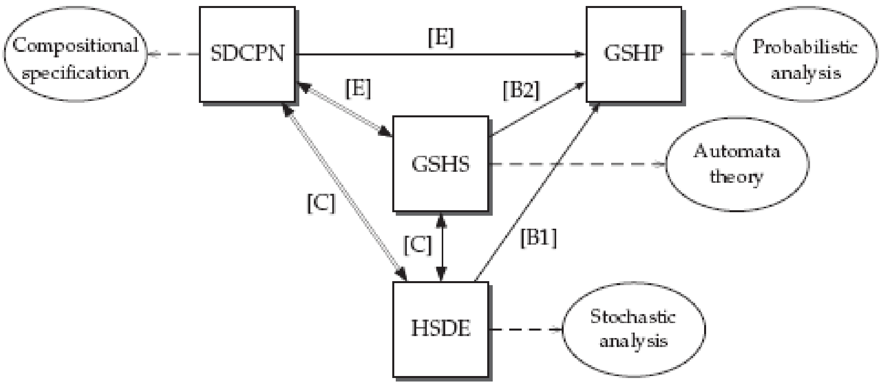
NLR-TP-2010-324

NLR Air Transport Safety Institute
Research & Consultancy

Executive summary

Hybrid state Petri nets which have the analysis power of stochastic hybrid systems and the formal verification power of automata



Problem area
In order to combine the compositional specification power of Petri nets with the analysis power of Markov processes, Malhotra & Trivedi (1994) Muppala & Fricks & Trivedi (2000) developed a power hierarchy of dependability models. In Everdij & Blom (2003, 2005), the power hierarchy was extended with dynamically coloured Petri nets (DCPN) and piecewise deterministic Markov processes (PDP). In Everdij Blom (2006), this power hierarchy was further extended by stochastically and dynamically coloured Petri nets (SDCPN) and general stochastic hybrid process (GSHP).

Description of work
In this paper the power-hierarchy has been further deepened by studying various ways to develop GSHP. We first define SDCPN and the resulting SDCPN process. Next, we study GSHP as an execution of a general stochastic hybrid system (GSHS). Subsequently, we define GSHP as a solution of a hybrid stochastic differential equation (HSDE) and explain the differences between GSHS and HSDE. Next, we show that GSHS, HSDE and SDCPN are bisimilar. Finally, the results are illustrated with an aircraft evolution example.

Report no.
NLR-TP-2010-324

Author(s)
M.H.C. Everdij
H.A.P. Blom

Report classification
UNCLASSIFIED

Date
February 2010

Knowledge area(s)
Vliegveiligheid (safety & security)

Descriptor(s)
Compositional modelling
Petri nets
Air Traffic Management
Safety
Hybrid automata

This report is based on chapter 12 in the book "Petri nets: applications", Pawel Pawlewski (editor), pages 227-252, February 2010, INTECH, ISBN 978-953-307-047-6.

Results and conclusions
The bisimilarities between SDCPN, GSHS and HSDE mean that each of them inherits the strengths of the other two formalisms. Hence, analysis tools designed for GSHS, HSDE and GSHP and their properties become available for SDCPN. Examples of GSHP properties are convergence in discretisation, existence of limits, existence of event probabilities, strong Markov properties, and reachability analysis. Examples of GSHS features are their connection to formal methods in automata theory and optimal control theory. Examples of HSDE features are stochastic analysis tools for semi-martingales. At the same time, numerous SDCPN features such as natural expression of causal dependencies, concurrency and synchronisation mechanism, hierarchical and modular construction, and graphical representation become available when modelling GSHS, HSDE and GSHP through SDCPN.
These complementary advantages of SDCPN, GSHS, HSDE and GSHP perspectives tend to increase with the complexity of the system considered.

NLR Air Transport Safety Institute

Anthony Fokkerweg 2, 1059 CM Amsterdam,
P.O. Box 90502, 1006 BM Amsterdam, The Netherlands
Telephone +31 20 511 35 00, Fax +31 20 511 32 10, Web site: http://www.nlr-atsi.nl

NLR Air Transport Safety Institute
Research & Consultancy

NLR-TP-2010-324

Hybrid state Petri nets which have the analysis power of stochastic hybrid systems and the formal verification power of automata

M.H.C. Everdij
H.A.P. Blom

This report is based on chapter 12 in the book "Petri nets: applications", Pawel Pawlewski (editor), pages 227-252, February 2010, INTECH, ISBN 978-953-307-047-6.

Customer	National Aerospace Laboratory NLR
Division	Air Transport
Owner	NLR
Distribution	Unlimited
Classification of title	Unclassified
	February 2010

Approved by:

Author	Reviewer	Managing department
[signature] 27/8/2010	Anonymous Peer Reviewers	[signature] 2/9/10

Contents

12

Hybrid state Petri nets which have the analysis power of stochastic hybrid systems and the formal verification power of automata

Mariken H.C. Everdij and Henk A.P. Blom
National Aerospace Laboratory NLR
The Netherlands

1. Introduction

For a large range of complex applications, governments and industries invest in the development of innovative new systems existing of many distributed components that interact in a dynamic way with many uncertainties. Before any such system can be introduced into practice, an evaluation needs to have shown that both the system and the way it is used in its new context realizes the applicable objectives. If the new complex system is in its interactions similar to a previous system, such investigation can be done by analysis judgement of capable and experienced experts who judge local behaviour and implicitly assume that the interactions are working as before. If the complex system is very different from the old system, then this expert judgement approach falls short. A valuable alternative is to develop a mathematical model that incorporates the interactions, analyse this model, mobilise domain experts to evaluate where the model is representative for reality and where it needs improvement, and learn to understand how the real system works by learning how the model works. This requires a growing need for modelling and analysis of stochastic hybrid systems.

Petri nets, e.g. (David & Alla, 1994), have shown to be useful for developing models of various complex applications. Typical Petri net features are concurrency and synchronisation mechanism, hierarchical and modular construction, and natural expression of causal dependencies, in combination with graphical and equational representation. Numerous extensions to the basic formalism have been developed that combine different modelling features in an integrated way, including various hybrid state Petri net versions, e.g. (Giua, 1999), which combine discrete and continuous system aspects.

As a powerful class of models that support stochastic analysis, (Davis, 1984; 1993) introduced *piecewise deterministic Markov processes* (PDPs) as the most general class of continuous-time hybrid state Markov processes which include both discrete and continuous processes, except diffusion. In (Bujorianu & Lygeros, 2003; Hu et al., 2000) the PDPs have been defined as stochastic hybrid automata. Subsequently, diffusion by means of Brownian motion has been incorporated (Bujorianu & Lygeros, 2006). This way, a formal connection is established between stochastic hybrid processes that are supported by powerful stochastic analysis tools (Davis, 1993; Elliott, 1982; Elliott et al., 1995) and the automata formalism to develop formal verification tools (Frehse, 2008; Kwiatkowska et al., 2004; Labinaz et al., 1997).

In order to combine the advantages of the Petri net modelling formalisms and those of the Markovian analysis formalism, (Malhotra & Trivedi, 1994) and (Muppala et al., 2000) started the development of establishing formal connections between Petri nets and stochastic processes. Their result is a hierarchy of various dependability models based on their modelling power. At the left-hand-side of this power hierarchy are Petri net models, with *generalised stochastic Petri nets* (GSPN) at the bottom, and *deterministic and stochastic Petri nets* (DSPN) at the top. At the right-hand-side of this power hierarchy are *Markov chains* at the bottom and *semi-Markov processes* at the top. Arrows between different formalisms indicate that mappings exist, i.e. that the elements of one formalism can be represented in terms of the elements of the other formalism, such that the executions, i.e. their solutions as a stochastic process, are equivalent. In a series of studies (Everdij & Blom, 2003; 2005; 2006) developed an extension of this power hierarchy in probabilistic modelling, see Fig. 1.

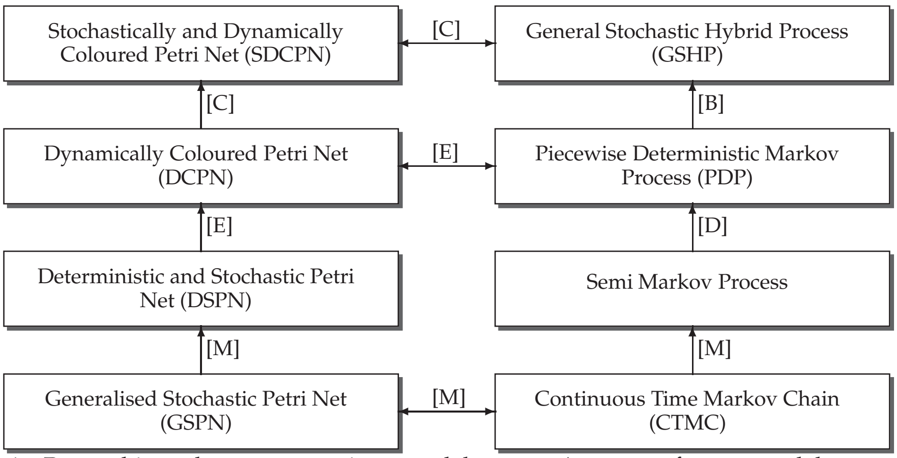


Fig. 1. Power hierarchy among various model types. An arrow from a model to another model indicates that the second model has more modelling power than the first model. The [M] arrows have been established in (Malhotra & Trivedi, 1994; Muppala et al., 2000). The [D] arrow is established in (Davis, 1984). The [B] arrow is established in (Bujorianu & Lygeros, 2006) and in (Blom, 2003). The [E] arrows are established in (Everdij & Blom, 2003; 2005). The [C] arrows are established in (Everdij & Blom, 2006) and the current chapter.

At the left hand side of this power hierarchy, we extended DSPN to *dynamically coloured Petri nets* (DCPN) and further to *stochastically and dynamically coloured Petri nets* (SDCPN). At the right hand side of the power hierarchy we extended semi Markov processes to *piecewise deterministic Markov processes* (PDP) and further to *general stochastic hybrid processes* (GSHP). In addition we showed effective ways how a DCPN can be mapped into PDP and the other way around, and how SDCPN can be mapped into GSHP and the other way around. DCPN and SDCPN are hybrid Petri net classes in which the tokens have Euclidean-valued colours that change through time (*dynamically*) while the tokens reside in their place. For DCPN, these colours follow ordinary differential equations, for SDCPN, the colours follow stochastic differential equations. The specific strength of (S)DCPN is their compositional specification power, which makes available a hierarchical modelling approach that separates local modelling issues from global modelling issues. This is illustrated for a large distributed example in air traffic management (Everdij et al., 2006), which covers many distributed agents each of which interacts in a dynamic way with the others. Through a series of studies (Strubbe & Van der Schaft, 2005) developed a powerful compositional specification approach for automaton of PDP type (i.e. without Brownian motion), but for the complex air traffic management example (S)DCPN was shown to be better at compositional specification (Strubbe & Van der Schaft, 2004, Section 5.2).

For the mappings developed in (Everdij & Blom, 2006) between SDCPN and GSHP we made use of the *general stochastic hybrid system* (GSHS) theoretical setting developed by (Bujorianu & Lygeros, 2006), where GSHP is defined as an execution of a GSHS. More specifically, this means that SDCPN and GSHS are *bisimilar* in the sense that executions of SDCPN and GSHS yield GSHPs which are probabilistically equivalent, see e.g. (Bujorianu et al., 2005; Van der Schaft, 2004). Because of this bisimilarity, each formalism can take advantage of the strengths of both of them (Everdij & Blom, 2008).

Although the progress in the development of GSHP as an execution of a GSHS has led to significant increase of available stochastic analysis tools, there are some remaining issues to be addressed:

- *Jump linear systems* are not well covered, which unfortunately excludes most existing work on stochastic hybrid systems.
- *Semi-martingale property* of GSHS execution is unknown, which prohibits the use of Itô's differentiation rule for semi-martingales.

In the current chapter, these issues are further developed by considering GSHP not only as GSHS executions, but also as solutions of *hybrid stochastic differential equations* (HSDE). The HSDE approach towards studying GSHP has been developed in a series of complementary studies (Blom, 2003; Blom et al., 2003; Krystul, 2006; Krystul et al., 2007). The aim of this chapter is to characterise the relations between SDCPN, GSHP, HSDE and GSHS and to show that SDCPN, GSHS and HSDE are bisimilar. Fig. 2 shows the relations between the formalisms, and the key tools available for each of them.

With these relations, the properties and advantages of the various approaches come within reach of each other. The compositional specification power of SDCPN makes it relatively easy to develop a model for a complex system with multiple interactions. Subsequently, in the analysis stage three alternative approaches can be taken. The first is direct execution of SDCPN and evaluation through e.g. Monte Carlo simulation. The second is mapping the SDCPN into a GSHS and evaluating its execution, with the advantages of connection to formal methods in automata theory and to optimal control theory (Bujorianu & Lygeros, 2004). The third is mapping the SDCPN into HSDE and evaluating its solution, with the advantages of stochastic analysis for semi-martingales (Elliott, 1982; Elliott et al., 1995). With the GSHP resulting from any of these three means, properties become available such as convergence of discretisation, existence of limits, existence of event probabilities, strong Markov properties, and reachability analysis (Bujorianu & Lygeros, 2006; Davis, 1993; Ethier & Kurtz, 1986).

The organisation of this chapter is as follows. Section 2 defines SDCPN and the related SDCPN process. Section 3 defines GSHS and its GSHS execution. Section 4 defines HSDE and its stochastic process solution. Section 5 shows that SDCPN, GSHS and HSDE are bisimilar. Section 6 gives an example SDCPN. Section 7 presents this SDCPN example by an HSDE and by a GSHS. Section 8 gives conclusions. The appendices provide proof for the theorems in Section 5.

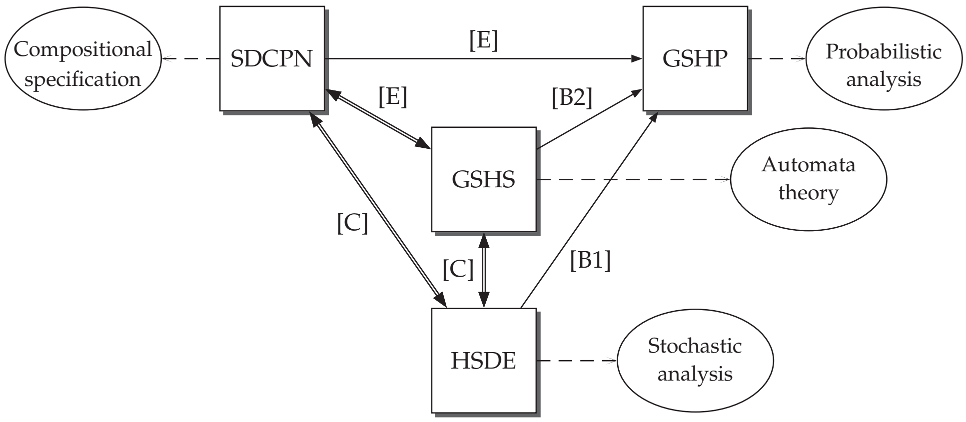


Fig. 2. Relationship between SDCPN, GSHS, GSHP and HSDE, and their key properties and advantages. The [B1] arrow is established in (Blom, 2003). The [B2] arrow is established in (Bujorianu & Lygeros, 2006). The [E] arrows are established in (Everdij & Blom, 2006). The [C] arrows are established in the current chapter, with bisimilarity relations having two-directional arrows.

2. SDCPN

This section presents a definition of *stochastically and dynamically coloured Petri net* (SDCPN).

Definition 2.1 (Stochastically and dynamically coloured Petri net)**.** *An SDCPN is a collection of elements* $(\mathcal{P}, \mathcal{T}, \mathcal{A}, \mathcal{N}, \mathcal{S}, \mathcal{C}, \mathcal{I}, \mathcal{V}, \mathcal{W}, \mathcal{G}, \mathcal{D}, \mathcal{F})$*, together with an SDCPN execution prescription which makes use of a sequence* $\{U_i; i = 0, 1, \ldots\}$ *of independent uniform* $U[0, 1]$ *random variables, of independent sequences of mutually independent standard Brownian motions* $\{B_t^{i,P}; i = 1, 2, \ldots\}$ *of appropriate dimensions, one sequence for each place P, and of five rules R0–R4 that solve enabling conflicts.*

The formal SDCPN definition provided below is organised as follows: Section 2.1 defines the SDCPN elements $(\mathcal{P}, \mathcal{T}, \mathcal{A}, \mathcal{N}, \mathcal{S}, \mathcal{C}, \mathcal{I}, \mathcal{V}, \mathcal{W}, \mathcal{G}, \mathcal{D}, \mathcal{F})$. Section 2.2 explains the SDCPN execution, which makes use of the rules R0–R4. Section 2.3 explains how the SDCPN execution defines a unique stochastic process.

2.1 SDCPN elements

The SDCPN elements $(\mathcal{P}, \mathcal{T}, \mathcal{A}, \mathcal{N}, \mathcal{S}, \mathcal{C}, \mathcal{I}, \mathcal{V}, \mathcal{W}, \mathcal{G}, \mathcal{D}, \mathcal{F})$ are defined as follows:

- $\mathcal{P}$ is a finite set of places.
- $\mathcal{T}$ is a finite set of transitions which consists of 1) a set $\mathcal{T}_G$ of guard transitions, 2) a set $\mathcal{T}_D$ of delay transitions and 3) a set $\mathcal{T}_I$ of immediate transitions.
- $\mathcal{A}$ is a finite set of arcs which consists of 1) a set $\mathcal{A}_O$ of ordinary arcs, 2) a set $\mathcal{A}_E$ of enabling arcs and 3) a set $\mathcal{A}_I$ of inhibitor arcs.

- $\mathcal{N} : \mathcal{A} \to \mathcal{P} \times \mathcal{T} \cup \mathcal{T} \times \mathcal{P}$ is a node function which maps each arc $A \in \mathcal{A}$ to a pair of ordered nodes $\mathcal{N}(A)$, where a node is a place or a transition[1]. The place of $\mathcal{N}(A)$ is denoted by $P(A)$, the transition of $\mathcal{N}(A)$ is denoted by $T(A)$, such that for all $A \in \mathcal{A}_E \cup \mathcal{A}_I$: $\mathcal{N}(A) = (P(A), T(A))$ and for all $A \in \mathcal{A}_O$: either $\mathcal{N}(A) = (P(A), T(A))$ or $\mathcal{N}(A) = (T(A), P(A))$. Further notation:
  - $A(T) = \{A \in \mathcal{A} \mid T(A) = T\}$ denotes the set of arcs connected to transition T, $A_{in}(T) = \{A \in A(T) \mid \mathcal{N}(A) = (P(A), T)\}$ is the set of input arcs of T, $A_{out}(T) = \{A \in A(T) \mid \mathcal{N}(A) = (T, P(A))\}$ is the set of output arcs of T, $A_{in,O}(T) = A_{in}(T) \cap \mathcal{A}_O$ is the set of ordinary input arcs of T, $A_{in,OE}(T) = A_{in}(T) \cap \{\mathcal{A}_E \cup \mathcal{A}_O\}$ is the set of input arcs of T that are either ordinary or enabling, and
  - $P(\mathcal{A}^{\subset}) = \{P(A); A \in \mathcal{A}^{\subset}\}$ is the multi-set of places connected to the subset of arcs $\mathcal{A}^{\subset} \subset \mathcal{A}$.

  Finally, $\{A_i \in \mathcal{A}_I \mid \exists A \in \mathcal{A}, A \neq A_i : \mathcal{N}(A) = \mathcal{N}(A_i)\} = \emptyset$, i.e., if an inhibitor arc points from a place P to a transition T, there is no other arc from P to T.
- $\mathcal{S} \subset \{\mathbb{R}^0, \mathbb{R}^1, \mathbb{R}^2, \ldots\}$ is a finite set of colour types, with $\mathbb{R}^0 \triangleq \emptyset$.
- $\mathcal{C} : \mathcal{P} \to \mathcal{S}$ is a colour type function which maps each place $P \in \mathcal{P}$ to a specific colour type in $\mathcal{S}$. Each token in P is to have a colour in $\mathcal{C}(P)$. Since $\mathcal{C}(P) \in \{\mathbb{R}^0, \mathbb{R}^1, \ldots\}$, there exists a function $n : \mathcal{P} \to \mathbb{N}$ such that $\mathcal{C}(P) = \mathbb{R}^{n(P)}$. If $\mathcal{C}(P) = \mathbb{R}^0 \triangleq \emptyset$ then a token in P has no colour. Further notation: if $P(\mathcal{A}^{\subset})$ contains more than one place, e.g., $P(\mathcal{A}^{\subset}) = \{P_{i_1}, \ldots, P_{i_k}\}$, then $\mathcal{C}(P(\mathcal{A}^{\subset}))$ is defined by $\mathcal{C}(P_{i_1}) \times \cdots \times \mathcal{C}(P_{i_k})$.
- $\mathcal{I} : \mathbb{N}^{|\mathcal{P}|} \times \mathcal{C}(\mathcal{P})^{\mathbb{N}} \to [0,1]$ is a probability measure, which defines the initial marking of the net: for each place it defines a number ≥ 0 of tokens initially in it and it defines their initial colours. Here, $\mathbb{N}^{|\mathcal{P}|} \triangleq \{(m_1, \ldots, m_{|\mathcal{P}|}); m_i \in \mathbb{N}, m_i < \infty, i = 1, \ldots, |\mathcal{P}|\}$ and $\mathcal{C}(\mathcal{P})^{\mathbb{N}} \triangleq \{\mathcal{C}(P_1)^{m_1} \times \cdots \times \mathcal{C}(P_{|\mathcal{P}|})^{m_{|\mathcal{P}|}}; m_i \in \mathbb{N}, m_i < \infty, i = 1, \ldots, |\mathcal{P}|\}$, where $\mathcal{C}(P_i)^{m_i} \triangleq \mathbb{R}^{m_i n(P_i)}$ for all $i = 1, \ldots, |\mathcal{P}|$, where $\mathcal{P}$ is denoted $\mathcal{P} = \{P_1, \ldots P_{|\mathcal{P}|}\}$. It is assumed that all tokens in a place are distinguishable by a unique identification tag which translates to a unique ordering/listing of tokens per place.
- $\mathcal{V} = \{\mathcal{V}_P; P \in \mathcal{P}, \mathcal{C}(P) \neq \mathbb{R}^0\}$ is a set of token colour functions. For each place $P \in \mathcal{P}$ for which $\mathcal{C}(P) \neq \mathbb{R}^0$, it contains a function $\mathcal{V}_P : \mathcal{C}(P) \to \mathcal{C}(P)$ that defines the drift coefficient of a differential equation for the colour of a token in place P.
- $\mathcal{W} = \{\mathcal{W}_P; P \in \mathcal{P}, \mathcal{C}(P) \neq \mathbb{R}^0\}$ is a set of token colour matrix functions. For each place $P \in \mathcal{P}$ for which $\mathcal{C}(P) \neq \mathbb{R}^0$, it contains a measurable mapping $\mathcal{W}_P : \mathcal{C}(P) \to \mathbb{R}^{n(P) \times h(P)}$ that defines the diffusion coefficient of a stochastic differential equation for the colour of a token in place P, where $h : \mathcal{P} \to \mathbb{N}$. It is assumed that $\mathcal{W}_P$ and $\mathcal{V}_P$ satisfy conditions that ensure a probabilistically unique solution of each stochastic differential equation.[2]

[1] The SDCPN arcs have no arc weights, but this node function definition leaves the freedom to define multiple arcs between the same pair of transition and place or place and transition (except if an inhibitor arc is involved).

[2] In the earlier definition by (Everdij & Blom, 2006) it was assumed that $\mathcal{V}_P$ and $\mathcal{W}_P$ satisfy local Lipschitz condition. This condition has now been relaxed to probabilistic uniqueness of solution of the related stochastic differential equation(s).

- $\mathcal{G} = \{\mathcal{G}_T; T \in \mathcal{T}_G\}$ is a set of transition guards. For each $T \in \mathcal{T}_G$, it contains a transition guard $\mathcal{G}_T$, which is an open subset in $\mathcal{C}(P(A_{in,OE}(T)))$ with boundary $\partial\mathcal{G}_T$. If $\mathcal{C}(P(A_{in,OE}(T))) = \mathbb{R}^0$ then $\partial\mathcal{G}_T = \emptyset$.[3] There is no requirement that $\mathcal{G}_T$ be connected.
- $\mathcal{D} = \{\mathcal{D}_T; T \in \mathcal{T}_D\}$ is a set of transition delay rates. For each $T \in \mathcal{T}_D$, it contains a locally integrable transition delay rate $\mathcal{D}_T : \mathcal{C}(P(A_{in,OE}(T))) \to \mathbb{R}^+$. If $\mathcal{C}(P(A_{in,OE}(T))) = \mathbb{R}^0$ then $\mathcal{D}_T$ is a constant function. [4]
- $\mathcal{F} = \{\mathcal{F}_T; T \in \mathcal{T}\}$ is a set of firing measures. For each $T \in \mathcal{T}$, it contains a firing measure $\mathcal{F}_T : (\{0,1\}^{|A_{out}(T)|} \times \mathcal{C}(P(A_{out}(T)))) \times \mathcal{C}(P(A_{in,OE}(T))) \to [0,1]$, which generates the number and colours of the tokens produced when transition T fires, given the value of the vector $\in \mathcal{C}(P(A_{in,OE}(T)))$ that collects all input tokens: For each output arc ($\in A_{out}(T)$), zero or one token is produced, and if the colours of the tokens produced are collected in a vector, this vector is $\in \mathcal{C}(P(A_{out}(T)))$. For each fixed $H \subset \mathcal{C}(P(A_{out}(T)))$, $\mathcal{F}_T(H; \cdot)$ is measurable. For any $c \in \mathcal{C}(P(A_{in,OE}(T)))$, $\mathcal{F}_T(\cdot; c)$ is a probability measure. Here, $\{0,1\}^{|A_{out}(T)|} \triangleq \{(e_1, \ldots, e_{|A_{out}(T)|}); e_i \in \{0,1\}, i = 1, \ldots, |A_{out}(T)|\}$.

For the places, transitions and arcs, the graphical notation is as in Figure 3.

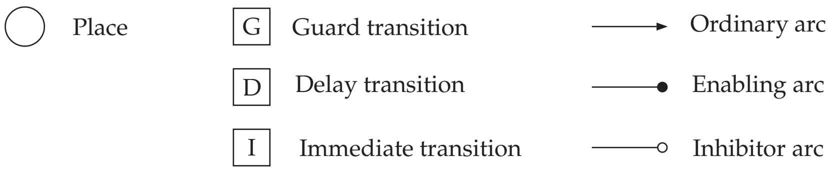


Fig. 3. Graphical notation for places, transitions and arcs in an SDCPN

2.2 SDCPN execution

The execution of an SDCPN provides a series of increasing stopping times, $0 = \tau_0 < \tau_1 < \tau_2 < \cdots$, with for $t \in (\tau_k, \tau_{k+1})$ a fixed number of tokens per place and per token a colour which is the solution of a stochastic differential equation. It uses a sequence $\{U_i; i = 0, 1, \ldots\}$ of independent uniform $U[0,1]$ random variables, and independent sequences of mutually independent standard Brownian motions $\{B_t^{i,P}; i = 1, 2, \ldots\}$ of appropriate dimensions, one sequence for each place P.

Initiation

The probability measure $\mathcal{I}$ characterises an initial marking at τ_0, i.e. it gives each place $P \in \mathcal{P}$ zero or more tokens and gives each token in P a colour in $\mathcal{C}(P)$, i.e. a Euclidean-valued vector. Define the inverse of $\mathcal{I}$ by a measurable function $\mathcal{I}^{inv} : [0,1] \to \mathbb{N}^{|\mathcal{P}|} \times \mathcal{C}(\mathcal{P})^{\mathbb{N}}$ such that $\mu_L\{u \mid \mathcal{I}^{inv}(u) \in H\} = \mathcal{I}(H)$, for H Borel measurable and μ_L the Lebesgue measure. Then the initial marking is a hybrid random vector characterised by $(M_0, C_0) = \mathcal{I}^{inv}(U_0)$. Here, M_0 is a $|\mathcal{P}|$-dimensional vector of non-negative integers, the ith component $M_{i,0}$ of which denotes

[3] In earlier SDCPN definitions, the transition guard was defined as a Boolean function that evaluated to True if the boundary of an open subset was hit by the input token colours. Without losing generality, the transition guard is now defined to be the open subset itself.

[4] In earlier SDCPN definitions, the transition delay was defined as a probability distribution function that made use of an integrable transition delay rate. Without losing generality, the transition delay is now defined to be the delay rate itself.

the number of tokens initially in place P_i, $i = 1, \dots |\mathcal{P}|$, and C_0 is a $\sum_{i=1}^{|\mathcal{P}|} M_{i,0} n(P_i)$-dimensional Euclidean-valued random vector which provides the colours of the initial tokens. If $M_{1,0} \geq 1$ then the first $n(P_1)$ components of C_0 are assigned to the first token in P_1. If $M_{1,0} \geq 2$ then the next $n(P_1)$ components of C_0 are assigned to the second token in P_1, etc., until all tokens in P_1 have their assigned colour. The following components of C_0 are assigned to tokens in places $P_2, \dots, P_{|\mathcal{P}|}$ in the same way. If $\mathcal{C}(P) = \mathbb{R}^0$ then the tokens in P get no colour.

Token colour evolution

For each token in each place P for which $\mathcal{C}(P) \neq \mathbb{R}^0$: if the colour of this token is equal to C_0^P at time $t = \tau_0$, and if this token is still in this place at time $t > \tau_0$, then the colour C_t^P of this token equals the probabilistically unique solution of the stochastic differential equation $dC_t^P = \mathcal{V}_P(C_t^P)dt + \mathcal{W}_P(C_t^P)dB_t^{i,P}$ with initial condition $C_{\tau_0}^P = C_0^P$, and with $\{B_t^{i,P}\}$ an $h(P)$-dimensional standard Brownian motion. The first token, if any, in place P uses Brownian motion $\{B_t^{1,P}\}$; the second token, if any, uses $\{B_t^{2,P}\}$, etc. Each token in a place for which $\mathcal{C}(P) = \mathbb{R}^0$ remains without colour.

Transition enabling

A transition T is *pre-enabled* if it has at least one token per incoming ordinary and enabling arc in each of its input places and has no token in places to which it is connected by an inhibitor arc. For each transition T that is pre-enabled at τ_0, consider one token per ordinary and enabling arc in its input places and write $C_t^T \in \mathcal{C}(P(A_{in,OE}(T)))$, $t \geq \tau_0$, as the column vector containing the colours of these tokens; C_t^T evolves through time according to its corresponding token colour functions of the places in $P(A_{in,OE}(T))$. If this vector is not unique (i.e., if one input place contains several tokens per arc), all possible such vectors are executed in parallel. Hence, a transition can be pre-enabled by multiple combinations of input tokens in parallel.
A transition T is *enabled* if it is pre-enabled and a second requirement holds true. For $T \in \mathcal{T}_I$, the second requirement automatically holds true at the time of pre-enabling. For $T \in \mathcal{T}_G$, the second requirement holds true when $C_t^T \in \partial\mathcal{G}_T$. For $T \in \mathcal{T}_D$, the second requirement holds true at $t = \tau_0 + \sigma_1^T$, where σ_1^T is generated from a probability distribution function $D_T(t - \tau_0) = 1 - \exp(-\int_{\tau_0}^t \mathcal{D}_T(C_s^T)ds)$, i.e. $\sigma_1^T = D_T^{inv}(U)$, where D_T^{inv} is the inverse of $D_T(t - \tau_0)$ defined by $D_T^{inv}(u) = \inf\{t - \tau_0 \mid \exp(-\int_{\tau_0}^t \mathcal{D}_T(C_s^T)ds) \leq u\}$, with $\inf\{\ \} = +\infty$. Each delay transition uses one new uniform random variable $U \sim U[0, 1]$ (per vector of input tokens) each time it becomes pre-enabled to determine its time of enabling.
In the case of competing enablings, the following rules apply:

- R0 The firing of an immediate transition has priority over the firing of a guard or a delay transition.
- R1 If one transition becomes enabled by two or more sets of input tokens at exactly the same time, and the firing of any one set will not disable one or more other sets, then it will fire these sets of tokens independently, at the same time.
- R2 If one transition becomes enabled by two or more sets of input tokens at exactly the same time, and the firing of any one set disables one or more other sets, then the set that is fired is selected randomly, with the same probability for each set.
- R3 If two or more transitions become enabled at exactly the same time and the firing of any one transition will not disable the other transitions, then they will fire at the same time.

R4 If two or more transitions become enabled at exactly the same time and the firing of any one transition disables some other transitions, then each combination of transitions that can fire independently without leaving enabled transitions gets the same probability of firing.

By these rules and their combinations, if a transition is enabled in a particular set of tokens, then it is either fired or it is disabled (in this set of tokens) by the firing of another transition.

Transition firing

If T is enabled, suppose this occurs at time τ_1 and in a particular vector of token colours $C^T_{\tau_1}$, it removes one token per arc in $A_{in,O}(T)$ corresponding with $C^T_{\tau_1}$ from each of its input places (i.e. tokens are not removed along enabling arcs). Next, T produces zero or one token along each output arc: If $(e^T_{\tau_1}, a^T_{\tau_1})$ is a random hybrid vector generated from probability measure $\mathcal{F}_T(\cdot; C^T_{\tau_1})$, then vector $e^T_{\tau_1} \in \{0,1\}^{|A_{out}(T)|}$ is an $|A_{out}(T)|$-dimensional vector of zeros and ones, where the ith vector element corresponds with the ith outgoing arc of transition T. An output place gets a token iff it is connected to an arc that corresponds with a vector element 1. Moreover, $a^T_{\tau_1} \in \mathcal{C}(P(A_{out}(T)))$ specifies the colours of the produced tokens, i.e. if the first 1 in $e^T_{\tau_1}$ corresponds with an arc from T to P_j, then the first $n(P_j)$ elements in vector $a^T_{\tau_1}$ are assigned to the token produced in output place P_j. The remaining elements in $a^T_{\tau_1}$ are assigned to other tokens in the same way. The random hybrid vector from $\mathcal{F}_T(\cdot; C^T_{\tau_1})$ is characterised by defining the inverse of $\mathcal{F}_T(\cdot; C^T_{\tau_1})$ as a measurable function $\mathcal{F}^{inv}_T : [0,1] \times \mathcal{C}(P(A_{in,OE}(T))) \to \{0,1\}^{|A_{out}(T)|} \times \mathcal{C}(P(A_{out}(T)))$ such that $\mu_L\{u \mid \mathcal{F}^{inv}_T(u,c) \in H\} = \mathcal{F}_T(H;c)$ for H in the Borel set of $\{0,1\}^{|A_{out}(T)|} \times \mathcal{C}(P(A_{out}(T)))$ and μ_L is the Lebesgue measure. Then $(e^T_{\tau_1}, a^T_{\tau_1}) = \mathcal{F}^{inv}_T(U, C^T_{\tau_1})$. Each firing transition uses one new uniform random variable $U \sim U[0,1]$ per firing to determine its output tokens.

Execution from first transition firing onwards

At $t = \tau_1$, zero or more transitions are pre-enabled (if this number is zero, no transitions will fire anymore). If these include immediate transitions, then these are fired without delay, but with use of rules R0–R4. If after this, still immediate transitions are enabled, then these are also fired, and so forth, until no more immediate transitions are enabled. Each of the immediate transitions that fire uses their firing measure and one uniform random variable (per firing) to determine the number and colours of their output tokens. Next, the SDCPN is executed in the same way as described above for the situation from τ_0 onwards.

In order to keep track of the identity of individual tokens, the tokens in a place are ordered according to the time at which they entered the place, or, if several tokens are produced for one place at the same time, according to the order within the set of arcs $\mathcal{A} = \{A_1, \ldots, A_{|\mathcal{A}|}\}$ along which these tokens were produced (the firing measure produces zero or one token along each output arc). If due to rule R1, a transition fires two or more tokens along one arc at the same time, their assigned order is according to the colours they have (smallest colour first). If under these conditions, two tokens have exactly the same colour, they are indistinguishable and the marking will not be dependent on their order.

2.3 SDCPN stochastic process

The marking of the SDCPN is given by the numbers of tokens in the places and the associated colour values of these tokens. Due to the uniquely defined order of the tokens, the marking is unique except possibly when one or more transitions fire (particularly, immediate transitions

fire without delay hence a sequence of immediate transitions firing will generate a sequence of markings at the same time instant). The SDCPN marking at each time instant can be mapped to a probabilistically unique SDCPN stochastic process $\{M_t, C_t\}$ as follows: For any $t \geq \tau_0$, let a token distribution be characterised by the vector $M'_t = (M'_{1,t}, \ldots, M'_{|\mathcal{P}|,t})$, where $M'_{i,t} \in \mathbb{N}$ denotes the number of tokens in place P_i at time t and $1, \ldots, |\mathcal{P}|$ refers to a unique ordering of places adopted for SDCPN. At times $t \in (\tau_{k-1}, \tau_k)$ when no transition fires, the token distribution is unique and the SDCPN discrete process state M_t is defined to be equal to M'_t. The associated colours of these tokens are gathered in a column vector C_t which first contains all colours of tokens in place P_1, next (i.e. below it) all colours of tokens in place P_2, etc, until place $P_{|\mathcal{P}|}$, where $1, \ldots, |\mathcal{P}|$ refers to a unique ordering of places adopted for SDCPN. Within a place the colours of the tokens are ordered according to the unique ordering of tokens within their place defined for SDCPN (see under SDCPN execution above).
If at time $t = \tau_k$ one or more transitions fire, then the set of applicable token distributions is collected in $\widetilde{M}_{\tau_k} = \{M'_{\tau_k} \mid M'_{\tau_k}$ is a token distribution at time $\tau_k\}$, and the SDCPN discrete process state at time τ_k is defined by $M_{\tau_k} = \{M'_{\tau_k} \mid M'_{\tau_k} \in \widetilde{M}_{\tau_k}$ and no transitions are enabled in $M'_{\tau_k}\}$. In other words, M_{τ_k} is defined to be the token distribution that occurs after all transitions that fire at time τ_k have been fired. The associated colours of these tokens are gathered in a column vector C_{τ_k} in the same way as described above. This construction ensures that the process $\{M_t, C_t\}$ has limits from the left and is continuous from the right, i.e., it satisfies the càdlàg property. If at a time t when one or more transitions fire, the process $\{M_t\}$ jumps to the same value again, and only C_t makes a jump, then the càdlàg property for $\{C_t\}$ (hence for $\{M_t, C_t\}$) is still maintained due to the timing construction of $\{M_t\}$ above and the direct coupling of $\{C_t\}$ with $\{M_t\}$.

3. GSHS

This section presents, following (Bujorianu & Lygeros, 2006), a definition of *general stochastic hybrid system* (GSHS) and its execution.

Definition 3.1 (General stochastic hybrid system)**.** *A GSHS is an automaton* ($\mathbf{K}$, d, $\mathcal{X}$, f, g, Init, λ, Q)*, where*

- $\mathbf{K}$ *is a countable set.*
- $d : \mathbf{K} \to \mathbb{N}$ *maps each* $\theta \in \mathbf{K}$ *to a natural number.*
- $\mathcal{X} : \mathbf{K} \to \{E_\theta; \theta \in \mathbf{K}\}$ *maps each* $\theta \in \mathbf{K}$ *to an open subset* E_θ *of* $\mathbb{R}^{d(\theta)}$*. With this, the hybrid state space is given by* $E \triangleq \{\{\theta\} \times E_\theta; \theta \in \mathbf{K}\}$.
- $f : E \to \{\mathbb{R}^{d(\theta)}; \theta \in \mathbf{K}\}$ *is a vector field.*
- $g : E \to \{\mathbb{R}^{d(\theta) \times h}; \theta \in \mathbf{K}\}$ *is a matrix field, with* $h \in \mathbb{N}$.
- Init: $\mathcal{B}(E) \to [0, 1]$ *is an initial probability measure, with* $\mathcal{B}(E)$ *the Borel* σ*-algebra on* E.
- $\lambda : E \to \mathbb{R}^+$ *is a jump rate function.*
- $Q : \mathcal{B}(E) \times (E \cup \partial E) \to [0, 1]$ *is a GSHS transition measure, where* $\partial E \triangleq \{\{\theta\} \times \partial E_\theta; \theta \in \mathbf{K}\}$ *is the boundary of* E*, in which* ∂E_θ *is the boundary of* E_θ.

Definition 3.2 (GSHS execution)**.** *A stochastic process* $\{\theta_t, X_t\}$ *is called a* GSHS execution *if there exists a sequence of stopping times* $0 = \tau_0 < \tau_1 < \tau_2 \cdots$ *such that for each* $k \in \mathbb{N}$*:*

- (θ_0, X_0) *is an* E*-valued random variable extracted according to probability measure* Init.

- *For $t \in [\tau_k, \tau_{k+1})$, $\theta_t = \theta_{\tau_k}$ and $X_t = X_t^k$, where for $t \geq \tau_k$, X_t^k is a solution of the stochastic differential equation $dX_t^k = f(\theta_{\tau_k}, X_t^k)dt + g(\theta_{\tau_k}, X_t^k)dB_t^{\theta_{\tau_k}}$ with initial condition $X_{\tau_k}^k = X_{\tau_k}$, and where $\{B_t^\theta\}$ is h-dimensional standard Brownian motion for each $\theta \in \mathbf{K}$.*
- *$\tau_{k+1} = \tau_k + \sigma_k$, where σ_k is chosen according to a survivor function given by $F(t) = \mathbf{1}_{(t<\tau^*)} \exp(-\int_0^t \lambda(\theta, X_s^k)ds)$. Here, $\tau^* = \inf\{t > \tau_k \mid X_t^k \in \partial E_{\theta_{\tau_k}}\}$ and $\mathbf{1}$ is indicator function.*
- *The probability distribution of $(\theta_{\tau_{k+1}}, X_{\tau_{k+1}})$, i.e. the hybrid state right after the jump, is governed by the law $Q(\cdot; (\theta_{\tau_k}, X_{\tau_{k+1}-}))$.*

(Bujorianu & Lygeros, 2006) show that under assumptions G1-G4 below, a GSHS execution is a strong Markov Process and has the càdlàg property (right continuous with left hand limits).

G1 $f(\theta, \cdot)$ and $g(\theta, \cdot)$ are Lipschitz continuous and bounded. This yields that for each initial state (θ, x) at initial time τ there exists a pathwise unique solution X_t to $dX_t = f(\theta, X_t)dt + g(\theta, X_t)dB_t$, where $\{B_t\}$ is h-dimensional standard Brownian motion.

G2 $\lambda : E \to \mathbb{R}^+$ is a measurable function such that for all $\xi \in E$, there is $\epsilon(\xi) > 0$ such that $t \to \lambda(\theta_t, X_t)$ is integrable on $[0, \epsilon(\xi))$.

G3 For each fixed $A \in \mathcal{B}(E)$, the map $\xi \to Q(A; \xi)$ is measurable and for any $(\theta, x) \in E \cup \partial E$, $Q(\cdot; \theta, x)$ is a probability measure.

G4 If $N_t = \sum_k \mathbf{1}_{(t \geq \tau_k)}$, then it is assumed that for every starting point (θ, x) and for all $t \in \mathbb{R}^+$, $\mathbb{E}N_t < \infty$. This means, there will be a finite number of jumps in finite time.

4. HSDE

This section presents, following (Blom, 2003) and (Blom et al., 2003), a definition of *hybrid stochastic differential equation* (HSDE) and gives conditions under which the HSDE has a pathwise unique solution. This pathwise unique solution is referred to as *HSDE solution process* or GSHP. The basic advantage of using HSDE in defining a GSHP over using GSHS is that with the HSDE approach the spontaneous jump mechanism is explicitly built on an underlying stochastic basis, whereas in GSHS the execution itself imposes an underlying stochastic basis. The differences are further discussed in Section 4.3.
For the HSDE setting we start with a complete stochastic basis $(\Omega, \Im, \mathbb{F}, \mathbb{P}, \mathbb{T})$, in which a complete probability space $(\Omega, \Im, \mathbb{P})$ is equipped with a right-continuous filtration $\mathbb{F} = \{\Im_t\}$ on the positive time line $\mathbb{T} = \mathbb{R}^+$. This stochastic basis is endowed with a probability measure μ_{θ_0, X_0} for the initial state, an independent h-dimensional standard Wiener process $\{W_t\}$ and an independent homogeneous Poisson random measure $p_P(dt, dz)$ on $\mathbb{T} \times \mathbb{R}^{d+1}$.

Definition 4.1 (Hybrid stochastic differential equation)**.** *An HSDE on stochastic basis $(\Omega, \Im, \mathbb{F}, \mathbb{P}, \mathbb{T})$, is defined as a set of equations (1)-(8) in which a collection of elements ($\mathbb{M}$, E, f, g, μ_{θ_0, X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$) appear.*

This section is organised as follows: Section 4.1 explains the elements and the equations (1)-(8) that define HSDE. Section 4.2 shows that under a number of HSDE conditions H1-H8, the HSDE has a pathwise unique solution which is a semi-martingale. Section 4.3 discusses the differences between GSHP as solution of HSDE and GSHP as execution of GSHS.

4.1 HSDE elements and equations

This section presents the elements and equations that define a HSDE on a hybrid state space. The elements ($\mathbb{M}$, E, f, g, μ_{θ_0,X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$) are defined as follows:

- $\mathbb{M} = \{\vartheta_1, \ldots, \vartheta_N\}$ is a finite set, $N \in \mathbb{N}, 1 \leq N < \infty$.
- $E = \{\{\theta\} \times E_\theta; \theta \in \mathbb{M}\}$ is the hybrid state space, where for each $\theta \in \mathbb{M}$, E_θ is an open subset of $\mathbb{R}^n$ with boundary ∂E_θ. The boundary of E is $\partial E = \{\{\theta\} \times \partial E_\theta; \theta \in \mathbb{M}\}$.
- $f : \mathbb{M} \times \mathbb{R}^n \to \mathbb{R}^n$ is a measurable mapping.
- $g : \mathbb{M} \times \mathbb{R}^n \to \mathbb{R}^{n \times h}$ is a measurable mapping.
- $\mu_{\theta_0,X_0} : \Omega \times \mathcal{B}(E) \to [0,1]$ is a probability measure for the initial random variables θ_0, X_0, which are defined on the stochastic basis; μ_{θ_0,X_0} is assumed to be invertible.
- $\Lambda : \mathbb{M} \times \mathbb{R}^n \to [0, \infty)$ is a measurable mapping.
- $\psi : \mathbb{M} \times \mathbb{M} \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$ is a measurable mapping such that $x + \psi(\vartheta, \theta, x, \underline{z}) \in E_\vartheta$ for all $x \in E_\theta$, $\underline{z} \in \mathbb{R}^d$, and $\vartheta, \theta \in \mathbb{M}$.
- $\rho : \mathbb{M} \times \mathbb{M} \times \mathbb{R}^n \to [0, \infty)$ is a measurable mapping such that $\sum_{i=1}^{N} \rho(\vartheta_i, \theta, x) = 1$ for all $\theta \in \mathbb{M}, x \in \mathbb{R}^n$.
- $\mu : \Omega \times \mathbb{R}^d \to [0,1]$ is a probability measure which is assumed to be invertible.
- $p_P : \Omega \times \mathbb{T} \times \mathbb{R}^{d+1} \to \{0,1\}$ is a homogeneous Poisson random measure on the stochastic basis, independent of (θ_0, X_0). The intensity measure of $p_P(dt, dz)$ equals $dt \cdot \mu_L(dz_1) \cdot \mu(d\underline{z})$, where $z = \text{Col}\{z_1, \underline{z}\}$ and μ_L is the Lebesgue measure.
- $W : \Omega \times \mathbb{T} \to \mathbb{R}^h$ such that $\{W_t\}$ is an h-dimensional standard Wiener process on the stochastic basis, and independent of (θ_0, X_0) and p_P.

Using these elements, the HSDE process $\{\theta_t^*, X_t^*\}$ is defined as follows:

$$\theta_t^* = \theta_t^k \text{ for all } t \in [\tau_k^b, \tau_{k+1}^b), k = 0, 1, 2, \ldots \tag{1}$$

$$X_t^* = X_t^k \text{ for all } t \in [\tau_k^b, \tau_{k+1}^b), k = 0, 1, 2, \ldots \tag{2}$$

Hence $\{\theta_t^*, X_t^*\}$ consists of a concatenation of processes $\{\theta_t^k, X_t^k\}$ which are defined by (3)-(8) below. If the system (1)-(8) has a solution in probabilistic sense, then the process $\{\theta_t^*, X_t^*\}$ is referred to as *HSDE solution process* or *GSHP*.

$$d\theta_t^k = \sum_{i=1}^{N} (\vartheta_i - \theta_{t-}^k) p_P(dt, (\Sigma_{i-1}(\theta_{t-}^k, X_{t-}^k), \Sigma_i(\theta_{t-}^k, X_{t-}^k)] \times \mathbb{R}^d) \tag{3}$$

$$dX_t^k = f(\theta_t^k, X_t^k)dt + g(\theta_t^k, X_t^k)dW_t + \int_{\mathbb{R}^d} \psi(\theta_t^k, \theta_{t-}^k, X_{t-}^k, \underline{z}) p_P(dt, (0, \Lambda(\theta_{t-}^k, X_{t-}^k)] \times d\underline{z}) \tag{4}$$

with $\theta_0^0 = \theta_0$, $X_0^0 = X_0$ and with Σ_0 through Σ_N measurable mappings satisfying, for $\theta \in \mathbb{M}$, $\vartheta_j \in \mathbb{M}$, $x \in \mathbb{R}^n$:

$$\Sigma_i(\theta, x) = \begin{cases} \Lambda(\theta, x) \sum_{j=1}^{i} \rho(\vartheta_j, \theta, x) & \text{if } i > 0 \\ 0 & \text{if } i = 0 \end{cases} \tag{5}$$

In addition, for $k = 0, 1, 2, \ldots$, with $\tau_0^b = 0$:

$$\tau_{k+1}^b \triangleq \inf\{t > \tau_k^b \mid (\theta_t^k, X_t^k) \in \partial E\} \tag{6}$$

$$\mathbb{P}\{\theta_{\tau_{k+1}^b}^{k+1} = \vartheta, X_{\tau_{k+1}^b}^{k+1} \in A \mid \theta_{\tau_{k+1}^b -}^k = \theta, X_{\tau_{k+1}^b -}^k = x\} = Q(\{\vartheta\} \times A; \theta, x) \tag{7}$$

for $A \in \mathcal{B}(\mathbb{R}^n)$, where Q is given by

$$Q(\{\vartheta\} \times A; \theta, x) = \rho(\vartheta, \theta, x) \int_{\mathbb{R}^d} \mathbf{1}_A(x + \psi(\vartheta, \theta, x, \underline{z}))\mu(d\underline{z}) \tag{8}$$

4.2 HSDE solution

This subsection shows that under a set of sufficient conditions H1-H8, the HSDE (1)-(8) has a pathwise unique solution. Note that the existence of a pathwise unique solution guarantees the existence of a unique solution in probabilistic sense.

Proposition 4.1. *Let conditions H1-H8 below hold true. Let* $(\theta_0^*(\omega), X_0^*(\omega)) = (\theta_0, X_0) \in E$ *for all* ω*. Then for every initial condition* (θ_0, X_0)*, (1)-(8) has a pathwise unique solution* $\{\theta_t^*, X_t^*\}$ *which is càdlàg and adapted and is a semi-martingale assuming values in the hybrid state space E.*

H1 For all $\theta \in \mathbb{M}$ there exists a constant $K(\theta)$ such that for all $x \in \mathbb{R}^n$, $|f(\theta,x)|^2 + \|g(\theta,x))\|^2 \leq K(\theta)(1+|x|^2)$, where $|a|^2 = \sum_i (a_i)^2$ and $||b||^2 = \sum_{i,j}(b_{ij})^2$.

H2 For all $r \in \mathbb{N}$ and for all $\theta \in \mathbb{M}$ there exists a constant $L_r(\theta)$ such that for all x and y in the ball $B_r = \{z \in \mathbb{R}^n \mid |z| \leq r+1\}$, $|f(\theta,x) - f(\theta,y)|^2 + \|g(\theta,x) - g(\theta,y)\|^2 \leq L_r(\theta)|x-y|^2$.

H3 For each $\theta \in \mathbb{M}$, the mapping $\Lambda(\theta,\cdot) : \mathbb{R}^n \to [0,\infty)$ is continuous and bounded, with upper bound a constant C_Λ.

H4 For all $(\theta,\vartheta) \in \mathbb{M}^2$, the mapping $\rho(\vartheta,\theta,\cdot) : \mathbb{R}^n \to [0,\infty)$ is continuous.

H5 For all $r \in \mathbb{N}$ there exists a constant $M_r(\theta)$ such that

$$\sup_{|x|\leq r} \int_{\mathbb{R}^d} |\psi(\vartheta,\theta,x,\underline{z})|\mu(d\underline{z}) \leq M_r(\theta), \text{ for all } \vartheta,\theta \in \mathbb{M}$$

H6 $|\psi(\theta,\theta,x,\underline{z})| = 0$ or > 1 for all $\theta \in \mathbb{M}$, $x \in \mathbb{R}^n$, $\underline{z} \in \mathbb{R}^d$

H7 $\{(\theta_t^*, X_t^*)\}$ hits the boundary ∂E a finite number of times on any finite time interval

H8 $|\vartheta_i - \vartheta_j| > 1$ for $i \neq j$, with $|\cdot|$ a suitable metric well defined on $\mathbb{M}$.

(Blom, 2003) has used (Lepeltier & Marchal, 1976) to prove a version of Proposition 4.1 where $E = \mathbb{M} \times \mathbb{R}^n$, i.e. there are no boundaries with instantaneous jumps. Subsequently, (Blom et al., 2003) have proven the proposition under H1-H8 and the additional condition that $\{\tau_k^b\}$ is a sequence of predictable stopping times. (Krystul, 2006; Krystul & Blom, 2005) have shown that this additional condition can be removed. An overview of various HSDE versions is given in (Krystul et al., 2007).

4.3 Discussion of HSDE versus GSHS

HSDE and GSHS have a lot of similarities. Both concatenate different solutions of SDEs with hybrid jumps at each moment of switching to another SDE. Hence the differences are of a rather technical nature. This section collects these technical differences between GSHS and its GSHP execution, versus HSDE and its GSHP solution:

1. For GSHS, the discrete state space is a countable space of discrete variables. For HSDE, the discrete state space is a finite set.
2. For GSHS, the continuous state is Euclidean with a dimension dependent on θ. For HSDE, the continuous state is Euclidean with constant dimension n.
3. The times of spontaneous jump of the GSHS execution are driven by a survivor function which imposes a stochastic basis. For HSDE, the times of spontaneous jumps are driven by a Poisson random measure endowed upon a given stochastic basis.
4. For GSHS, the size of jump is driven by a transition measure Q. For HSDE, the jump size is determined by probability measure μ and measurable mappings ψ and ρ.

5. GSHS involves $|\mathbf{K}|$ Brownian motions. HSDE involves one Wiener process only.
6. For GSHS, the drift and diffusion coefficient are assumed (globally) Lipschitz and bounded. For HSDE, the drift and dissusion coefficient are locally Lipschitz and are allowed to grow with the continuous state.

For 1) and 2), GSHS has as advantage of being more general than HSDE. HSDE however has significant advantages regarding issues 3)-6): Regarding 3)-5), HSDE has the advantage that this allows to establish the *semi-martingale property*. Regarding 6), HSDE removes the particular restriction of GSHS which excludes *jump linear systems*.

5. SDCPN, GSHS and HSDE are bisimilar

This section shows that for each SDCPN there exists a GSHS which is bisimular, and there exists a HSDE which is bisimular. This is shown in the four theorems below.

Theorem 5.1. *Consider an arbitrary GSHS* ($\mathbf{K}$, d, $\mathcal{X}$, f, g, Init, λ, Q) *with a finite domain* $\mathbf{K}$. *If for each* θ *and initial value* X_0, *the stochastic differential equation* $dX_t = f(\theta, X_t)dt + g(\theta, X_t)dB_t$ *has a unique solution in probabilistic sense, then this GSHS can be mapped into an SDCPN* ($\mathcal{P}$, $\mathcal{T}$, $\mathcal{A}$, $\mathcal{N}$, $\mathcal{S}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{V}$, $\mathcal{W}$, $\mathcal{G}$, $\mathcal{D}$, $\mathcal{F}$) *satisfying R0-R4. If the resulting SDCPN is executed on a probability space endowed with standard Brownian motion (one for each place), then the resulting SDCPN process and the GSHS execution are probabilistically equivalent.*

Proof. See (Everdij & Blom, 2006). □

Theorem 5.2. *Consider an arbitrary SDCPN* ($\mathcal{P}$, $\mathcal{T}$, $\mathcal{A}$, $\mathcal{N}$, $\mathcal{S}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{V}$, $\mathcal{W}$, $\mathcal{G}$, $\mathcal{D}$, $\mathcal{F}$) *satisfying R0-R4. If in the initial marking no immediate transition is enabled, and if the number of tokens remains finite for* $t \to \infty$, *then this SDCPN can be mapped into a GSHS* ($\mathbf{K}$, d, $\mathcal{X}$, f, g, Init, λ, Q). *If the original SDCPN is executed on a probability space endowed with Brownian motion (one for each place) then the resulting GSHS execution and the SDCPN process are probabilistically equivalent.*

Proof. See (Everdij & Blom, 2006). □

Theorem 5.3 (HSDE into SDCPN)**.** *Consider an arbitrary HSDE (1)-(8) with elements* ($\mathbb{M}$, E, f, g, μ_{θ_0,X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$). *If for each* θ *the stochastic differential equation* $dX_t = f(\theta, X_t)dt + g(\theta, X_t)dW_t$ *has a unique solution in probabilistic sense and if* Λ *is bounded, then the elements of this HSDE can be mapped into an SDCPN* ($\mathcal{P}$, $\mathcal{T}$, $\mathcal{A}$, $\mathcal{N}$, $\mathcal{S}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{V}$, $\mathcal{W}$, $\mathcal{G}$, $\mathcal{D}$, $\mathcal{F}$) *satisfying R0–R4. If the resulting SDCPN is executed on a probability space endowed with sequences of standard Brownian motions (one sequence for each place), then the resulting SDCPN process and the HSDE solution process are probabilistically equivalent.*

Proof. See Appendix A. □

Theorem 5.4 (SDCPN into HSDE)**.** *Consider an arbitrary SDCPN* ($\mathcal{P}$, $\mathcal{T}$, $\mathcal{A}$, $\mathcal{N}$, $\mathcal{S}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{V}$, $\mathcal{W}$, $\mathcal{G}$, $\mathcal{D}$, $\mathcal{F}$) *satisfying R0–R4. If in the initial marking no immediate transition is enabled, if the delay rates* $\mathcal{D}_T$ *are bounded, and if the number of tokens remains finite for* $t \to \infty$, *then this SDCPN can be mapped into a HSDE (1)-(8) with elements* ($\mathbb{M}$, E, f, g, μ_{θ_0,X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$). *If the original SDCPN is executed on a probability space which is endowed with sequences of standard Brownian motions (one sequence for each place), then the resulting HSDE solution process and the SDCPN process are probabilistically equivalent.*

Proof. See Appendix B. □

Theorems 5.1 and 5.2 imply that SDCPN and GSHS are bisimilar. Theorems 5.3 and 5.4 imply that SDCPN and HSDE are bisimilar. The implications are that GSHS and HSDE are also bisimilar and that the strengths of all three formalisms come within reach of each other. The use of this bisimilarity is illustrated by an example in the following two sections.

6. SDCPN example

To illustrate the advantages of SDCPN when modelling a complex system, consider a simplified model of the evolution of an aircraft in one sector of airspace. The deviation of this aircraft from its intended path is affected by its engine system and its navigation system. Each of these aircraft systems can be in either *Working* (functioning properly) or *Not working* (operating in some failure mode). Both systems switch between these modes independently and with exponentially distributed sojourn times, with finite rates δ_3 (engine repaired), δ_4 (engine fails), δ_5 (navigation repaired) and δ_6 (navigation fails), respectively. If both systems are *Working*, the aircraft evolves in *Nominal* mode and the position Y_t and velocity S_t of the aircraft are determined by $dX_t = \mathcal{V}_1(X_t)dt + \mathcal{W}_1 dW_t$, where $X_t = (Y_t, S_t)'$. If either one, or both, of the systems is *Not working*, the aircraft evolves in *Non-nominal* mode and the position and velocity of the aircraft are determined by $dX_t = \mathcal{V}_2(X_t)dt + \mathcal{W}_2 dW_t$. The factors $\mathcal{W}_1$ and $\mathcal{W}_2$ are determined by wind fluctuations. Initially, the aircraft has position Y_0 and velocity S_0, while both its systems are *Working*. The evaluation of this process may be stopped when the aircraft has *Landed*, i.e. its vertical position and velocity are equal to zero.

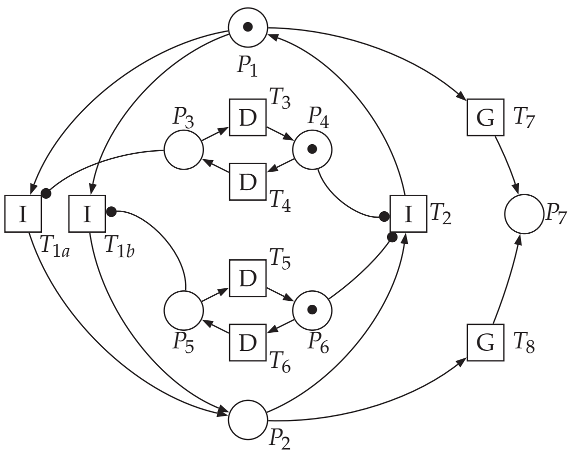


Fig. 4. SDCPN graph for the aircraft evolution example

Fig. 4 shows the SDCPN graph for this example, where,

- P_1 denotes aircraft evolution *Nominal*, i.e. evolution is according to $\mathcal{V}_1$ and $\mathcal{W}_1$.
- P_2 denotes aircraft evolution *Non-nominal*, i.e. evolution is according to $\mathcal{V}_2$ and $\mathcal{W}_2$.
- P_3 and P_4 denote engine system *Not working* and *Working*, respectively.
- P_5 and P_6 denote navigation system *Not working* and *Working*, respectively.
- P_7 denotes the aircraft has landed.
- T_{1a} and T_{1b} denote a transition of aircraft evolution from *Nominal* to *Non-nominal*, due to engine system or navigation system *Not working*, respectively.

- T_2 denotes a transition of aircraft evolution from *Non-nominal* to *Nominal*, due to engine system and navigation system both *Working* again.
- T_3 through T_6 denote transitions between *Working* and *Not working* of the engine and navigation systems.
- T_7 and T_8 denote transitions of the aircraft landing.

The graph in Fig. 4 completely defines SDCPN elements $\mathcal{P}$, $\mathcal{T}$, $\mathcal{A}$ and $\mathcal{N}$, where $\mathcal{T}_G = \{T_7, T_8\}$, $\mathcal{T}_D = \{T_3, T_4, T_5, T_6\}$ and $\mathcal{T}_I = \{T_{1a}, T_{1b}, T_2\}$. The other SDCPN elements are specified below:

$\mathcal{S}$: Two colour types are defined; $\mathcal{S} = \{\mathbb{R}^0, \mathbb{R}^6\}$.

$\mathcal{C}$: $\mathcal{C}(P_1) = \mathcal{C}(P_2) = \mathcal{C}(P_7) = \mathbb{R}^6$, i.e. tokens in P_1, P_2 and P_7 have colours in $\mathbb{R}^6$; the colour components model the 3-dimensional position and 3-dimensional velocity of the aircraft. $\mathcal{C}(P_3) = \mathcal{C}(P_4) = \mathcal{C}(P_5) = \mathcal{C}(P_6) = \mathbb{R}^0 \triangleq \emptyset$.

$\mathcal{I}$: Place P_1 initially has a token with colour $X_0 = (Y_0, S_0)'$, with $Y_0 \in \mathbb{R}^2 \times (0, \infty)$ and $S_0 \in \mathbb{R}^3 \setminus \text{Col}\{0, 0, 0\}$. Places P_4 and P_6 initially each have a token without colour.

$\mathcal{V}$, $\mathcal{W}$: The token colour functions for places P_1, P_2 and P_7 are determined by $(\mathcal{V}_1, \mathcal{W}_1)$, $(\mathcal{V}_2, \mathcal{W}_2)$, and $(\mathcal{V}_7, \mathcal{W}_7)$, respectively, where $(\mathcal{V}_7, \mathcal{W}_7) = (0, 0)$. For places P_3 – P_6 there is no token colour function.

$\mathcal{G}$: Transitions T_7 and T_8 have a guard defined by $\mathcal{G}_{T_7} = \mathcal{G}_{T_8} = \mathbb{R}^2 \times (0, \infty) \times \mathbb{R}^2 \times (0, \infty)$.

$\mathcal{D}$: The jump rates for transitions T_3, T_4, T_5 and T_6 are $\mathcal{D}_{T_3}(\cdot) = \delta_3$, $\mathcal{D}_{T_4}(\cdot) = \delta_4$, $\mathcal{D}_{T_5}(\cdot) = \delta_5$ and $\mathcal{D}_{T_6}(\cdot) = \delta_6$.

$\mathcal{F}$: Each transition has a unique output place, to which it fires a token with a colour (if applicable) equal to the colour of the token removed.

7. Mapping of SDCPN example to HSDE and GSHS

Next we transform the SDCPN of Section 6 into an HSDE. The first step is to construct the state space $\mathbb{M}$ for the HSDE discrete process $\{\theta_t\}$. This is done by identifying the SDCPN *reachability graph*. Nodes in the reachability graph provide the number of tokens in each of the SDCPN places. Arrows connect these nodes as they represent transitions firing. The SDCPN of Fig. 4 has seven places hence the reachability graph for this example has elements that are vectors of length 7. These nodes, excluding the nodes that enable immediate transitions, form the HSDE discrete state space.

The reachability graph is shown in Fig. 5, with nodes that form the HSDE discrete state space in Bold typeface, i.e. $\mathbb{M} = \{V_1, \ldots, V_8\}$, with $V_1 = (1, 0, 0, 1, 0, 1, 0)$, $V_2 = (0, 1, 1, 0, 0, 1, 0)$, $V_3 = (0, 1, 1, 0, 1, 0, 0)$, $V_4 = (0, 1, 0, 1, 1, 0, 0)$, $V_5 = (0, 0, 0, 1, 0, 1, 1)$, $V_6 = (0, 0, 1, 0, 0, 1, 1)$, $V_7 = (0, 0, 1, 0, 1, 0, 1)$, $V_8 = (0, 0, 0, 1, 1, 0, 1)$. Since initially there is a token in places P_1, P_4 and P_6, the HSDE initial mode equals $\theta_0 = V_1 = (1, 0, 0, 1, 0, 1, 0)$. The HSDE initial continuous state value equals the vector containing the initial colours of all initial tokens. Since the initial colour of the token in Place P_1 equals X_0, and the tokens in places P_4 and P_6 have no colour, the HSDE initial continuous state value equals $\text{Col}\{X_0, \emptyset, \emptyset\} = X_0$. The HSDE drift coefficient f is given by $f(\theta, \cdot) = \mathcal{V}_1(\cdot)$ for $\theta = V_1$, $f(\theta, \cdot) = \mathcal{V}_2(\cdot)$ for $\theta \in \{V_2, V_3, V_4\}$, and $f(\theta, \cdot) = 0$ otherwise. For the diffusion coefficient, $g(\theta, \cdot) = \mathcal{W}_1$ for $\theta = V_1$, $g(\theta, \cdot) = \mathcal{W}_2$ for $\theta \in \{V_2, V_3, V_4\}$, and $g(\theta, \cdot) = 0$ otherwise. The hybrid state space is given by $E = \{\{\theta\} \times E_\theta; \theta \in \mathbb{M}\}$, where for $\theta \in \{V_1, V_2, V_3, V_4\}$: $E_\theta = \mathbb{R}^2 \times (0, \infty) \times \mathbb{R}^2 \times (0, \infty)$ and for $\theta \in \{V_5, V_6, V_7, V_8\}$: $E_\theta = \mathbb{R}^6$. Always two delay transitions are pre-enabled: either T_3 or T_4 and either T_5 or T_6. This yields $\Lambda(V_1, \cdot) = \Lambda(V_5, \cdot) = \delta_4 + \delta_6$, $\Lambda(V_2, \cdot) = \Lambda(V_6, \cdot) = \delta_3 + \delta_6$, $\Lambda(V_3, \cdot) = \Lambda(V_7, \cdot) = \delta_3 + \delta_5$,

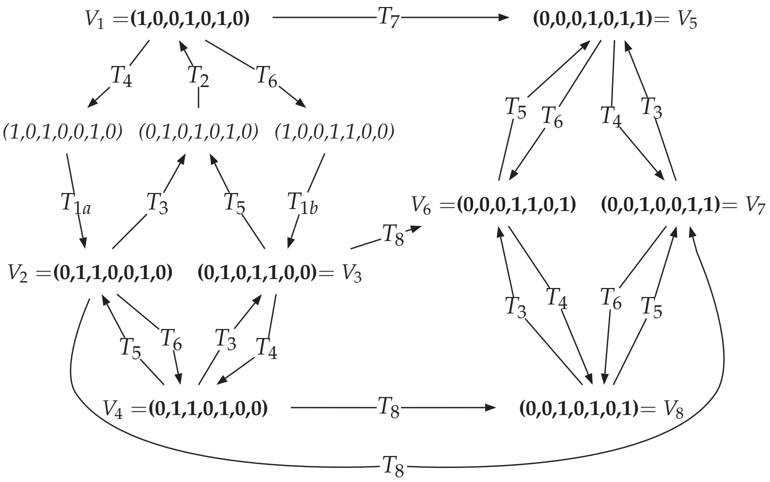


Fig. 5. Reachability graph for the SDCPN of Fig. 4. The nodes in bold type face correspond with the elements of the HSDE discrete state space $\mathbb{M}$.

$\Lambda(V_4,\cdot) = \Lambda(V_8,\cdot) = \delta_4 + \delta_5$. For the determination of elements ψ, ρ and μ, we first construct a probability measure P_Q, by making use of the reachability graph, the sets $\mathcal{D}$, $\mathcal{G}$ and $\mathcal{F}$ and the rules R0–R4. In Table 1, $P_Q(\theta',x';\theta,x) = p$ denotes that if (θ,x) is the value of the HSDE state before the hybrid jump, then, with probability p, (θ',x') is the value of the HSDE state immediately after the jump. Since the continuous valued process jumps to the same value with probability 1, we find that $\psi(V^i,V^j,x,\underline{z}) = 0$ for all V^i, V^j, x, $\underline{z}$. Moreover, $\rho(V^i,V^j,x) = P_Q(V^i,x,V^j,x)$ and μ may be any given invertible probability measure.

Table 1. Example probability measure for size of jump

For $x \notin \partial E_{V_1}$:	$P_Q(V_2,x;V_1,x) = \frac{\delta_4}{\delta_4+\delta_6}$,	$P_Q(V_4,x;V_1,x) = \frac{\delta_6}{\delta_4+\delta_6}$
For $x \in \partial E_{V_1}$:	$P_Q(V_5,x;V_1,x) = 1$	
For $x \notin \partial E_{V_2}$:	$P_Q(V_3,x;V_2,x) = \frac{\delta_6}{\delta_3+\delta_6}$,	$P_Q(V_1,x;V_2,x) = \frac{\delta_3}{\delta_3+\delta_6}$
For $x \in \partial E_{V_2}$:	$P_Q(V_6,x;V_2,x) = 1$	
For $x \notin \partial E_{V_3}$:	$P_Q(V_4,x;V_3,x) = \frac{\delta_3}{\delta_3+\delta_5}$,	$P_Q(V_2,x;V_3,x) = \frac{\delta_5}{\delta_3+\delta_5}$
For $x \in \partial E_{V_3}$:	$P_Q(V_7,x;V_3,x) = 1$	
For $x \notin \partial E_{V_4}$:	$P_Q(V_3,x;V_4,x) = \frac{\delta_4}{\delta_4+\delta_5}$,	$P_Q(V_1,x;V_4,x) = \frac{\delta_5}{\delta_4+\delta_5}$
For $x \in \partial E_{V_4}$:	$P_Q(V_8,x;V_4,x) = 1$	
For all x:	$P_Q(V_6,x;V_5,x) = \frac{\delta_4}{\delta_4+\delta_6}$,	$P_Q(V_8,x;V_5,x) = \frac{\delta_6}{\delta_4+\delta_6}$
For all x:	$P_Q(V_7,x;V_6,x) = \frac{\delta_6}{\delta_3+\delta_6}$,	$P_Q(V_5,x;V_6,x) = \frac{\delta_3}{\delta_3+\delta_6}$
For all x:	$P_Q(V_8,x;V_7,x) = \frac{\delta_3}{\delta_3+\delta_5}$,	$P_Q(V_6,x;V_7,x) = \frac{\delta_5}{\delta_3+\delta_5}$
For all x:	$P_Q(V_7,x;V_8,x) = \frac{\delta_4}{\delta_4+\delta_5}$,	$P_Q(V_5,x;V_8,x) = \frac{\delta_5}{\delta_4+\delta_5}$

With this, the SDCPN of the aircraft evolution example is uniquely mapped to an HSDE. If in addition, we want to make use of the HSDE properties of Proposition 4.1, i.e. the resulting HSDE solution process being adapted and a semi-martingale, we need to make sure that

HSDE conditions H1-H8 are satisfied. It is shown below that they are, under the following sufficient condition D1 for the example SDCPN.

D1 For $P \in \{P_1, P_2\}$, there exist K_P^v, L_P^v, K_P^w and L_P^w such that for all $c, a \in \mathcal{C}(P)$,
$|\mathcal{V}_P(c)|^2 \leq K_P^v(1+|c|^2)$ and $|\mathcal{V}_P(c) - \mathcal{V}_P(a)|^2 \leq L_P^v|c-a|^2$ and
$\|\mathcal{W}_P(c)\|^2 \leq K_P^w(1+|c|^2)$ and $\|\mathcal{W}_P(c) - \mathcal{W}_P(a)\|^2 \leq L_P^w|c-a|^2$.

We verify that under condition D1, HSDE conditions H1-H8 hold true in this example.

H1: From the construction of f and g above we have for $\theta = V_1$: $|f(\theta,x)|^2 + \|g(\theta,x)\|^2 = |\mathcal{V}_1(x)|^2 + \|\mathcal{W}_1(x)\|^2 \leq K_{P_1}^v(1+|x|^2) + K_{P_1}^w(1+|x|^2) = K(\theta)(1+|x|^2)$, with $K(\theta) = (K_{P_1}^v + K_{P_1}^w)$. For $\theta = V_2, V_3, V_4$ the verification is with replacing $\mathcal{V}_1$, $\mathcal{W}_1$ by $\mathcal{V}_2$, $\mathcal{W}_2$.

H2: From the construction of f and g above we have for $\theta = V_1$: $|f(\theta,x) - f(\theta,y)|^2 + \|g(\theta,x) - g(\theta,y)\|^2 = |\mathcal{V}_1(x) - \mathcal{V}_1(y)|^2 + \|\mathcal{W}_1(x) - \mathcal{W}_1(y)\|^2 \leq L_{P_1}^v|x-y|^2 + L_{P_1}^w|x-y|^2 = L_r(\theta)|x-y|^2$ with $L_r(\theta) = L_{P_1}^v + L_{P_1}^w$. For $\theta = V_2, V_3, V_4$ replace $\mathcal{V}_1$, $\mathcal{W}_1$ by $\mathcal{V}_2$, $\mathcal{W}_2$.

H3: Since δ_3–δ_6 are constant, for all θ, $\Lambda(\theta,\cdot)$ is bounded and continuous, with upper bound $C_\Lambda = \max\{\delta_4+\delta_6, \delta_3+\delta_6, \delta_3+\delta_5, \delta_4+\delta_5\}$.

H4: Since for all θ, ϑ, $P_Q(\vartheta, \cdot; \theta, x)$ is constant, we find $\rho(\vartheta, \theta, x) = P_Q(\vartheta, x, \theta, x)$ is continuous.

H5 and H6: These are satisfied due to $\psi(V^i, V^j, x, \underline{z}) = 0$ for all V^i, V^j, x, $\underline{z}$.

H7: This condition holds due to δ_3–δ_6 being finite and the fact that in this SDCPN example, there is no firing sequence of more than one guard transition.

H8: This condition holds for all $V_1, \ldots, V_8$, with metric $|a|^2 = \sum_i (a_i)^2$.

Thanks to this bisimilarity mapping we can now use HSDE tools to analyse the GSHP that is defined by the execution of the SDCPN model for the example.
In (Everdij & Blom, 2008) we showed how the SDCPN for the aircraft evolution example above is mapped to a GSHS. The main difference is that the GSHS transition measure Q is defined by the probability measure P_Q in Table 1 and that GSHS does not use elements ψ, ρ and μ, but apart of these details the differences with the mapping of SDCPN elements into HSDE elements are small. Thanks to this bisimilarity mapping, we can also use the automata framework to analyse the GSHS that is defined by the SDCPN model.

8. Conclusions

In order to combine the compositional specification power of Petri nets with the analysis power of Markov processes, (Malhotra & Trivedi, 1994) and (Muppala et al., 2000) developed a power hierarchy of dependability models. In (Everdij & Blom, 2003; 2005), the power hierarchy was extended with *dynamically coloured Petri nets* (DCPN) and *piecewise deterministic Markov processes* (PDP). In (Everdij & Blom, 2006), this power hierarchy was further extended by *stochastically and dynamically coloured Petri nets* (SDCPN) and *general stochastic hybrid process* (GSHP).
In this chapter the power-hierarchy has been further deepened by studying various ways to develop GSHP. We started in Section 2 by defining SDCPN and the resulting SDCPN process. In Section 3 we studied GSHP as an execution of a *general stochastic hybrid system* (GSHS). In Section 4 we defined GSHP as a solution of a *hybrid stochastic differential equation* (HSDE) and explained the differences between GSHS and HSDE. Next, in Section 5 we showed that GSHS, HSDE and SDCPN are bisimilar. In Sections 6-7, the results were illustrated with an aircraft

evolution example. The bisimilarities between SDCPN, GSHS and HSDE mean that each of them inherits the strengths of the other two formalisms. This has been depicted in Fig. 2 in the introduction. Hence, analysis tools designed for GSHS, HSDE and GSHP and their properties become available for SDCPN. Examples of GSHP properties are convergence in discretisation, existence of limits, existence of event probabilities, strong Markov properties, reachability analysis. Examples of GSHS features are their connection to formal methods in automata theory and optimal control theory. Examples of HSDE features are stochastic analysis tools for semi-martingales. At the same time, numerous SDCPN features such as natural expression of causal dependencies, concurrency and synchronisation mechanism, hierarchical and modular construction, and graphical representation become available when modelling GSHS, HSDE and GSHP through SDCPN. And these complementary advantages of SDCPN, GSHS, HSDE and GSHP perspectives tend to increase with the complexity of the system considered.

9. References

Blom, H. (2003). Stochastic hybrid processes with hybrid jumps, *Proceedings IFAC conference on analysis and design of hybrid system (ADHS), Saint-Malo, Brittany, France*, pp. 361–365.

Blom, H., Bakker, G., Everdij, M. & Van der Park, M. (2003). Stochastic analysis background of accident risk assessment for air traffic management, *Hybridge report*, D2.2. http://hosted.nlr.nl/public/hosted-sites/hybridge/.

Bujorianu, M. & Lygeros, J. (2003). Reachability questions in piecewise deterministic Markov processes, *in* O. Maler & A. Pnueli (eds), *Proceedings 6th international workshop on hybrid systems: computation and control (HSCC), Prague, Czech Republic*, Vol. 2623 of *Lecture notes in computer science (LNCS)*, Springer, pp. 126–140.

Bujorianu, M. & Lygeros, J. (2004). General stochastic hybrid systems: modelling and optimal control, *Proceedings 43rd conference on decision and control (CDC), Nassau, Bahamas*.

Bujorianu, M. & Lygeros, J. (2006). Toward a general theory of stochastic hybrid systems, *in* H. Blom & J. Lygeros (eds), *Stochastic hybrid systems: theory and safety critical applications*, Vol. 337 of *Lectures notes in control and information sciences (LNCIS)*, Springer, pp. 3–30.

Bujorianu, M., Lygeros, J. & Bujorianu, M. (2005). Different approaches on bisimulation for stochastic hybrid systems, *in* M. Morari & L. Thiele (eds), *Proceedings 8th international workshop on hybrid systems: computation and control (HSCC), Zürich, Switzerland*, Vol. 3414 of *Lecture notes in computer science (LNCS)*, pp. 198–214.

David, R. & Alla, H. (1994). Petri nets for modeling of dynamic systems - a survey, *Automatica* **30**(2): 175–202.

Davis, M. (1984). Piecewise deterministic Markov processes: a general class of non-diffusion stochastic models, *Journal royal statistical society (B)* **46**(3): 353–388.

Davis, M. (1993). *Markov models and optimization*, Vol. 49 of *Monographs on statistics and applied probability*, Chapman and Hall, London.

Elliott, R. (1982). *Stochastic calculus and applications*, Vol. 18 of *Applications of mathematics: Stochastic modelling and applied probability*, Springer-Verlag.

Elliott, R., Aggoun, L. & Moore, J. (1995). *Hidden Markov models: estimation and control*, Vol. 29 of *Applications of mathematics: stochastic modelling and applied probability*, Springer-Verlag.

Ethier, S. & Kurtz, T. (1986). *Markov processes, characterization and convergence*, Wiley series in probability and mathematical statistics, John Wiley & Sons, New York.

Everdij, M. & Blom, H. (2003). Petri nets and hybrid state Markov processes in a power-hierarchy of dependability models, *Proceedings IFAC conference on analysis and design of hybrid system (ADHS), Saint-Malo, Brittany, France*, pp. 355–360.

Everdij, M. & Blom, H. (2005). Piecewise deterministic Markov processes represented by dynamically coloured Petri nets, *in* S. Jacka (ed.), *Stochastics: an international journal of probability and stochastic processes*, Vol. 77, number 1, Taylor & Francis, pp. 1–29.

Everdij, M. & Blom, H. (2006). Hybrid Petri nets with diffusion that have into-mappings with generalised stochastic hybrid processes, *in* H. Blom & J. Lygeros (eds), *Stochastic hybrid systems: theory and safety critical applications*, Vol. 337 of *Lectures notes in control and information sciences (LNCIS)*, Springer, pp. 31–63.

Everdij, M. & Blom, H. (2008). Enhancing hybrid state Petri nets with the analysis power of stochastic hybrid processes, *in* B. Lennartson, M. Fabian, K. Åkesson, A. Giua & R. Kumar (eds), *Proceedings 9th International Workshop on Discrete Event Systems (WODES), Goeteborg, Sweden*, pp. 400–405.

Everdij, M., Klompstra, M., Blom, H. & Klein Obbink, B. (2006). Compositional specification of a multi-agent system by stochastically and dynamically coloured Petri nets, *in* H. Blom & J. Lygeros (eds), *Stochastic hybrid systems: theory and safety critical applications*, Vol. 337 of *Lectures notes in control and information sciences (LNCIS)*, Springer, pp. 325–350.

Frehse, G. (2008). PHAVer: algorithmic verification of hybrid systems past HyTech, *International journal on software tools for technology transfer* **10**(3): 263–279.

Giua, A. (1999). Bibliography on hybrid Petri nets. http://bode.diee.unica.it/∼hpn/.

Hu, J., Lygeros, J. & Sastry, S. (2000). Towards a theory of stochastic hybrid systems, *in* N. Lynch & B. Krogh (eds), *Proceedings 3rd international workshop on hybrid systems: computation and control (HSCC), Pittsburgh, Pennsylvania, USA*, Vol. 1790 of *Lecture notes in computer science (LNCS)*, Springer Verlag, pp. 160–173.

Krystul, J. (2006). *Modelling of stochastic hybrid systems with applications to accident risk assessment*, PhD thesis, University of Twente, The Netherlands.

Krystul, J. & Blom, H. (2005). Generalised stochastic hybrid processes as strong solutions of stochastic differential equations, *Hybridge report*, D2.3. http://hosted.nlr.nl/public/hosted-sites/hybridge/.

Krystul, J., Blom, H. & Bagchi, A. (2007). *Stochastic hybrid systems*, number 24 in *Control engineering series*, Taylor and Francis / CRC Press, chapter 2: Stochastic differential equations on hybrid state spaces, pp. 15–45.

Kwiatkowska, M., Norman, G. & Parker, D. (2004). Probabilistic symbol model checking with PRISM: a hybrid approach, *International journal on software tools for technology transfer* **6**(2): 128–142.

Labinaz, G., Bayoumi, M. & Rudie, K. (1997). A survey of modeling and control of hybrid systems, *Annual reviews of control* **21**: 79–92.

Lepeltier, J. & Marchal, B. (1976). Problème des martingales et équations différentielles stochastiques associées à un opérateur intégro-différentiel, *Annales de l'Institute Henri Poincaré* **Section B - XII**(1): 43–103.

Malhotra, M. & Trivedi, K. (1994). Power-hierarchy of dependability-model types, *IEEE transactions on reliability* **R-43**(3): 493–502.

Muppala, J., Fricks, R. & Trivedi, K. (2000). Techniques for system dependability evaluation, *in* W. Grasman (ed.), *Computational probability*, Kluwer academic publishers, The Netherlands, pp. 445–480.

Strubbe, S. & Van der Schaft, A. (2004). Semantics, bisimulation and interaction-structures for the CPDP model, *Hybridge report*, D4.3. http://hosted.nlr.nl/public/hosted-sites/hybridge/.

Strubbe, S. & Van der Schaft, A. (2005). Bisimulation for communicating piecewise deterministic Markov processes (CPDPs), *in* M. Morari & L. Thiele (eds), *Proceedings 8th international workshop on hybrid systems: computation and control (HSCC), Zürich, Switzerland*, Vol. 3414 of *Lecture notes in computer science (LNCS)*, pp. 623–639.

Van der Schaft, A. (2004). Equivalence of dynamical systems by bisimulation, *IEEE transactions on automatic control* **49**(12): 2160–2172.

Appendix A: Proof of Theorem 5.3

Consider an arbitrary HSDE (1)-(8) with elements ($\mathbb{M}$, E, f, g, μ_{θ_0,X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$). We assume that the stochastic differential equations defined by f and g have probabilistically unique solutions and that Λ is bounded. First, we characterise SDCPN elements ($\mathcal{P}$, $\mathcal{T}$, $\mathcal{A}$, $\mathcal{N}$, $\mathcal{S}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{V}$, $\mathcal{W}$, $\mathcal{G}$, $\mathcal{D}$, $\mathcal{F}$) in terms of HSDE elements ($\mathbb{M}$, E, f, g, μ_{θ_0,X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$). The thus constructed SDCPN is referred to as SDCPN^{HSDE}. Subsequently, we show that the SDCPN^{HSDE} stochastic process is probabilistically equivalent to the stochastic process defined by the original HSDE.

A.1 Construction of SDCPN^{HSDE} elements

We provide an into-mapping that characterises SDCPN elements ($\mathcal{P}$, $\mathcal{T}$, $\mathcal{A}$, $\mathcal{N}$, $\mathcal{S}$, $\mathcal{C}$, $\mathcal{I}$, $\mathcal{V}$, $\mathcal{W}$, $\mathcal{G}$, $\mathcal{D}$, $\mathcal{F}$) in terms of HSDE elements ($\mathbb{M}$, E, f, g, μ_{θ_0,X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$).

$\mathcal{P} = \{P_\theta; \theta \in \mathbb{M}\}$. Hence, for each $\theta \in \mathbb{M}$, there is one place P_θ. The places are ordered $P_{\vartheta_1}, \ldots, P_{\vartheta_N}$ according to $\mathbb{M} = \{\vartheta_1, \ldots, \vartheta_N\}$.

$\mathcal{T} = \mathcal{T}_G \cup \mathcal{T}_D \cup \mathcal{T}_I$, with $\mathcal{T}_I = \emptyset$, $\mathcal{T}_G = \{T_\theta^G; \theta \in \mathbb{M}\}$, $\mathcal{T}_D = \{T_\theta^D; \theta \in \mathbb{M}\}$. Hence, for each $\theta \in \mathbb{M}$ there is one guard transition T_θ^G and one delay transition T_θ^D.

$\mathcal{A} = \mathcal{A}_O \cup \mathcal{A}_E \cup \mathcal{A}_I$, with $|\mathcal{A}_I| = 0$, $|\mathcal{A}_E| = 0$, and $|\mathcal{A}_O| = 2N + 2N^2$, where $N = |\mathbb{M}|$. Hence, there are no inhibitor arcs or enabling arcs in this SDCPN^{HSDE} constructed, and the number of ordinary arcs is $2N + 2N^2$.

$\mathcal{N}$: The node function maps each arc in $\mathcal{A} = \mathcal{A}_O$ to a pair of nodes. These connected pairs of nodes are: $\{(P_\theta, T_\theta^G); \theta \in \mathbb{M}\} \cup \{(P_\theta, T_\theta^D); \theta \in \mathbb{M}\} \cup \{(T_\theta^G, P_\vartheta); \theta, \vartheta \in \mathbb{M}\} \cup \{(T_\theta^D, P_\vartheta); \theta, \vartheta \in \mathbb{M}\}$. Hence, each place P_θ ($\theta \in \mathbb{M}$) has two outgoing arcs: one to guard transition T_θ^G and one to delay transition T_θ^D. Each transition has N outgoing arcs: one arc to each place in $\mathcal{P}$.

$\mathcal{S} = \{\mathbb{R}^n\}$.

$\mathcal{C}$: For all $\theta \in \mathbb{M}$, $\mathcal{C}(P_\theta) = \mathbb{R}^n$

$\mathcal{I}$: For all $\theta_0 \in \mathbb{M}$ and $X_0 \in \mathcal{C}(P_{\theta_0}) = \mathbb{R}^n$, $\mathcal{I}(M^{\theta_0}, X_0) = \mu_{\theta_0,x_0}(\theta_0, X_0)$, where M^θ is the $|\mathcal{P}|$-dimensional vector that has a one at the element corresponding to place P_θ and zeros elsewhere.

$\mathcal{V}$: For all $\theta \in \mathbb{M}$, $\mathcal{V}_{P_\theta}(\cdot) = f(\theta, \cdot)$.

$\mathcal{W}$: For all $\theta \in \mathbb{M}$, $\mathcal{W}_{P_\theta}(\cdot) = g(\theta, \cdot)$.

$\mathcal{G}$: For all $\theta \in \mathbb{M}$, $\mathcal{G}_{T_\theta^G} = E_\theta$.

$\mathcal{D}$: For all $\theta \in \mathbb{M}$, $\mathcal{D}_{T_\theta^D}(\cdot) = \Lambda(\theta, \cdot)$. Since we assumed that Λ is bounded, e.g. $\Lambda(\theta, \cdot) \leq C_\Lambda$, we find that $\mathcal{D}_{T_\theta^D}(\cdot)$ is bounded as well, and its upperbound is $C_\delta = C_\Lambda$.

$\mathcal{F}$: Define for particular transition T, $e^{\vartheta'}$ as the vector of length N containing a one at the component corresponding with the arc from transition T to place $P_{\vartheta'}$ and zeros elsewhere. Then for all $\theta \in \mathbb{M}$, and for $T \in \{T_\theta^G, T_\theta^D\}$, $\mathcal{F}_T(e^{\vartheta'}, x'; x) = F_T^Q(\vartheta', x'; \theta, x)$, for all $x \in E_\theta \cup \partial E_\theta$, $\vartheta' \in \mathbb{M}$ and $x' \in E_{\vartheta'}$, where F_T^Q is defined through

$$F_T^Q(\{\vartheta'\} \times A'; \theta, x) = \rho(\vartheta', \theta, x) \int_{\mathbb{R}^d} \mathbf{1}_{A'}(x + \psi(\vartheta', \theta, x, \underline{z}))\mu(d\underline{z}) \quad (9)$$

A.2 Probabilistic equivalence

Next, we show that the SDCPN^{HSDE} stochastic process is probabilistically equivalent to the stochastic process defined by the original HSDE. This is done by showing: Equivalence of initial states; Equivalence of continuous evolution until first jump; Equivalence of time of jumps; Equivalence of size of jumps; Equivalence of processes after the first jump.

Equivalence of initial states:

The initial marking of the SDCPN^{HSDE} is defined by $\mathcal{I}(M^{\theta_0}, X_0) = \mu_{\theta_0, X_0}(\theta_0, X_0)$, where M^θ is the N-dimensional vector that has a one at the element corresponding to place P_θ and zeros elsewhere. Therefore, with probability $\mathcal{I}(M^{\theta_0}, X_0)$, at time $t = \tau_0$ there is one token in place P_{θ_0} which has colour X_0. The initial state of the HSDE is (θ_0, X_0) with probability $\mu_{\theta_0, X_0}(\theta_0, X_0)$. Due to the mapping between the places $P_\theta \in \mathcal{P}$ and the modes $\theta \in \mathbb{M}$, the initial states of SDCPN^{HSDE} and HSDE are probabilistically equivalent.

Equivalence of continuous evolution until first jump:

The continuous part of the SDCPN^{HSDE} stochastic process equals the vector that collects all token colours. Since there is only one token in the constructed SDCPN^{HSDE} at all times, this vector equals the colour of this single token. Until the first jump, this colour follows the stochastic differential equation $dC_t^{P_{\theta_0}} = \mathcal{V}_{P_{\theta_0}}(C_t^{P_{\theta_0}})dt + \mathcal{W}_{P_{\theta_0}}(C_t^{P_{\theta_0}})dW_t^{P_{\theta_0}}$ which has probabilistically unique solution $C_t^{P_{\theta_0}}$.

In the original HSDE solution process, the continuous process until the first jump follows stochastic differential equation $dX_t^0 = f(\theta_t^0, X_t^0)dt + g(\theta_t^0, X_t^0)dW_t + \int_{\mathbb{R}^d} \psi(\theta_t^0, \theta_{t-}^0, X_{t-}^0, \underline{z})$ $p_P(dt, (0, \Lambda(\theta_{t-}^0, X_{t-}^0)] \times d\underline{z})$ where $d\theta_t^0 = \sum_{i=0}^N (\vartheta_i - \theta_{t-}^0) p_P(dt, (\Sigma_{i-1}(\theta_{t-}^0, X_{t-}^0), \Sigma_i(\theta_{t-}^0, X_{t-}^0)] \times \mathbb{R}^d)$. Until the first jump, the Poisson terms in the stochastic differential equations above are equal to zero. What remains is: $d\theta_t^0 = 0$ and $dX_t^0 = f(\theta_t^0, X_t^0)dt + g(\theta_t^0, X_t^0)dW_t$, which are assumed to have a probabilistically unique solution θ_t^0 and X_t^0.

Due to equivalence of initial states $M^{\theta_0} \equiv \theta_0$ and $C_0 = X_0$, equivalence of drift coefficients $\mathcal{V}_{P_{\theta_0}}(\cdot) = f(\theta_0, \cdot)$ and equivalence of diffusion coefficients $\mathcal{W}_{P_{\theta_0}}(\cdot) = g(\theta_0, \cdot)$, as long as no jumps occur, we derive that for $t \geq \tau_0 = 0$, $M^{\theta_t} = \theta_t^0$ and $X_t^0 = C_t^{P_{\theta_0}}$.

Equivalence of time of jumps:

For the SDCPN^{HSDE}, for each arbitrary place in which the initial token may reside, two transitions are pre-enabled: a guard transition and a delay transition. If either of them becomes enabled and fires, then the other becomes disabled. The time until the guard transition is enabled is $t_*(M^{\theta_0}, C_0) \triangleq \inf\{t - \tau_0 > 0 \mid C_t^{P_{\theta_0}} \in \partial\mathcal{G}_{T_{\theta_0}^G}\}$. The time until the delay transition is

enabled is $\sigma_1^{T^D_{\theta_0}} = D^{inv}_{T^D_{\theta_0}}(U_1)$, with $D^{inv}_{T^D_{\theta_0}}(u) = \inf\{t - \tau_0 \mid \exp(-\int_{\tau_0}^{t} \mathcal{D}_{T^D_{\theta_0}}(C_s^{P_{\theta_0}})ds) \leq u\}$ and $U_1 \sim U[0,1]$.
For HSDE, from Equation (6), using $k = 0$ and $\tau_0^b = \tau_0$, the time at which the continuous state first hits the boundary of its state space is $\tau_1^b \triangleq \inf\{t > \tau_0 \mid (\theta_t^0, X_t^0) \in \{\{\theta\} \times \partial E_\theta; \theta \in \mathbb{M}\}\}$. It is easily seen that as long as $\theta_t^0 = \theta_0$, then due to $X_t^0 = C_t^{P_{\theta_0}}$ and the equality $\partial \mathcal{G}_{T^G_{\theta_0}} = \partial E_{\theta_0}$, we have that $\inf\{t > \tau_0 \mid (\theta_t^0, X_t^0) \in \{\{\theta\} \times \partial E_\theta; \theta \in \mathbb{M}\}\} = \tau_0 + \inf\{t - \tau_0 > 0 \mid C_t^{P_{\theta_0}} \in \partial \mathcal{G}_{T^G_{\theta_0}}\}$, hence $\tau_1^b = \tau_0 + t_*(M^{\theta_0}, C_0)$. However, there is a possibility that at some time $\tau_1^p < \tau_1^b$, the HSDE solution process state makes a jump due to the Poisson random measure generating a point: Consider Equations (3) and (4), for $k = 0$. A jump is generated when $\sum_{i=1}^N (\vartheta_i - \theta_{t-}^0) p_P(dt, (\Sigma_{i-1}(\theta_{t-}^0, X_{t-}^0), \Sigma_i(\theta_{t-}^0, X_{t-}^0)] \times \mathbb{R}^d) \neq 0$ or when $\int_{\mathbb{R}^d} \psi(\theta_t^0, \theta_{t-}^0, X_{t-}^0, \underline{z}) p_P(dt, (0, \Lambda(\theta_{t-}^0, X_{t-}^0)] \times d\underline{z}) \neq 0$, or both. Consider the Poisson random measure in Equation (4), i.e. $p_P(dt, (0, \Lambda(\theta_{t-}^0, X_{t-}^0)] \times d\underline{z})$, which is equal to zero, except at singular times when it generates a multivariate point $(\{\tau_1^p\}, \{z_1\}, \{\underline{z}\})$. Due to the Poisson random measure being homogeneous and due to $\Lambda(\theta_{t-}^0, X_{t-}^0) \leq C_\Lambda$, the point $(\{\tau_1^p\}, \{z_1\}, \{\underline{z}\})$ is generated as follows: Generate a triple $(\varepsilon_1, \nu_1, \underline{\nu}_1)$, with $\varepsilon_1 \sim \text{Exp}(C_\Lambda)$, $\nu_1 \sim U[0, C_\Lambda]$ and $\underline{\nu}_1 \sim \mu$. Accept this triple if $\nu_1 \leq \Lambda(\theta^0_{\tau_0+\varepsilon_1-}, X^0_{\tau_0+\varepsilon_1-})$, otherwise reject it. If it is accepted then $\tau_1^p = \tau_0 + \varepsilon_1$, $z_1 = \nu_1$ and $\underline{z} = \underline{\nu}_1$. If it is not accepted then another triple $(\varepsilon_2, \nu_2, \underline{\nu}_2)$ is generated with $\varepsilon_2 \sim \text{Exp}(C_\Lambda)$, $\nu_2 \sim U[0, C_\Lambda]$ and $\underline{\nu}_2 \sim \mu$, and this triple is accepted if $\nu_2 \leq \Lambda(\theta^0_{\tau_0+\varepsilon_1+\varepsilon_2-}, X^0_{\tau_0+\varepsilon_1+\varepsilon_2-})$. If it is accepted then $\tau_1^0 = \tau_0 + \varepsilon_1 + \varepsilon_2$, $z_1 = \nu_2$ and $\underline{z} = \underline{\nu}_2$. If it is not accepted then another triple $(\varepsilon_3, \nu_3, \underline{\nu}_3)$ is generated, and so on. Hence if $(\varepsilon_r, \nu_r, \underline{\nu}_r)$ is the first triple that is accepted then $\tau_1^p = \tau_0 + \sum_{n=1}^r \varepsilon_n$ and $z_1 = \nu_r$ and $\underline{z} = \underline{\nu}_r$. The interarrival times of the triples accepted through this mechanism are exponential with intensity Λ. In addition, due to $\mathcal{D}_{T^D_\theta}(\cdot) = \Lambda(\theta, \cdot)$, we find that $\tau_1^p - \tau_0$ is probabilistically equivalent to $\sigma_1^{T^D_{\theta_0}}$. For HSDE, the time of the first jump is equal to the minimum of τ_1^b and τ_1^p. Due to the reasoning above, this time of first jump is probabilistically equivalent to the time of first jump of the SDCPN^{HSDE}.

Equivalence of size of jumps
For the SDCPN^{HSDE}, the jump size is determined by the firing measure $\mathcal{F}_T$ of the enabled transition T: for all $\theta \in \mathbb{M}$ and $T \in \{T_\theta^G, T_\theta^D\}$, $\mathcal{F}_T(e^{\vartheta'}, x'; x) = F_T^Q(\vartheta', x'; \theta, x)$, for all $x \in E_\theta \cup \partial E_\theta$, $\vartheta' \in \mathbb{M}$ and $x' \in E_{\vartheta'}$, where F_T^Q is defined through

$$F_T^Q(\{\vartheta'\} \times A'; \theta, x) = \rho(\vartheta', \theta, x) \int_{\mathbb{R}^d} \mathbf{1}_{A'}(x + \psi(\vartheta', \theta, x, \underline{z}))\mu(d\underline{z})$$

For HSDE, the size of jumps is generated as follows: In case of a jump generated by Poisson random measure at time $t = \tau_1^p$, the size of jump in $\{\theta_t^0\}$ is given by

$$\theta^0_{\tau_1^p} - \theta^0_{\tau_1^p-} = \sum_{i=1}^N (\vartheta_i - \theta^0_{\tau_1^p-}) p_P(dt, (\Sigma_{i-1}(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}), \Sigma_i(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-})] \times \mathbb{R}^d)$$

and the size of jump in $\{X_t^0\}$ is given by

$$X^0_{\tau_1^p} - X^0_{\tau_1^p-} = \int_{\mathbb{R}^d} \psi(\theta^0_{\tau_1^p}, \theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}, \underline{z}) p_P(dt, (0, \Lambda(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-})] \times d\underline{z})$$

Now use that the Poisson random measure has generated a point $(\{\tau_1^p\}, \{z_1\}, \{\underline{z}\})$, with $z_1 = \nu_r$ and $\underline{z} = \underline{\nu}_r$ as described above. Random variable z_1 is used as follows: Notice that by Equation (5) and definition of ρ, for all $\theta \in \mathbb{M}$ and all $x \in \mathbb{R}^n$, the interval $(0, \Lambda(\theta, x)]$ is divided into subintervals $(\Sigma_{i-1}(\theta, x), \Sigma_i(\theta, x)]$, i.e. $(0, \Lambda(\theta, x)] = (\Sigma_0(\theta, x), \Sigma_1(\theta, x)] \cup (\Sigma_1(\theta, x), \Sigma_2(\theta, x)] \cup \cdots \cup (\Sigma_{N-1}(\theta, x), \Sigma_N(\theta, x)]$, where $\Sigma_0(\theta, x) = 0$ and $\Sigma_N(\theta, x) = \Lambda(\theta, x)$. The ith interval, i.e. $(\Sigma_{i-1}(\theta, x), \Sigma_i(\theta, x)]$ has a weight $\rho(\vartheta_i, \theta, x) = (\Sigma_i(\theta, x) - \Sigma_{i-1}(\theta, x))/\Lambda(\theta, x)$, with $\sum_{i=1}^N \rho(\vartheta_i, \theta, x) = 1$. Due to $z_1 \in (0, \Lambda(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-})]$, there exists $j \in \{1, \ldots, N\}$ such that $z_1 \in (\Sigma_{j-1}(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}), \Sigma_j(\theta^0_{\tau_1^p}, X^0_{\tau_1^p})]$. This makes $p_P(dt, (\Sigma_{i-1}(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}), \Sigma_i(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-})] \times \mathbb{R}^d) = 1$ if $i = j$ and $= 0$ for $i \neq j$. Therefore, $\theta^0_{\tau_1^p} - \theta^0_{\tau_1^p-} = \vartheta_j - \theta^0_{\tau_1^p-}$, i.e. at time τ_1^p, θ_t jumps from $\theta^0_{\tau_1^p-} = \theta_0$ to $\theta^0_{\tau_1^p} = \vartheta_j$. Next, the random variable $\underline{z}$ is used to determine $X^0_{\tau_1^p} - X^0_{\tau_1^p-}$, i.e. in $(\{\tau_1^p\}, \{z_1\}, \{\underline{z}\})$, $p_P(dt, (0, \Lambda(\theta^0_{t-}, X^0_{t-})] \times d\underline{z}) = 1$ and is zero elsewhere. Therefore, $X^0_{\tau_1^p} - X^0_{\tau_1^p-} = \psi(\vartheta_j, \theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}, \underline{z})$. This gives that at time τ_1^p, X_t jumps from $X^0_{\tau_1^p-}$ to $X^0_{\tau_1^p-} + \psi(\vartheta_j, \theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}, \underline{z})$. From this, we find that the probability for (θ_t, X_t) to jump into $(\{\vartheta_j\}, A)$, given that the state right before the jump is $(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-})$, is equal to the probability that z_1 is in $(\Sigma_{j-1}(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}), \Sigma_j(\theta^0_{\tau_1^p-}, X^0_{\tau_1^p-})]$, times the probability that $X^0_{\tau_1^p-} + \psi(\vartheta_j, \theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}, \underline{z})$ is in A. This probability is equal to

$$\rho(\vartheta_j, \theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}) \int_{\mathbb{R}^d} \mathbf{1}_A(X^0_{\tau_1^p-} + \psi(\vartheta_j, \theta^0_{\tau_1^p-}, X^0_{\tau_1^p-}, \underline{z}))\mu(d\underline{z})$$

which is equal to $Q(\{\vartheta_j\} \times A; \theta^0_{\tau_1^p-}, X^0_{\tau_1^p-})$, according to Equation (8).
For boundary hitting type of jumps, the size of jump is given by Equation (7), i.e.

$$\mathbb{P}\{\theta^1_{\tau_1^b} = \vartheta, X^1_{\tau_1^b} \in A \mid \theta^0_{\tau_1^b-} = \theta, X^0_{\tau_1^b-} = x\} = Q(\{\vartheta\} \times A; \theta, x)$$

This shows that the jump size mechanisms for Poisson random measure type of jumps and boundary hitting type of jumps are the same. Also note that for all ϑ', x', θ and x, and $T \in \mathcal{T}_D \cup \mathcal{T}_G$, $F_T^Q(\vartheta', x'; \theta, x) = Q(\vartheta', x'; \theta, x)$. This means that the SDCPN^{HSDE} state after the jump and the HSDE solution process state after the jump are probabilistically equivalent.

Equivalence of processes after the first jump:
From $\tau_1 = \min\{\tau_1^b, \tau_1^p\}$ onwards, the probabilistic equivalence of the HSDE and SDCPN^{HSDE} processes is shown in the same way. If $\tau_1 = \tau_1^p$, then Equations (3) and (4) are used for $k = 0$; if $\tau_1 = \tau_1^b$ then these equations are used for $k = 1$. From stopping time τ_{n-1} to stopping time τ_n the HSDE solution process and the associated SDCPN^{HSDE} process have probabilistically equivalent paths and probabilistically equivalent stopping times. Due to the unique definition of the SDCPN^{HSDE} stochastic process at times when transitions fire, the SDCPN^{HSDE} state at stopping times is also equivalent to the HSDE solution process state at the stopping times and both processes are càdlàg.
This completes the proof of Theorem 5.3.

Appendix B: Proof of Theorem 5.4

Consider an arbitrary SDCPN $(\mathcal{P}, \mathcal{T}, \mathcal{A}, \mathcal{N}, \mathcal{S}, \mathcal{C}, \mathcal{I}, \mathcal{V}, \mathcal{W}, \mathcal{G}, \mathcal{D}, \mathcal{F})$ that satisfies rules R0–R4. It is assumed that in the initial marking no immediate transitions are enabled, that the delay rates $\mathcal{D}_T$ are bounded, and that for $t \to \infty$, the number of tokens remains finite. First, we characterise the HSDE elements ($\mathbb{M}$, E, f, g, μ_{θ_0,X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$), in terms of SDCPN elements, where it is assumed that μ is given. The thus constructed HSDE is referred to as HSDE^{SDCPN}. Subsequently, we show that the HSDE^{SDCPN} solution process is probabilistically equivalent to the stochastic process defined by the original SDCPN.

B.1 Construction of HSDE^{SDCPN} elements

We provide an into-mapping that characterises HSDE^{SDCPN} elements ($\mathbb{M}$, E, f, g, μ_{θ_0,X_0}, Λ, ψ, ρ, μ, p_P, $\{W_t\}$) in terms of SDCPN elements ($\mathcal{P}, \mathcal{T}, \mathcal{A}, \mathcal{N}, \mathcal{S}, \mathcal{C}, \mathcal{I}, \mathcal{V}, \mathcal{W}, \mathcal{G}, \mathcal{D}, \mathcal{F}$).

$\mathbb{M}$ The characterisation of $\mathbb{M}$ in terms of SDCPN elements is by means of the reachability graph (RG). The nodes in the RG are token distributions, written as row vectors $(m_1, \ldots, m_{|\mathcal{P}|})$, where m_i is the number of tokens in place P_i. Arrows between nodes are labelled by transitions, and indicate how the number of tokens in the places change due to transition firings. Then $\mathbb{M}$ is composed of those nodes in the reachability graph that do not enable an immediate transition, and $N = |\mathbb{M}|$.

E For each $\theta \in \mathbb{M}$, corresponding with node $m = (m_1, \ldots, m_{|\mathcal{P}|})$ in the RG, define $d(\theta) = \sum_{i=1}^{|\mathcal{P}|} m_i n(P_i)$. If under token distribution θ, no guard transitions are pre-enabled, then $E_\theta = \mathbb{R}^{d(\theta)}$. If under token distribution θ, one or more guard transitions are pre-enabled, then $E_\theta = \mathbb{R}^{d(\theta)} \setminus \partial E_\theta$, where ∂E_θ is constructed as follows: Without loss of generality, suppose that under token distribution θ, the multi-set of pre-enabled guard transitions is $T_1, \ldots, T_k$. This set may contain one transition multiple times, if such transition evaluates multiple input token vectors in parallel. Suppose $\{P_{i_1}, \ldots, P_{i_{r_i}}\} = P(A_{in,OE}(T_i))$ are the input places of T_i that are connected to T_i by means of ordinary or enabling arcs. This set may contain one place multiple times if such place is connected to T_i by multiple arcs (input arcs of T_i). Define $d_i = \sum_{j=1}^{r_i} n(P_{i_j})$, then $\partial E_\theta = \partial \mathcal{G}'_{T_1} \cup \ldots \cup \partial \mathcal{G}'_{T_k}$, where $\mathcal{G}'_{T_i} = [\mathcal{G}_{T_i} \times \mathbb{R}^{d(\theta)-d_i}] \in \mathbb{R}^{d(\theta)}$. Here $[\cdot]$ denotes a special ordering of all vector elements: Vector elements are ordered according to the unique ordering of places and to the unique ordering of tokens within their place defined for SDCPN. Finally, $E = \{\{\theta\} \times E_\theta; \theta \in \mathbb{M}\}$.

f For each $\theta \in \mathbb{M}$ and $x \in E_\theta$, $f(\theta, x) = \text{Col}_{i=1}^{|\mathcal{P}|} \left\{ \text{Col}_{j=1}^{m_i} \{\mathcal{V}_{P_i}(c_{ij})\} \right\}$, where $x = \text{Col}_{i=1}^{|\mathcal{P}|} \left\{ \text{Col}_{j=1}^{m_i} \{c_{ij}\} \right\}$ and θ corresponds to $(m_1, \ldots, m_{|\mathcal{P}|})$.

g: For each $\theta \in \mathbb{M}$ and $x \in E_\theta$,
$g(\theta, x) = \text{Row}\{\text{Diag}_{i=1}^{|\mathcal{P}|} \left\{ \text{Diag}_{j=1}^{m_i} \{\mathcal{W}_{P_i}(c_{ij})\} \right\}, O^{\Sigma_{i=1}^{|\mathcal{P}|}(m_i^{max} - m_i)h(P_i)}\}$, where

- $O^{\Sigma_{i=1}^{|\mathcal{P}|}(m_i^{max} - m_i)h(P_i)}$ is a square matrix of dimension $(\Sigma_{i=1}^{|\mathcal{P}|}(m_i^{max} - m_i)h(P_i)) \times (\Sigma_{i=1}^{|\mathcal{P}|}(m_i^{max} - m_i)h(P_i))$ that contains only zeros. In the $g(\theta, \cdot)$ constructed above it is put to the right of the block that contains the matrices $\mathcal{W}_{P_i}$.
- $m_i^{max} = \max_{\theta \in \mathbb{M}}\{m_i \mid \theta = (m_1, \ldots, m_{|\mathcal{P}|})\}$ is the maximum number of tokens that exists in place P_i. This maximum m_i^{max} exists due to the condition that for $t \to \infty$ the number of tokens remains finite.

μ_{θ_0,X_0}: $\mu_{\theta_0,X_0}(M_0, C_0) = \mathcal{I}(M_0, C_0)$ for all M_0 and C_0, where $M_0 = (M_{1,0}, \ldots, M_{|\mathcal{P}|,0})$, with $M_{i,0}$ the initial number of tokens in place P_i, with the places ordered according to the unique ordering adopted for SDCPN, and $C_0 \in \mathbb{R}^{d(\theta_0)}$ containing the colours of these tokens. Due to the condition that no immediate transitions are enabled in the initial marking (which prevents vanishing token distributions to be current at the initial time), the constructed M_0 and C_0 are uniquely defined.

Λ: For each $\theta \in \mathbb{M}$ and $x \in E_\theta$, $\Lambda(\theta, x) = \sum_{n=1}^{k} \mathcal{D}_{T_n}(c^{T_n})$, where $T_1, \ldots, T_k$ refers to the multi-set of transitions in $\mathcal{T}_D$ that, under token distribution θ, are pre-enabled, and c^{T_n} are the respective elements of x that are used to pre-enable these transitions. This set $T_1, \ldots, T_k$ may contain one transition multiple times, if multiple input token vectors are evaluated in parallel. If the set of pre-enabled delay transitions is empty in θ, then $\Lambda(\theta, \cdot) = 0$.

ψ, ρ, μ: we make use of the assumption that μ is given. As part of the construction, define a probability measure $P_Q(\theta', A; \theta, x)$, the value of which equals the probability that if a jump occurs, and if the value of the HSDE solution process just prior to the jump is (θ, x), then the value of the HSDE solution process just after the jump is in (θ', A). Probability $P_Q(\theta', A; \theta, x)$ is characterised in terms of the SDCPN by the reachability graph (RG), elements $\mathcal{D}$, $\mathcal{G}$ and Rules R0–R4 and the set $\mathcal{F}$. This is done in four steps, precisely following the characterisation of the GSHS transition measure Q in terms of SDCPN elements in the appendix of (Everdij & Blom, 2006). Next, we characterise ψ and ρ in terms of this result: For HSDE, due to Equation (7), the probability that given a jump from (θ, x), the state after the jump is in (θ', A) is given by $Q(\{\theta'\} \times A; \theta, x)$ hence we find that $P_Q = Q$. Here, Q is given by Equation (8). From this, we find

$$\rho(\theta', \theta, x) = Q(\{\theta'\} \times \mathbb{R}^n; \theta, x)$$

Next write, for any x',

$$\begin{aligned} Q(\{\theta'\}, x'; \theta, x) &= \rho(\theta', \theta, x) \cdot \mathbb{P}\{x + \psi(\theta', \theta, x, \underline{z}) = x'\} \\ &= \rho(\theta', \theta, x) \cdot \mathbb{P}\{\underline{z} = \psi^{inv}(x' - x)\} \\ &= \rho(\theta', \theta, x) \cdot \mu(\psi^{inv}(x' - x)) \end{aligned}$$

where ψ^{inv} is such that $\mu_L\{u \mid \psi^{inv}(u) \in B\} = \psi(\theta', \theta, x, B)$. Therefore,

$$\mu(\psi^{inv}(x' - x)) = \frac{Q(\{\theta'\}, x'; \theta, x)}{\rho(\theta', \theta, x)}$$

and ψ is finally defined by

$$\psi^{inv}(x' - x)) = \mu^{inv}\left(\frac{Q(\{\theta'\}, x'; \theta, x)}{\rho(\theta', \theta, x)}\right)$$

$\{W_t\}$: This is generated according to the standard mechanism to generate Wiener processes. An h-dimensional Wiener process is constructed by collecting a number of $h = \sum_{i=1}^{|\mathcal{P}|} m_i^{max} h(P_i)$ independent one-dimensional Wiener processes in a vector.

Adding zeros and transforming discrete state vectors We add a sufficient number of zeros to some of the elements in order to create a constant dimension for the HSDE^{SDCPN} hybrid state space. Denote $n = \max_\theta d(\theta)$, 0^a as a column vector of zeros in $\mathbb{R}^a$ and $0^{a \times b}$ as a matrix of zeros in $\mathbb{R}^{a \times b}$, then E is redefined as $E = \{\{\theta\} \times (E_\theta \times \mathbb{R}^{n-d(\theta)}); \theta \in \mathbb{M}\}$; f is redefined as $\text{Col}\{f, 0^{n-d(\theta)}\}$; g is redefined as $\text{Col}\{g, 0^{(n-d(\theta)) \times \Sigma_i m_i^{max} \cdot h(P_i)}\}$; X_0 is redefined as $\text{Col}\{X_0, 0^{n-d(\theta)}\}$ and ψ is redefined as $\text{Col}\{\psi, 0^{n-d(\theta)}\}$.

This shows that all HSDE^{SDCPN} elements can be characterised uniquely in terms of SDCPN elements.

B.2 Probabilistic equivalence

Subsequently, we show that the solution of the constructed HSDE^{SDCPN} delivers a stochastic process which is probabilistically equivalent to the process defined by the SDCPN. This is done by showing: Equivalence of initial states; Equivalence of continuous evolution until first jump; Equivalence of time of jumps; Equivalence of size of jumps; Equivalence of processes after the first jump.

Equivalence of initial states:

The initial HSDE^{SDCPN}-process state (θ_0, X_0) at $t = \tau_0$ is equivalent to the initial SDCPN state through the mapping constructed above. If $\mathcal{I}^{inv}$ denotes the inverse of $\mathcal{I}$ and $\mu^{inv}_{\theta_0,X_0}$ denotes the inverse of μ_{θ_0,X_0}, then the random variable $(M_0, C_0) = \mathcal{I}^{inv}(U)$ is equivalent to the random variable $(\theta_0, X_0) = \mu^{inv}_{\theta_0,X_0}(U)$. Due to equivalence between $\mathcal{I}$ and μ_{θ_0,X_0}, the initial states are probabilistically equivalent.

Equivalence of continuous evolution until first jump:

At times t when no jump occurs, the HSDE^{SDCPN}-process evolves according to f and g, driven by a Wiener process, and the SDCPN-process evolves according to $\mathcal{V} = \{\mathcal{V}_P; P \in \mathcal{P}\}$ and $\mathcal{W} = \{\mathcal{W}_P; P \in \mathcal{P}\}$, driven by Brownian motion. Through the mappings between f and $\mathcal{V}$ and between g and $\mathcal{W}$ developed above, and through the probabilistic equivalence between Brownian motions and Wiener processes, these evolutions provide probabilistically equivalent processes X_t and C_t for all $t > \tau_0$, until the first jump.

Equivalence of time of jumps:

The times of jumps are generated by forced jumps and spontaneous jumps. In SDCPN, the forced jumps are represented by guard transitions; in HSDE, the forced jumps are represented by continuous state space boundary hits. Due to the mapping between the boundary of the HSDE^{SDCPN} state space ∂E_θ and the boundaries of the transition guards of the guard transitions $\{\partial\mathcal{G}_T; T \in \mathcal{T}_G\}$, the HSDE^{SDCPN} forced jumps and the SDCPN forced jumps occur at the same time. The HSDE^{SDCPN} spontaneous jumps are generated by a Poisson random measure that uses a rate Λ. The SDCPN spontaneous jumps are generated by the delay transitions that use rates $\{\mathcal{D}_T; T \in \mathcal{T}_D\}$. Due to the mapping between Λ and $\{\mathcal{D}_T; T \in \mathcal{T}_D\}$, the time of spontaneous jump is according to the same rate for both HSDE^{SDCPN} and SDCPN.

Equivalence of size of jumps:

At times when a jump occurs, the HSDE^{SDCPN}-process makes a jump generated by ψ, ρ and μ, while the SDCPN-process makes a jump generated by $\mathcal{F}$. Through the mapping between ψ, ρ, μ and $\mathcal{F}$, these jumps provide probabilistically equivalent processes.

Equivalence of processes after the first jump:

After the first jump, equivalence is shown in a similar way as above.
This completes the proof of Theorem 5.4.