition is that the original model should have smooth nonlinearities and the linearized model will be valid only within a more-or-less narrow region of a nominal operating point.

One may apply linearization for the total model or for a part of the model only depending on the nonlinearities along the variable-coordinate directions.

Properties of the elementary model reduction steps. It can be shown that both the reduction of the number state equations and the linearization as model transformations possess the following basic properties:

- They *preserve the basic dynamic properties* (e.g. controllability, observability and stability) of the model [Leitold *et al.* (2002)].
- Their actual result is strongly *steady-state point dependent*.
- They are both applicable in the *"initial phase" of a fault* when the system is in the neighbourhood of a steady-state nominal operating point.

3 Prediction-Based Diagnosis and Loss Prevention

Prediction of a system's behaviour is used for deriving the consequences of a state of the system in time that is usually performed in process engineering by dynamic simulation. With the help of prediction, however, the faulty mode of the system can also be detected based on the comparison between the real plant data and the predicted values generated by a suitable dynamic model. This type of fault detection and diagnosis is called prediction-based diagnosis [Venkatasubramanian *et al.* (2003)]. Because of the high complexity of multiscale process models, however, the computational load of performing the prediction can be substantial. Therefore, the need of model reduction arises.

3.1 Modelling goal for diagnosis

If one intends to construct a process model for diagnostic purposes this modelling aim has important implications on the model and its variables as follows.

- The model should be dynamic and should be able to produce *dynamic input-output behaviour* with the measurable quantities as output variables.
- One usually defines *symptoms* from the measurable output signals, that are qualitative performance output signals of the model.
- The *actuator input* variables correspond to manipulable input variables, that can be used for preventing dangerous consequences of the considered faults.
- Root causes are usually considered as *disturbances* that may determine the "failure modes" of the system with possibly different process models (i.e. a hybrid model is often needed).

3.2 Elements of prediction-based diagnosis

The most important notions of prediction-based diagnosis are briefly summarized here.

Symptoms. Similarly to medical diagnosis, the diagnosis of process systems is usually based on symptoms. Loosely speaking, symptoms are deviations from a well-defined "normal behaviour". Symptoms are formally described by using Boolean-valued *predicates* that contains a *measurable variable*, e.g. T, such as

$$p_{T_{low}} = (T < T_{ss})$$

Because of the dependence of the symptom on a measurable signal, its value is *time-dependent*, and can be regarded as a qualitative-valued *performance output* of the process model.

A *family of symptoms* is a set of symptoms that are defined over the same measurable variable.

Diagnostic scenario. Similarly to an input-output scenario, that is a finite record of related input and output signals, a *diagnostic scenario is a timed sequence of symptoms from the same family* (i.e. over the same measurable output signal).

If one associates to the underlying measurable variable the symptoms defined thereon as qualitative variables, then a *diagnostic scenario can be regarded as a qualitative-valued output signal* of the system.

Root cause. In model-based fault detection and diagnosis one usually assigns a so called root cause to every faulty mode of the system, the variation of which acts as a cause of the fault. Root causes are most often not measurable and have discrete value (indicator variables) thus a root cause is described as an unmeasurable *disturbance* in a process model for diagnosis.

Preventive action. If a fault occurs it is usually possible to take actions in the initial phase of the transient to avoid serious consequences or to try to drive the system back to its original normal state. Dedicated *input signal(s)* serve for this purpose separately for each fault (identified by its root cause) where the *preventive action is a prescribed scenario for the manipulated input signal*.

3.3 Prediction for diagnosis

In model-based diagnosis *the model of the process system is assumed to describe the behaviour of the system in each of the considered faulty mode.* This is a quite severe requirement, thus one usually narrows the domain of the model by ***assuming that the "normal" operating mode of the system is steady-state and only the initial deviations are considered***.

The model is used for predicting *diagnostic scenarios* of the measured output variables for at least two purposes related to diagnosis.

- **Fault isolation**
  When the occurrence of a fault is detected the first task is to find out which is the root cause of the fault, i.e. to isolate the fault. For this task one uses the observed *diagnostic scenario* and tries to match it with the generated diagnostic scenarios by using *every possible root cause*. The generation of the possible diagnostic scenarios can be done by model-based prediction using the dynamic model of the process system.

- **Testing preventive actions**
  Having isolated the fault, i.e. assigned a root cause to the observed diagnostic scenario, one has to determine the course of actions to remedy the situation. This can be done by performing *"what-if"* type prediction by using the dynamic model of the process system and applying *every possible action*. The selection of the suitable preventive action can then be performed by comparing the final state of the system with the "normal" operation.

Because process systems are highly nonlinear and their model can be drastically changed depending on the actual fault mode, simple **reduced** models are needed separately for every *(root cause, diagnostic scenario, preventive action)* triplets.

3.4 External diagnostic knowledge: HAZOP and FMEA

HAZOP. Hazard and operability study (HAZOP), formalized by Imperial of Chemical Industries (ICI) at the end of the 1960s, is the most widely used methodology for hazard identification. HAZOP [Knowlton (1989)] is a systematic procedure for determining the causes of process deviations from normal behaviour and their consequences. The main idea behind HAZOP is that hazards in process plants arise as a result of deviations from normal operating conditions. A group of experts systematically identifies every conceivable deviations in a plant, finds all the possible abnormal causes, and the adverse consequences of that deviation. During the HAZOP these deviations are systematically analyzed by applying guide expressions (for example NONE, MORE OF, LESS OF, PART OF, MORE THAN, OTHER, ...) in conjunction with process variables and parameters. Driven by these guide words, failure causes and their effects are listed in a systematical way.

The results of the HAZOP analysis are collected in a HAZOP result table. A HAZOP analysis table (the structure of which is shown in Table 1) defines logical (static) cause-consequence relationships between symptoms and potential causes that can be traced to root causes of the deviation. These can be used for fault detection and isolation.

FMEA. Fault mode and effect analysis (FMEA) [Jordan (1972)] is a qualitative analysis of hazard identification, universally applicable in a wide variety of industries. Its use in the process industries has been more limited with HAZOP as one of the main contenders for the preferred hazard identification tool. FMEA is a tabulation of each piece of equipment, noting the various modes by which the equipment can fail, and the corresponding consequences (effects) of the failures. FMEA focuses on individual components and their failure modes. Thus, each failure mode is only considered once, and all of its effects and controls are listed together. This allows a more accurate assessment of the risk associated with each component failure

Guide word	**Deviation**	**Possible causes**	**Consequences**	**Action required**
Fresh Feed Flow	NONE	(1) Feed hopper empty	◇ loss of production ◇ shift in GSD ◇ decrease in recycle and output	a) feed the hopper b) check the hopper
		(2) Feed chute blockage	◇ Covered by (1)	c) check the hopper
	⋮	⋮	⋮	⋮

Symptoms : Guide word $\bigoplus$ Deviation
Root Cause : Possible causes
Action : Preventive actions
Scenario : Consequences

Table 1: An example of a HAZOP result table

4 Multiscale Process Models and Model Reduction

In the case of large and/or complex systems, the use of a multiscale modelling [Ingram *et al.* (2004), Ingram and Cameron (2004)] approach is recommended. The basis of multiscale modelling is to divide a complex problem into a family of sub-problems that exist at different scales. Multiscale models of a system can be organized along various scales depending on the system and on the intended use of the model. Generally, we distinguish between the length, time and detail scales, but diagnosis requires to have a multiscale model organized along the time scale [Németh *et al.* (2005)].

A multiscale model is then an ordered collection of partial models or sub-models that are connected by a so-called *integration framework*. The serial, subroutine-like organization [Ingram *et al.* (2004)], i.e. the simplest way the integration framework integrates the partial models, is used in this paper.

4.1 Multiscale modelling: the length and time scales

Traditionally, multiscale models are built along the length scale because the mechanisms that determine a model based on first engineering principles drive the model building. The levels of the multiscale model are found if one looks on the separation of the characteristic scales, if such separation exists.

In the case of process plant we generally distinguish at least the *molecular level* for chemical kinetics, *particle level* when applicable, *operating unit level* and *plant level* along the length scale. We try, if possible, build a multiscale model where we can solve the sub-models or partial models separately, i.e. when the sub-models are integrated using a *serial integration framework* [Ingram *et al.* (2004)]. This means that the solution of a sub-model in a lower level is used for developing so called *correlations*, i.e. static algebraic relationships between the variables on the higher level. In a recent PhD thesis [Ingram (2005)] the kernel functions of the granule bed level model are found by solving the equations of motion for a set of granules in the bed.

For model-based diagnosis, however, we focus on the time-dependent behaviour of our system, thus we have to arrange our sub-models along the *time scale*. Fortunately, the characteristic times that define the levels along the time scale in a multiscale model usually follow the separation of characteristic lengths, i.e. the characteristic time belonging to a particular length scale level is an order of magnitude larger than that belonging to a higher length level. For example, the characteristic time on a molecular level is in the order of seconds, while the characteristic time constants on an operating unit level are in the order of minutes (when particle level does not exists).

For a particular process system one can construct a so-called ***scale-map*** that relates the identified time and length scales and connects them to the mechanisms considered in the model. Fig. 3 shows

an example of a scale-map constructed for a granulator circuit.

It will be important for model-based diagnosis of multiscale process systems, that ***each variable is associated to a level in a model hierarchy*** that is determined by the mechanism or governing conservation balance/constitutive equation the particular variable is assigned to (i.e. is determined by). One can refine the scale-map of a multiscale process system by denoting regions associated to particular variables that are important from the viewpoint of diagnosis. Such a refined scale-map is called the ***model structure map***. An example of such a refined scale-map is seen in Fig. 7.

4.2 Diagnosis of multiscale models: the model reduction problem

As we have already seen in subsection 3.2, there is at least three characteristic variable that determines a *diagnostic scenario*: a root cause, a symptom variable that generates the symptoms and an input signal that is used for implementing preventive actions. Associated to these variables we have the following levels of interest along a scale, usually along the length scale:

- root *cause* level,
- *target* or symptom level,
- *control* or preventive action level.

Of course, the ideal case is when all the above three levels coincide, that is we select symptom variables and preventive actions from the same level as the root cause level is on.

Now we are ready to formulate the variant of the problem statement of model reduction applicable for multiscale models as follows.

Given:

1. a *multiscale process model* that is able to describe the systems behaviour under the considered faulty model (generated by the set of root causes),
2. a nominal *steady-state operating point*,
3. a *(root cause, symptom variable, preventive action)* triplet,
4. *reference input-output behaviour*: in the form of diagnostic scenarios with and without preventive actions,
5. *simplicity index*: the number of state variables and linearity.

Determine:

- a *reduced process model* on a single time scale level that is linear and generates the given reference behaviour for fault effect prediction.

Conceptual steps of the solution:

(1) Determine the *target level on the time scale* from the variables present in the symptoms of interest. If there is more than one time-level corresponding to the target length level, choose the one that corresponds to the given diagnostic scenarios.

(2) Select the dynamic conservation balances from the target, cause and control levels from the length scale that belong to the target level on the time scale. These together with their constitutive equations form the *scale-reduced nonlinear model*.

(3) Linearize the above nonlinear model (possibly in a DAE form) around the given nominal steady-state operating point. Form a standard linear state-space model by substituting the algebraic equations into the differential ones. The obtained model determines the structure of the *scale-reduced linear model*.

(4) Determine the model parameters of the above obtained scale-reduced linear model by standard parameter estimation methods either by using real measured data or by using data generated by simulation.

5 Case study: model reduction of a granule bed for diagnosis

The proposed model reduction method is demonstrated on a commercial fertilizer granulation drum example, where the granulator system is used for the production of mono-ammonium and di-ammonium phosphate (MAP, DAP) [Cameron *et al.* (2005)].

A typical granulation plant consists of a granulator drum, a dryer, screens and a crusher arranged in a granulator circuit as seen in Fig. 2. Details of the technology and the equipments involved are found elsewhere (see e.g. [Balliu (2004), Cameron *et al.* (2005)]).

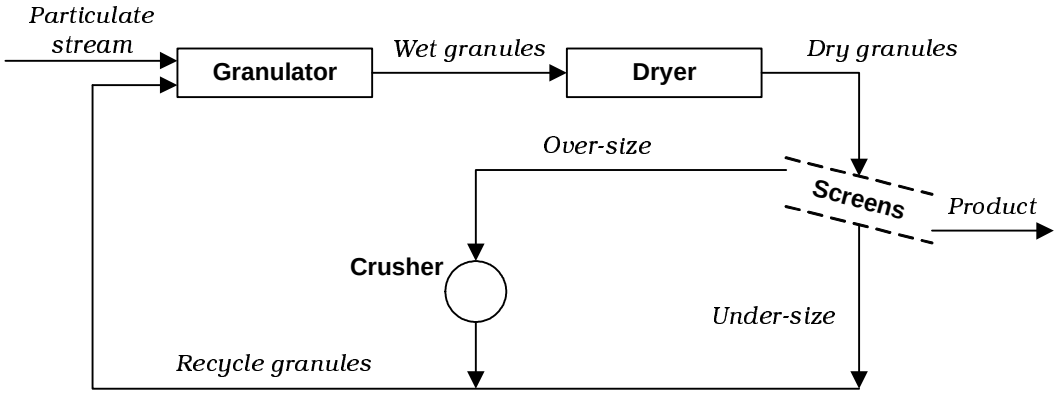

Figure 2: A flowsheet of a granulator circuit

5.1 Scale-map of a granulator circuit

If one considers the characteristic lengths or times of the various phenomena taking part in a granulator circuit then 5 scale levels can be distinguished along both scales as shown in Fig. 3 (adopted from [Ingram and Cameron (2004)]). The lower 4 levels belong to the granulator drum itself, where we concentrate our attention.

It is important to observe that there is a close relationship between the characteristic length and time scales of a granulator circuit. Generally, *the higher the level of length is the higher level of characteristic times corresponds to it.* An exception to this rule is the time-length scale relationship of the levels **Granule bed**, **Vessel** and **Granule**. Here the mechanisms characterizing the granules have a direct substantial influence on the dynamic behaviour of the variables on higher levels.

There is a clear separation of the characteristic lengths with a little overlap between the **Particle** and **Granule** levels. This enables to separate the overall model into sub-models along the time scale, this is how the multiscale modelling of granulation processes takes place (see e.g. [Ingram (2005)])

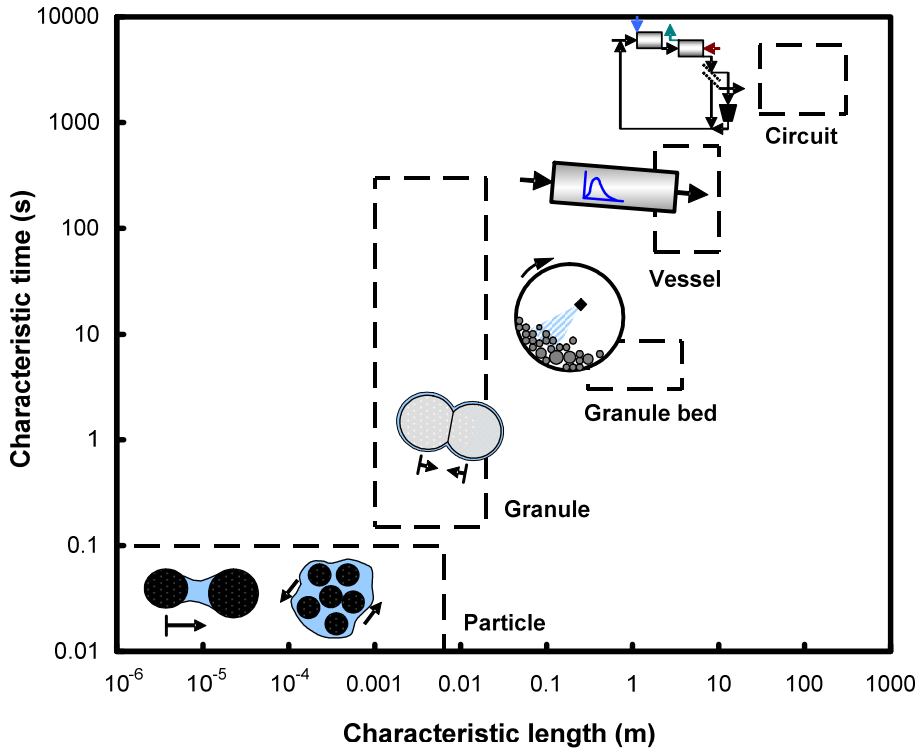

Figure 3: Scale-map of a granulator circuit

5.2 Granulator drum: levels, variables and mechanisms

As seen in Fig. 3 there are four scales or levels that can be identified in a granulator drum as follows.

- **Vessel** level: the whole vessel
  Variables: concentrations in the granules (pseudo-solid) and in the binder (liquid) phases
  Mechanisms: convection and phase transfer

- **Granule bed** level: slice of the vessel
  Variables: particle size distribution (PSD) of granules, *component*: a size range
  Mechanism: phase transfer, particle (solids) convection

- **Granule** level: a single granule
  Variables: size and composition of the granules together with their position in time
  Mechanisms: reaction, agglomeration, breakage, growth, coupled with collisions with each other and with the mechanical parts of the equipment

- **Particle** level: inter-granule processes, adsorption-desorption, reaction on the surface etc.

Variables and symptoms of the granulator drum model
Table 2 shows the identified variables and symptoms of the drum based on the results of the HAZOP studies.

Based on the scale-map of the granulator drum one can associate the symptoms and their variables listed in Table 2 to the levels of the multiscale model seen in Fig. 4.

Variable	Symptom
Binder Flow	NONE
Binder Flow	MORE
Binder Flow	LESS
Binder Viscosity	MORE
Binder Viscosity	LESS
Solids Feed PSD	NARROW
Solids Feed PSD	WIDE
Solids Feed Flow	NONE
Solids Feed Flow	MORE
Solids Feed Flow	LESS
Solids Feed Size	MORE
Solids Feed Size	LESS
Granulator Drum Speed	NONE
Granulator Drum Speed	MORE
Granulator Drum Speed	LESS
Granulator Exit Distribution	NARROW
Granulator Exit Distribution	WIDE
Granulator Exit Flow	NONE
Granulator Exit Flow	MORE
Granulator Exit Flow	LESS
Granulator Exit Size	MORE
Granulator Exit Size	LESS

Table 2: The list of variables and symptoms connected to the drum

5.3 Granulator drum: the model structure map

In order to develop the refined scale-map of the granulator drum to be used for identifying the target, cause and control levels on the time scale we need to have the detailed model of the system. The following model equations are considered adopted from [Balliu (2004)].

Components mass balances in the liquid phase: the change of *MAP, DAP* and H_2O mass over time

$$\frac{dm_{MAP}}{dt} = F_{L,in}^{MAP} + F_{SL}^{MAP} - F_{L,out}^{MAP} - c_1 \cdot r_{MAP/DAP} \tag{2}$$

$$\frac{dm_{DAP}}{dt} = F_{L,in}^{DAP} + F_{SL}^{DAP} - F_{L,out}^{DAP} - \dot{m}_{crystals} + c_2 \cdot r_{MAP/DAP} \tag{3}$$

$$\frac{dm_{H_2O}}{dt} = F_{L,in}^{H_2O} + (1-\varphi)F_{SL}^{H_2O} - F_{L,out}^{H_2O} - F_{H_2O}^{evap} \tag{4}$$

Overall mass balances

- liquid phase:

$$\frac{dM_L}{dt} = F_{L,in} + F_{SL} + f_{NH_3} - F_{H_2O}^{evap} - F_{L,out} - \dot{m}_{crystals} \tag{5}$$

- solids phase:

$$\frac{dM_S(i)}{dt} = F_{S,in}(i) + F_{SL}^{MAP,sol}(i) + F_{SL}^{DAP,sol}(i) - F_{S,out}(i) + \dot{m}_{crystals}(i) + Agg(i) + Lay(i) - Break(i) \tag{6}$$

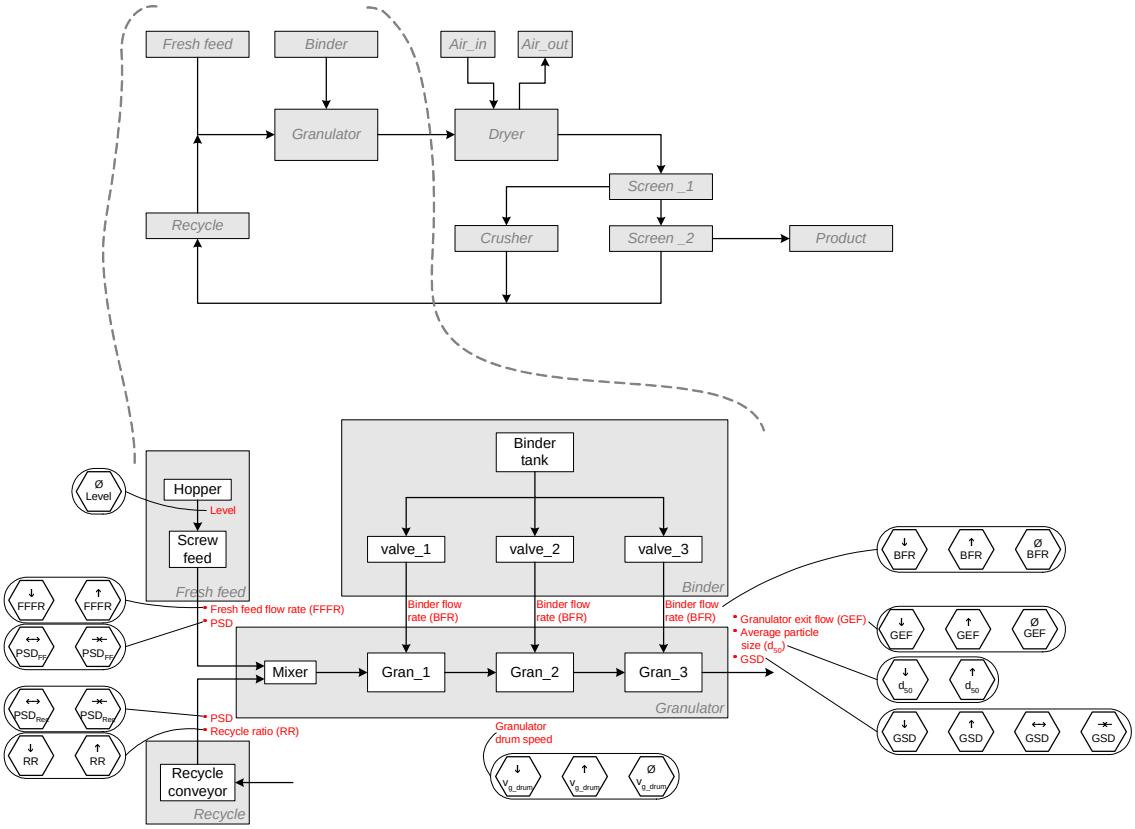

Figure 4: The hierarchy of symptoms

Energy balances

- liquid phase:

$$\frac{dE_L}{dt} = E_{L,in} + E_{SL} + E_{NH_3} + c_2 \cdot \Delta H \cdot r_{MAP/DAP} - E^{evap}_{H_2O} - E_{L,out} - E_{LS} - \dot{m}_{crystals} \cdot \Delta H_{crys} \quad (7)$$

- solids phase:

$$\frac{dE_S}{dt} = E_{S,in} + E^{MAP,sol}_{SL} + E^{DAP,sol}_{SL} - E_{S,out}(i) + E_{LS} + E_{crystals} \quad (8)$$

In the above equations the notation in Table 3 is used:

State-space model It is seen from the above model equations that Eqs. (2) - (8) form the state equations of the state-space model with the state vector

$$x = \left[\begin{array}{cccccccc} m_{MAP} & m_{DAP} & m_{H_2O} & M_L & M_S(i) & E_L & E_S \end{array} \right]^T$$

The input variables (to be manipulated or to have as disturbances) are

$$u = \left[\begin{array}{cccccccc} F^{MAP}_{SL} & F^{DAP}_{SL} & F^{NH_3}_{SL} & F^{H_2O}_{SL} & F^{MAP,sol}_{SL} & F^{DAP,sol}_{SL} & T_{SL} & \\ F^{MAP}_{SL} & F^{DAP}_{SL} & F^{NH_3}_{SL} & F^{H_2O}_{SL} & T_{L,in} & F_{NH_3} & T_{NH_3} & F_{S,in}(i) \quad T_{S,in} \quad T^{H_2O}_{S,in} \end{array} \right]^T$$

Variable	**Meaning**
m_{MAP}	mass of MAP in the liquid phase
$F_{L,in}^{MAP}$	flow of MAP into the liquid phase
F_{SL}^{MAP}	flow of MAP solution with the slurry stream
$F_{L,out}^{MAP}$	flow of MAP with the liquid phase out of the drum section
c_1	coefficient of reaction rate
$r_{MAP/DAP}$	reaction rate between ammonia and MAP
m_{DAP}	mass of DAP in the liquid phase
$F_{L,in}^{DAP}$	flow of DAP into the liquid phase
F_{SL}^{DAP}	flow of DAP solution with the slurry stream
$F_{L,out}^{DAP}$	flow of DAP with the liquid phase out of the drum section
$\dot{m}_{crystals}$	mass rate of crystallization onto existing solid phase
c_2	coefficient of reaction rate
m_{H_2O}	mass of H_2O in the liquid phase
$F_{L,in}^{H_2O}$	flow of H_2O into the liquid phase
φ	flash fraction of water from slurry flow as it exits the spray nozzle
$F_{SL}^{H_2O}$	flow of H_2O solution with the slurry stream
$F_{L,out}^{H_2O}$	flow of H_2O with the liquid phase out of the drum section
$F_{H_2O}^{evap}$	flow of water evaporated
M_L	mass holdup of liquid phase
$F_{L,in}$	flow of liquid phase in the drum section
F_{SL}	flow of slurry into the liquid phase section
f_{NH_3}	flow of ammonia uptaken into the liquid phase
$F_{L,out}$	flow of liquid phase out of the drum section
$M_S(i)$	mass holdup of the solids in each particle size interval i
$F_{L,in}(i)$	flow of solids in each particle size interval i flowing in each section
$F_{SL}^{MAP,sol}(i)$	flow of MAP crystals from the slurry stream deposited onto each size range i
$F_{SL}^{DAP,sol}(i)$	flow of DAP crystals from the slurry stream deposited onto each size range i
$\dot{m}_{crystals}(i)$	mass rate of crystallization into each size range i onto existing solid phase
$F_{S,out}(i)$	flow of solids in each particle size interval i flowing out of each section
$Agg(i)$	birth and death agglomeration in each particle size interval i
$Lay(i)$	layering in each particle size interval i
$Break(i)$	breakage into each particle size interval i
E_L	energy content in the liquid phase
$E_{L,in}$	energy content in flow of liquid phase in the drum section
E_{SL}	energy content in flow of slurry into the liquid phase section
E_{NH_3}	energy content in flow of ammonia uptaken into the liquid phase
ΔH	heat of reaction
$E_{H_2O}^{evap}$	energy content in flow of water evaporated
$E_{L,out}$	energy content in flow of liquid phase in the drum section
E_{LS}	energy transferred between the liquid phase and the solid phase
ΔH_{crys}	heat of crystallization
E_S	energy content in the solids
$E_{S,in}$	energy content in total flow of solids in each drum section
$E_{SL}^{MAP,sol}$	energy content in flow of MAP crystals in the slurry stream
$E_{SL}^{DAP,sol}$	energy content in flow of DAP crystals in the slurry stream
$E_{S,out}(i)$	energy content in total flow of solids out of each drum section
$E_{crystals}$	energy content of crystallization

Table 3: The list of variables in the granulator drum model

The structure of the state equations can be described by the following linear qualitative differential equation:

$$\dot{x} = Ax + Bu$$

where the matrices A and B are structure matrices with either fixed 0 or nonzero $\star$ elements as follows:

$$A = \begin{bmatrix} 0 & 0 & \star & \star & \star & 0 & 0 \\ \star & \star & \star & \star & \star & \star & 0 \\ 0 & 0 & \star & \star & \star & \star & 0 \\ \star & \star & \star & \star & \star & \star & 0 \\ \star & \star & 0 & \star & \star & \star & 0 \\ \star & \star & \star & \star & \star & \star & \star \\ \star & \star & 0 & \star & \star & \star & \star \end{bmatrix}$$

$$B = \begin{bmatrix} \star & 0 & 0 & 0 & 0 & 0 & 0 & \star & 0 & 0 & 0 & 0 & \star & 0 & 0 & 0 & 0 \\ 0 & \star & 0 & 0 & 0 & 0 & 0 & 0 & \star & 0 & 0 & 0 & \star & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \star & 0 & 0 & \star & 0 & 0 & 0 & \star & 0 & 0 & 0 & 0 & 0 & 0 \\ \star & \star & \star & \star & 0 & 0 & \star & \star & \star & \star & \star & 0 & \star & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \star & \star & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \star & 0 & 0 \\ \star & \star & \star & \star & 0 & 0 & \star & \star & \star & \star & \star & \star & \star & \star & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \star & \star & \star & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \star & \star & \star \end{bmatrix}$$

The above structure shows that the overall model is highly coupled in the general case, with little chance to apply black-box model reduction techniques (such as `modred` in MATLAB).

5.4 Fault scenarios and time-scale separation

In order to investigate the possibilities of time-scale separation and model reduction, and their dependence on the root cause of the faults, two fault scenarios have been considered:

(A) the increase in the binder flow ("Binder_flow=MORE") that acts primarily on the overall mass and energy variables,

(B) the increase in the width of the particle size density ("PSD=WIDE") that acts on every mechanisms and balance considered,

In both cases we have waited till a steady-state operating condition developed and then issued a step-like disturbance in the relevant variable to the system and observed the transient responses.

(A) Binder Flow – MORE the total amount of ammonia feed (F_{NH_3}) to the granulator section is increased at $t = 700s$. The simulation result can be seen in Figure 5.

A time-scale separation between the masses of the particle size classes (slow variables) and the rest of the state variables can be observed in Fig. 5 with at least an order of magnitude difference in the dominant time constant. This is in a good agreement with our engineering expectations that the increase in the NH_3 feed acts primarily on the overall mass and energy variables and only a slower, secondary effect is expected in the PSD variables on the granule level.

(B) Solids Feed PSD – WIDE the particle size distribution of solids feed to the granulator is made wider (more smaller size particles) at $t = 700s$. The Figure 6 shows the transient of the state variables.

There is no time-scale separation in this case.

Time-scale separation Based on the step-response scenarios above, we can construct a refined scale-map of the drum model (see in Fig. 7) by indicating the time-scale of the change in the particular variables effected by the root cause of increasing the binder flowrate.

6 Conclusions and future work

A systematic approach is proposed in this paper in order to reduce the number of state variables and parameters of multiscale process models in prediction-based fault detection and diagnosis. The method requires a well-documented multiscale process model that is able to describe all considered faulty modes of the system. In addition, the list of faulty modes together with their characteristic time scales and dominant mechanisms are needed that drive the reduction procedure applied for each faulty mode.

The proposed model structure-driven model reduction method applies a refined scale-map, the so called model structure map to form the scale-reduced nonlinear model of the system that only contains the dynamic balance equations on the target time scale with all the other dynamic equations reduced by steady-state assumptions. Thereafter local linearization is applied to obtain the final reduced model.

The resulted set of reduced models preserves the engineering meaning of the remaining state variables, while the number of state variables and parameters is decreased significantly.

The proposed model can only be applied if the following conditions hold.

- The system is in the initial phase of a fault.
- The operating point is steady-state.
- There is a time-scale separation in the set of dynamic conservation balance equations to be reduced.

References

[Baldea and Daoutidis (2005)]
Baldea, M., Daoutidis M.: Dynamics and control of integrated process network with multi-rate reactions. Prep. IFAC World Congress, Prague, Czech Republic (2005)

[Balliu (2004)]
Balliu, N.: An object-oriented approach to the modelling and dynamics of granulation circuits, PhD Thesis, School of Engineering, The University of Queensland, Australia 4072. (2004)

[Cameron *et al.* (2005)]
Cameron, I. T., Wang, F. Y., Immanuel, C. D., Stepanek, F.: Process systems modelling and applications in granulation: A review. Chemical Engineering Science, **60**, (2005) 3723–3750

[Hahn *et al.* (2003)] Hahn, J., Edgar, T.F., Marquardt, W.: Controllability and observability covariance matrices for the analysis and order reduction of stable nonlinear systems. Journal of Process Control, **13**, (2003) 115–127

[Hangos *et al.* (2005)]
Lakner, R., Hangos, K. M., Cameron, I. T.: On minimal models of process systems. Chemical Engineering Science, **60**, (2005), 1127–1142

[Hangos and Cameron (2001)]
Hangos, K. M., Cameron, I. T.: Process modelling and model analysis. Academic Press, London (2001)

[Ingram *et al.* (2004)]
Ingram, G. D., Cameron, I. T., Hangos, K. M.: Classification and analysis of integrating frameworks in multiscale modelling. Chemical Engineering Science **59** (2004) 2171–2187

[Ingram and Cameron (2004)]
Ingram, G. D., Cameron, I. T.: Challenges in Multiscale Modelling and its Application to Granulation Systems. Development in Chemical Engineering and Mineral Processing **12(3/4)** (2004) 293–308

[Ingram (2005)]
Ingram, G. D.: Challenges in Multiscale Modelling (to be made more precise) PhD Thesis, The University of Queensland (2005)

[Jordan (1972)]
Jordan, W.: Failure modes, effects and criticality analyses. In: Proceedings of the Annual Reliability and Maintainability Symposium. IEEE Press (1972) 30 – 37

[Knowlton (1989)]
Knowlton, R. E.: Hazard and operability studies : the guide word approach. Vancouver: Chematics International Company (1989)

[Kumar and Daoutidis (2002)]
Kumar, A., Daoutidis, P.: Nonlinear dynamics and control of process systems with recycle. Journal of Process Control **12** (2002) 475 – 484

[Leitold *et al.* (2002)]
Leitold, A., Hangos, K. M. and Tuza, Z.: Structure simplification of dynamic process models. J. Process Control **12** (2002) 69–83

[Németh *et al.* (2005)]
Németh, E., Cameron, I. T., Hangos, K. M.: Diagnostic goal driven modelling and simulation of multiscale process systems. Computers and Chemical Engineering **29** (2005) 783–796

[Venkatasubramanian *et al.* (2003)]
Venkatasubramanian, V., Rengaswamy, R., Kavuri, S. N.: A review of process fault detection and diagnosis Part II: Qualitative models and search strategies. Computers and Chemical Engineering **27** (2003) 313–326

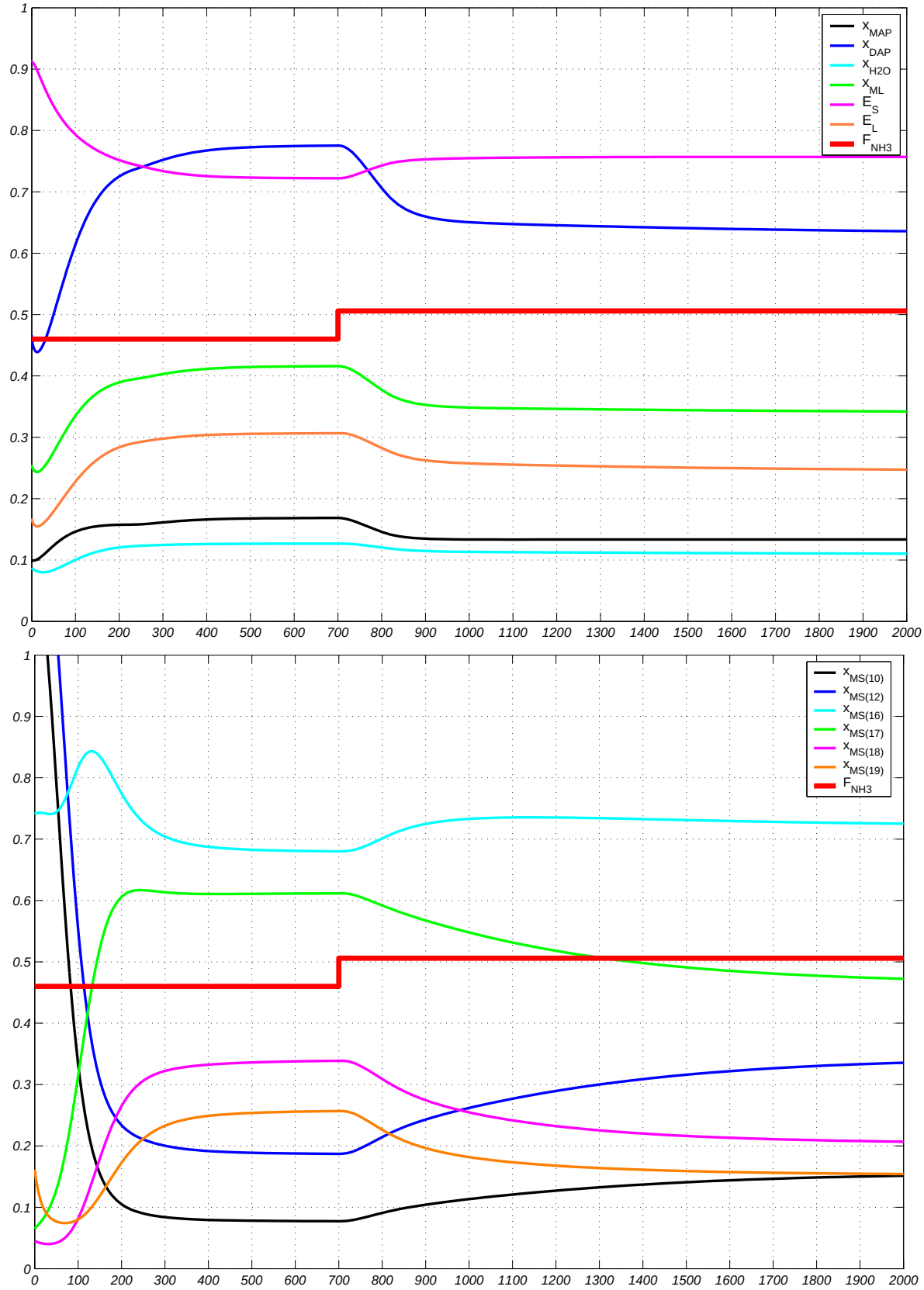

Figure 5: Simulation result of the increased binder flow rate

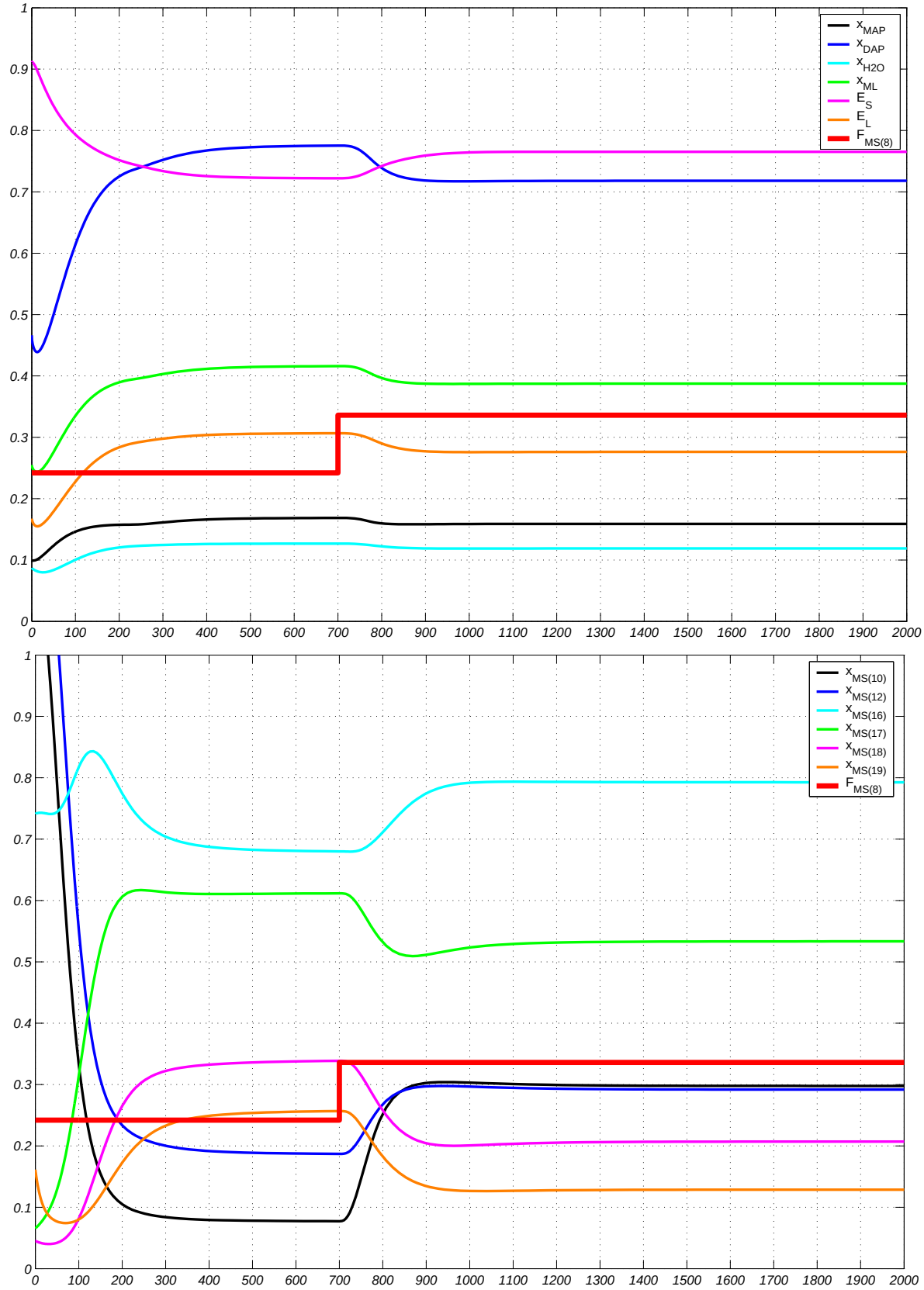

Figure 6: Simulation result of the PSD changing of the granulator solids feed

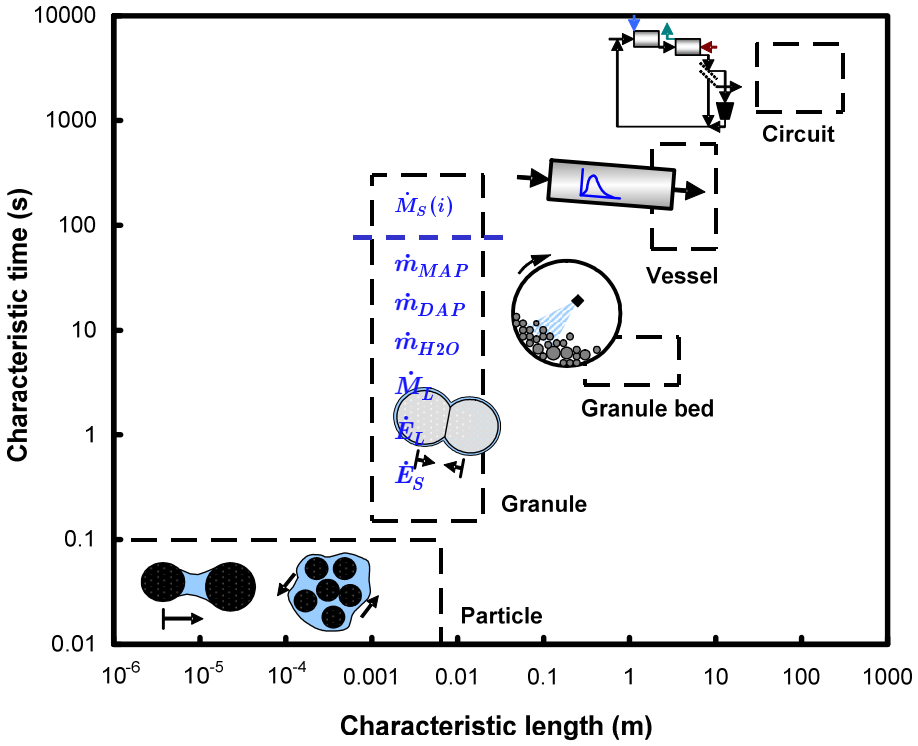

Figure 7: The model structure map (refined scale-map) of a granulator circuit