Stochastic Differential Dynamic Logic for Stochastic Hybrid Programs

André Platzer

April 2011
CMU-CS-11-111

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

This material is based upon work supported by the National Science Foundation by NSF CAREER Award CNS-1054246, NSF EXPEDITION CNS-0926181, CNS-0931985, CNS-1035800, by ONR N00014-10-1-0188 and DARPA FA8650-10C-7077. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of any sponsoring institution or government.
A conference version of this report has appeared at CADE [Pla11].

Keywords: Dynamic logic, proof calculus, stochastic differential equations, stochastic hybrid systems, stochastic processes, compositional verification

Abstract

Logic is a powerful tool for analyzing and verifying systems, including programs, discrete systems, real-time systems, hybrid systems, and distributed systems. Some applications also have a stochastic behavior, however, either because of fundamental properties of nature, uncertain environments, or simplifications to overcome complexity. Discrete probabilistic systems have been studied using logic. But logic has been chronically underdeveloped in the context of stochastic hybrid systems, i.e., systems with interacting discrete, continuous, and stochastic dynamics. We aim at overcoming this deficiency and introduce a dynamic logic for stochastic hybrid systems. Our results indicate that logic is a promising tool for understanding stochastic hybrid systems and can help taming some of their complexity. We introduce a compositional model for stochastic hybrid systems. We prove adaptivity, càdlàg, and Markov time properties, and prove that the semantics of our logic is measurable. We present compositional proof rules, including rules for stochastic differential equations, and prove soundness.

1 Introduction

Logic has been used very successfully for verifying several classes of system models, including programs [Pra76], discrete systems, real-time systems [Dut95], hybrid systems [Pla10a], distributed systems, and distributed hybrid systems [Pla10b]. This gives us confidence in the power of logic. Not all aspects of real systems can be represented faithfully by these models, however. Some systems are inherently uncertain, either because of fundamental properties of nature, because they operate in an uncertain environment, or because deterministic models are simply too complex. Such systems have a stochastic dynamics. Nondeterministic overapproximations may be too inaccurate for a meaningful analysis, e.g., because a worst-case analysis would let bad environment actions happen always, which is very unlikely. Discrete probabilistic systems have been studied using logic. Yet, complex systems are driven by joint discrete, continuous, and stochastic dynamics. Logic has been chronically underdeveloped in the context of these stochastic hybrid systems.

Classical logic is about boolean truth and yes/no answers. That is why it is tricky to use logic for systems with stochastic effects. Logic has reached out into probabilistic extensions at least for discrete programs [Koz81, Koz85, FH84] and for first-order logic over a finite domain [RD06]. Logic has been used for the purpose of specifying system properties in model checking finite Markov chains [YKNP06] and probabilistic timed automata [KNSW07]. Stochastic hybrid systems, instead, are a domain where logic and especially proof calculi have so far been more conspicuous by their absence. Given how successful logic has been elsewhere, we want to change that.

Stochastic hybrid systems [BL06, CL06, HLS00] are systems with interacting discrete, continuous, and stochastic dynamics. There is not just one canonical way to add stochastic behavior to a system model. Stochasticity might be restricted to the discrete dynamics, as in piecewise deterministic Markov decision processes, restricted to the continuous and switching behavior as in switching diffusion processes [GAM97], or allowed in many parts as in so-called General Stochastic Hybrid Systems; see [BL06, CL06] for an overview. Several different forms of combinations of probabilities with hybrid systems and continuous systems have been considered, both for model checking [FTE10, KR08, CL06] and for simulation-based validation [MS06, ZPC10].

We develop a very different approach. We consider logic and theorem proving for stochastic hybrid systems[1] to transfer the success that logic has had in other domains. Our approach is partially inspired by probabilistic PDL [Koz85] and by barrier certificates for continuous dynamics [PJP07]. We follow the arithmetical view that Kozen identified as suitable for probabilistic logic [Koz85].

Classical analysis is provably inadequate [KP10] for analyzing even simple continuous stochastic processes. We heavily draw on both stochastic calculus and logic. It is not possible to present all mathematical background exhaustively here. But we provide basic definitions and intuition and refer to the literature for details and proofs of the main results of stochastic calculus [KS91, Øks07, KP10].

[1] Note that there is a model called Stochastic Hybrid Systems [HLS00]. We do not mean this specific model in the narrow sense but refer to stochastic hybrid systems as the broader class of systems that share discrete, continuous, and stochastic dynamics.

Our most interesting contributions are:

1. We present the new model of stochastic hybrid programs (SHPs) and define a compositional semantics of SHP executions in terms of stochastic processes.
2. We prove that the semantic processes are adapted, have almost surely càdlàg paths, and that their natural stopping times are Markov times.
3. We introduce a new logic called stochastic differential dynamic logic (SdL) for specifying and verifying properties of SHPs.
4. We define a semantics and prove that it is measurable such that probabilities are well-defined and probabilistic questions become meaningful.
5. We present proof rules for SdL and prove their soundness.
6. We identify the requirements for using Dynkin's formula for proving properties using the infinitesimal generator of stochastic differential equations.

SdL makes the rich semantical complexity and deep theory of stochastic hybrid systems accessible in a simple syntactic language. This makes the verification of stochastic hybrid systems possible with elementary syntactic proof principles.

2 Preliminaries: Stochastic Processes

We fix a dimension $d \in \mathbb{N}$ for the Euclidean state space $\mathbb{R}^d$ equipped with its *Borel σ-algebra* $\mathcal{B}$, i.e., the σ-algebra generated by all open subsets. A *σ-algebra* on a set Ω is a nonempty set $\mathcal{F} \subseteq 2^\Omega$ that is closed under complement and countable union. We axiomatically fix a *probability space* $(\Omega, \mathcal{F}, P)$ with a σ-algebra $\mathcal{F} \subseteq 2^\Omega$ of events on space Ω and a probability measure P on $\mathcal{F}$ (i.e., $P : \mathcal{F} \to [0,1]$ is countable additive with $P \geq 0, P(\Omega) = 1$). We assume the probability space has been completed, i.e., every subset of a null set (i.e., $P(A) = 0$) is measurable. A property holds *P-almost surely* (*a.s.*) if it holds with probability 1. A *filtration* is a family $(\mathcal{F}_t)_{t\geq 0}$ of σ-algebras that is increasing, i.e., $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s < t$. Intuitively, $\mathcal{F}_t$ are the events that can be discriminated at time t. We always assume a filtration $(\mathcal{F}_t)_{t\geq 0}$ that has been completed to include all null sets and that is right-continuous, i.e., $\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u$ for all t. We generally assume the compatibility condition that $\mathcal{F}$ coincides with the σ-algebra $\mathcal{F}_\infty := \sigma\left(\bigcup_{t\geq 0} \mathcal{F}_t\right)$, i.e., the σ-algebra generated by all $\mathcal{F}_t$.

For a σ-algebra Σ on a set D and the Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}^d$, function $f : D \to \mathbb{R}^d$ is *measurable* iff $f^{-1}(B) \in \Sigma$ for all $B \in \mathcal{B}$ (or, equivalently, for all open $B \subseteq \mathbb{R}^d$). An $\mathbb{R}^d$-valued *random variable* is an $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}^d$. All sets and functions definable in first-order logic over real arithmetic are Borel-measurable. A *stochastic process* X is a collection $\{X_t\}_{t\in T}$ of $\mathbb{R}^d$-valued random variables X_t indexed by some set T for time. That is, $X : T \times \Omega \to \mathbb{R}^d$ is a function such that for all $t \in T$, $X_t = X(t,\cdot) : \Omega \to \mathbb{R}^d$ is a random variable. Process X is *adapted* to filtration $(\mathcal{F}_t)_{t\geq 0}$ if X_t is $\mathcal{F}_t$-measurable for each t. That is, the process does not depend on future

events. We consider only adapted processes (e.g., using the completion of the natural filtration of a process or the completion of the optional σ-algebra for $\mathcal{F}$ [KS91]). A process X is *càdlàg* iff its *paths* $t \mapsto X_t(\omega)$ (for each $\omega \in \Omega$) are càdlàg a.s., i.e., right-continuous ($\lim_{s \searrow t} X_s(\omega) = X_t(\omega)$) and left limits ($\lim_{s \nearrow t} X_s(\omega)$) exist.

We further need an e-dimensional *Brownian motion* W (i.e., W is a stochastic process starting at 0 that is almost surely continuous and has independent increments that are normally distributed with mean 0 and variance equal to the time difference). Brownian motion is mathematically extremely complex. Its paths are almost surely continuous everywhere but differentiable nowhere and of unbounded variation. Intuitively, W can be understood as the limit of a random walk. We denote the Euclidean vector norm by $|x|$ and use the Frobenius norm $|\sigma| := \sqrt{\sum_{i,j} \sigma_{ij}^2}$ for matrices $\sigma \in \mathbb{R}^{d \times e}$.

3 Stochastic Differential Equations

We consider stochastic differential equations [Øks07, KP10] to describe stochastic continuous system dynamics. They are like ordinary differential equations but have an additional diffusion term that varies the state stochastically. Stochastic differential equations are of the form $dX_t = b(X_t)dt + \sigma(X_t)dW_t$. We consider Itō stochastic differential equations, whose solutions are defined by the stochastic Itō integral [Øks07, KP10], which is again a stochastic process. Like in an ordinary differential equation, the drift coefficient $b(X_t)$ determines the deterministic part of how X_t changes over time as a function of its current value. As a function of X_t, the diffusion coefficient $\sigma(X_t)$ determines the stochastic influence by integration with respect to the Brownian motion process W_t. See Fig. 1 for two sample paths. Ordinary differential equations are retained for $\sigma = 0$. We focus on the time-homogenous case, where b and σ are time-independent, because time could be added as an extra state variable.

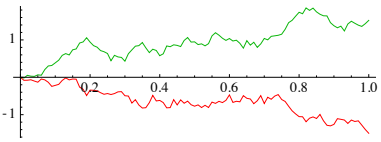

Figure 1: Sample paths with $b = 1$ (top) and $b = 0$ (bottom), $\sigma = 1$

Definition 1 (Stochastic differential equation) *A stochastic process $X : [0, \infty) \times \Omega \to \mathbb{R}^d$ solves the (Itō)* stochastic differential equation

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t \tag{1}$$

with $X_0 = Z$, if $X_t = Z + \int b(X_t)dt + \int \sigma(X_t)dW_t$, where $\int \sigma(X_t)dW_t$ is an Itō integral process [Øks07, KP10].

For simplicity, we always assume $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times e}$ to be measurable and locally Lipschitz-continuous:

$$\forall N \exists C \forall x, y : \ |x|, |y| \leq N \ \Rightarrow \ |b(x) - b(y)| \leq C|x - y|, |\sigma(x) - \sigma(y)| \leq C|x - y|$$

As an integral of an a.s. continuous process, solution X has almost surely continuous paths [Øks07]. A.s. continuous solution X is pathwise unique [KP10, Ch 4.5]. Process X is a strong Markov process for each initial value x [Øks07, Theorem 7.2.4].

4 Stochastic Hybrid Programs

As a system model for stochastic hybrid system, we introduce stochastic hybrid programs (SHPs). SHPs combine stochastic differential equations for describing the stochastic continuous system dynamics with program operations to describe the discrete switching, jumps, and discrete stochastic choices. These primitive dynamics can be combined programmatically in flexible ways. All basic terms in stochastic hybrid programs and stochastic differential dynamic logic are polynomial terms built over real-valued variables and rational constants. Our approach is sound for more general settings, but first-order real arithmetic is decidable [Tar51].

4.1 Syntax

Stochastic hybrid programs (SHPs) are formed by the following grammar (where x_i is a variable, x a vector of variables, θ a term, b a vector of terms, σ a matrix of terms, H is a quantifier-free first-order real arithmetic formula, $\lambda, \nu \geq 0$ are rational numbers):

$$\alpha \ ::= \ x_i := \theta \mid x_i := * \mid ?H \mid dx = bdt + \sigma dW \,\&\, H \mid \lambda\alpha \ \oplus \ \nu\beta \mid \alpha;\beta \mid \alpha^*$$

Assignment $x_i := \theta$ deterministically assigns term θ to variable x_i instantaneously. *Random assignment* $x_i := *$ randomly updates variable x_i, but unlike in classical dynamic logic [Pra76], we assume a probability distribution for x. As one example for a probability distribution, we consider uniform distribution in the interval [0,1], but other distributions can be used as long as they are computationally tractable, e.g., definable in first-order real arithmetic.

Most importantly, $dx = bdt + \sigma dW \,\&\, H$ represents a stochastic continuous evolution along a stochastic differential equation, restricted to the evolution domain region H, i.e., the stochastic process will not continue when it leaves H. We assume that $dx = bdt + \sigma dW$ satisfies the assumptions of stochastic differential equations according to Def. 1. In particular, the dimensions of the vectors x, b, matrix σ, and (vectorial) Brownian motion W fit together and b, σ are globally Lipschitz-continuous (which is first-order definable for polynomial terms and, thus, decidable by quantifier elimination [Tar51]).

Test $?H$ represents a stochastic process that fails (disappears into an absorbing state) if H is not satisfied yet continues unmodified otherwise. *Linear combination* $\lambda\alpha \ \oplus \ \nu\beta$ evolves like α in λ percent of the cases and like β otherwise. We simply assume $\lambda + \nu = 1$. *Sequential composition* $\alpha;\beta$ and *repetition* α^* work similarly to dynamic logic [Pra76], except that they combine SHPs.

4.2 Stochastic Process Semantics

The semantics of a SHP is the stochastic process that it generates. The semantics $[\![\alpha]\!]$ of a SHP α consists of a function $[\![\alpha]\!] : (\Omega \to \mathbb{R}^d) \to ([0,\infty) \times \Omega \to \mathbb{R}^d)$ that maps any $\mathbb{R}^d$-valued random variable Z describing the initial state to a stochastic process $[\![\alpha]\!]^Z$ together with a function $(\!|\alpha|\!) : (\Omega \to \mathbb{R}^d) \to (\Omega \to \mathbb{R})$ that maps any $\mathbb{R}^d$-valued random variable Z describing the initial state to a stopping time $(\!|\alpha|\!)^Z$ indicating when to stop $[\![\alpha]\!]^Z$. Often, an $\mathcal{F}_0$-measurable random variable Z or deterministic state is used to describe the initial state. We assume independence of Z from subsequent stochastic processes like Brownian motions occurring in the definition of $[\![\alpha]\!]^Z$.

For an $\mathbb{R}^d$-valued random variable Z, we denote by $\hat{Z}$ the stochastic process

$$\hat{Z} : \{0\} \times \Omega \to \mathbb{R}^d; (0, \omega) \mapsto \hat{Z}_0(\omega) := Z(\omega)$$

that is stuck at Z. We write $\hat{x}$ for the random variable Z that is a deterministic state $Z(\omega) := x$ for all $\omega \in \Omega$. We write $[\![\alpha]\!]^x$ and $(\!|\alpha|\!)^x$ for $[\![\alpha]\!]^Z$ and $(\!|\alpha|\!)^Z$ then.

In order to simplify notation, we assume that all variables are uniquely identified by an index, i.e., the only occurring variables are $x_1, x_2, \ldots, x_d$. We write $Z(\omega) \models H$ if state $Z(\omega)$ satisfies first-order real arithmetic formula H and $Z(\omega) \not\models H$ otherwise. In the semantics we will use a family of random variables $\{U_i\}_{i \in I}$ that are distributed uniformly in $[0, 1]$ and independent of other U_j and all other random variables and stochastic processes in the semantics. Hence, U satisfies $P(\{\omega \in \Omega : U(\omega) \leq s\}) = \int_{-\infty}^{s} \mathcal{I}_{[0,1]} dt$ with the usual extensions to other Borel subsets. To describe this situation, we just say that "$U \sim \mathcal{U}(0, 1)$ is i.i.d. (independent and identically distributed)", meaning that U is furthermore independent of all other random variables and stochastic processes in the semantics. We denote the *characteristic function* of a set S by $\mathcal{I}_S$, which is defined by $\mathcal{I}_S(x) := 1$ if $x \in S$ and $\mathcal{I}_S(x) := 0$ otherwise.

Definition 2 (Stochastic hybrid program semantics) *The* semantics *of SHP α is defined by*

$$[\![\alpha]\!] : (\Omega \to \mathbb{R}^d) \to ([0, \infty) \times \Omega \to \mathbb{R}^d); Z \mapsto [\![\alpha]\!]^Z = ([\![\alpha]\!]^Z_t)_{t \geq 0}$$

and

$$(\!|\alpha|\!) : (\Omega \to \mathbb{R}^d) \to (\Omega \to \mathbb{R}); Z \mapsto (\!|\alpha|\!)^Z$$

These functions are inductively defined for random variable Z by

1. $[\![x_i := \theta]\!]^Z = \hat{Y}$ *where* $Y(\omega)_i = [\![\theta]\!]^{Z(\omega)}$ *and* $Y_j = Z_j$ *for all* $j \neq i$. *Further,* $(\!|x_i := \theta|\!)^Z = 0$.
2. $[\![x_i := *]\!]^Z = \hat{U}$ *where* $U_j = Z_j$ *for all* $j \neq i$ *and* $U_i \sim \mathcal{U}(0, 1)$ *is i.i.d. and* $\mathcal{F}_0$*-measurable. Further,* $(\!|x_i := *|\!)^Z = 0$.
3. $[\![?H]\!]^Z = \hat{Z}$ *on the event* $\{Z \models H\}$ *and* $(\!|?H|\!)^Z = 0$ *(on all events* $\omega \in \Omega$*). Note that* $[\![?H]\!]^Z$ *is not defined on the event* $\{Z \not\models H\}$.
4. $[\![dx = bdt + \sigma dW \,\&\, H]\!]^Z$ *is the process* $X : [0, \infty) \times \Omega \to \mathbb{R}^d$ *that solves the (Itō) stochastic differential equation* $dX_t = [\![b]\!]^{X_t} dt + [\![\sigma]\!]^{X_t} dB_t$ *with* $X_0 = Z$ *on the event* $\{Z \models H\}$, *where* B_t *is a fresh* e*-dimensional Brownian motion if* σ *has* e *columns. We assume that* Z *is independent of the* σ*-algebra generated by* $(B_t)_{t \geq 0}$.
   Further, $(\!|dx = bdt + \sigma dW \,\&\, H|\!)^Z = \inf\{t \geq 0 : X_t \notin H\}$. *Note that* X *is not defined on the event* $\{Z \not\models H\}$.
5. $$[\![\lambda\alpha \,\oplus\, \nu\beta]\!]^Z = \mathcal{I}_{U \leq \lambda} [\![\alpha]\!]^Z + \mathcal{I}_{U > \lambda} [\![\beta]\!]^Z = \begin{cases} [\![\alpha]\!]^Z & \text{on the event } \{U \leq \lambda\} \\ [\![\beta]\!]^Z & \text{on the event } \{U > \lambda\} \end{cases}$$
   $$(\!|\lambda\alpha \,\oplus\, \nu\beta|\!)^Z = \mathcal{I}_{U \leq \lambda} (\!|\alpha|\!)^Z + \mathcal{I}_{U > \lambda} (\!|\beta|\!)^Z$$
   where $U \sim \mathcal{U}(0, 1)$ *is i.i.d. and* $\mathcal{F}_0$*-measurable.*

6. $\quad [\![\alpha;\beta]\!]^Z_t = \begin{cases} [\![\alpha]\!]^Z_t & \text{on } \{t < (\!|\alpha|\!)^Z\} \\ [\![\beta]\!]^{[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}}_{t-(\!|\alpha|\!)^Z} & \text{on } \{t \geq (\!|\alpha|\!)^Z\} \end{cases}$ *and* $(\!|\alpha;\beta|\!)^Z = (\!|\alpha|\!)^Z + (\!|\beta|\!)^{[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}}$

7. $\quad [\![\alpha^*]\!]^Z_t = [\![\alpha^n]\!]^Z_t$ *on the event* $\{(\!|\alpha^n|\!)^Z > t\}$ *and* $(\!|\alpha^*|\!)^Z = \lim_{n\to\infty} (\!|\alpha^n|\!)^Z$

where $\alpha^0 \equiv ?\mathit{true}$, $\alpha^1 \equiv \alpha$, *and* $\alpha^{n+1} \equiv \alpha;\alpha^n$.

For Case 7, note that $(\!|\alpha^n|\!)^Z$ is monotone in n, hence the limit $(\!|\alpha^*|\!)^Z$ exists and is finite if the sequence is bounded. The limit is ∞ otherwise. Note that $[\![\alpha^*]\!]^Z_t$ is independent of the choice of n on the event $\{(\!|\alpha^n|\!)^Z > t\}$ (but not necessarily independent of n on the event $\{(\!|\alpha^n|\!)^Z \geq t\}$, because α might start with a jump after α^n). Observe that $[\![\alpha^*]\!]^Z_t$ is not defined on the event $\{\forall n\, (\!|\alpha^n|\!)^Z \leq t\}$, which happens, e.g., for Zeno executions violating divergence of time. It would still be possible to give a semantics in this case, e.g., at $t = (\!|\alpha^n|\!)^Z$, but we do not gain much from introducing those technicalities.

In the semantics of $[\![\alpha]\!]^Z$, time is allowed to end. We explicitly consider $[\![\alpha]\!]^Z_t$ as not defined for a realization ω if a part of this process is not defined, because of failed tests in α. The process may be explicitly not defined when $t > (\!|\alpha|\!)^Z$. Explicitly being not defined can be viewed as being in a special absorbing state that can never be left again, as in killed processes. The stochastic process $[\![\alpha]\!]^Z$ is only intended to be used until time $(\!|\alpha|\!)^Z$. We stop using $[\![\alpha]\!]^Z$ after time $(\!|\alpha|\!)^Z$.

A *Markov time* (a.k.a. stopping time) is a non-negative random variable τ such that $\{\tau \leq t\} \in \mathcal{F}_t$ for all t. For a Markov time τ and a stochastic process X_t, the following process is called *stopped process* X^τ

$$X^\tau_t := X_{t\sqcap\tau} = \begin{cases} X_t & \text{if } t < \tau \\ X_\tau & \text{if } t \geq \tau \end{cases} \quad \text{where} \quad t \sqcap \tau := \min\{t, \tau\}$$

A class $\mathcal{C}$ of processes is *stable under stopping* if $X \in \mathcal{C}$ implies $X^\tau \in \mathcal{C}$ for every Markov time τ. Right continuous adapted processes, and processes satisfying the strong Markov property are stable under stopping [Dyn65, Theorem 10.2].

Most importantly, we show that the semantics is well-defined. We prove that the natural stopping times $(\!|\alpha|\!)^Z$ are actually Markov times so that it is meaningful to stop process $[\![\alpha]\!]^Z$ at $(\!|\alpha|\!)^Z$ and useful properties of $[\![\alpha]\!]^Z$ inherit to the stopped process $[\![\alpha]\!]^Z_{t\sqcap(\!|\alpha|\!)^Z}$. Furthermore, we show that the process $[\![\alpha]\!]^Z$ is adapted (does not look into the future) and càdlàg, which will be important to define a semantics for formulas. We give a proof of the following theorem in Appendix A.1.

Theorem 1 (Adaptive càdlàg process with Markov times) *For each SHP α and any $\mathbb{R}^d$-valued random variable Z, $[\![\alpha]\!]^Z$ is an a.s. càdlàg process and adapted (to the completed filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by Z and the constituent Brownian motion $(B_s)_{s\leq t}$ and uniform U processes) and $(\!|\alpha|\!)^Z$ is a Markov time (for $(\mathcal{F}_t)_{t\geq 0}$). In particular, the end value $[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}$ is again $\mathcal{F}_{(\!|\alpha|\!)^Z}$-measurable.*

Note in particular, that the event $\{(\!|\alpha^n|\!)^Z \geq t\}$ is $\mathcal{F}_t$-measurable, thus, by [KS91, Prop 1.2.3], the event $\{(\!|\alpha^n|\!)^Z > t\}$ in Case 7 of Def. 2 is $\mathcal{F}_t$-measurable. As a corollary to Theorem 1, $[\![\alpha]\!]^Z$ is progressively measurable [KS91, Prop 1.1.13].

5 Stochastic Differential Dynamic Logic

For specifying and analyzing properties of SHPs, we introduce *stochastic differential dynamic logic* SdℒC.

5.1 Syntax

Function terms of *stochastic differential dynamic logic* SdℒC are formed by the grammar (F is a primitive measurable function definable in first-order real arithmetic, e.g., the characteristic function $\mathcal{I}_S$ of a measurable set S definable in first-order real arithmetic, B is a boolean combination of such characteristic functions using operators $\wedge, \vee, \neg, \rightarrow$ from Fig. 2, λ, ν are rational numbers):

$$
\begin{aligned}
0 &\equiv \mathcal{I}_{\emptyset} \\
1 &\equiv \mathcal{I}_{\mathbb{R}^d} \\
\neg f &\equiv 1 - f \\
A \wedge B &\equiv AB \\
A \vee B &\equiv A + B - AB \\
A \rightarrow B &\equiv 1 - A + AB \\
\texttt{if}(H)\,\{\alpha\}\texttt{else}\{\beta\} &\equiv \frac{1}{2}(?H;\alpha) \,\oplus\, \frac{1}{2}(?\neg H;\beta) \\
\texttt{while}(H)\,\{\alpha\} &\equiv (?H;\alpha)^*;?\neg H \\
[\alpha]f &\equiv \neg\langle\alpha\rangle\neg f
\end{aligned}
$$

Figure 2: Common SdℒC and SHP abbreviations

$$f, g \;::=\; F \mid \lambda f + \nu g \mid Bf \mid \langle\alpha\rangle f$$

These are for linear ($\lambda f + \nu g$) or boolean product (Bf) combinations of terms. Term $\langle\alpha\rangle f$ represents the supremal value of f along the process belonging to α. The syntactic abbreviations in Fig. 2 can be useful. Formulas of SdℒC are simple, because SdℒC function terms are powerful. SdℒC formulas express equational and inequality relations between SdℒC function terms f, g. They are of the form:

$$\phi \;::=\; f \leq g \mid f = g$$

5.2 Measurable Semantics

The semantics of classical logics maps an interpretation to a truth-value. This does not work for stochastic logic, because the state evolution of SHPs contained in SdℒC formulas is stochastic, not deterministic. Instead, we define the semantics of an SdℒC function term as a random variable.

Definition 3 (SdℒC semantics) *The semantics $[\![f]\!]$ of a function term f is a function*

$$[\![f]\!] : (\Omega \rightarrow \mathbb{R}^d) \rightarrow (\Omega \rightarrow \mathbb{R})$$

that maps any $\mathbb{R}^d$-valued random variable Z describing the current state to a random variable $[\![f]\!]^Z$. It is defined by

1. $[\![F]\!]^Z = F^\ell(Z)$*, i.e.,* $[\![F]\!]^Z(\omega) = F^\ell(Z(\omega))$ *where function F denotes F^ℓ*
2. $[\![\lambda f + \nu g]\!]^Z = \lambda[\![f]\!]^Z + \nu[\![g]\!]^Z$
3. $[\![Bf]\!]^Z = [\![B]\!]^Z * [\![f]\!]^Z$*, i.e., multiplication* $[\![Bf]\!]^Z(\omega) = [\![B]\!]^Z(\omega) * [\![f]\!]^Z(\omega)$

4. $[\![\langle\alpha\rangle f]\!]^Z = \sup\{[\![f]\!]^{[\![\alpha]\!]^Z_t} : 0 \leq t \leq (\!|\alpha|\!)^Z\}$

When Z is not defined (results from a failed test), then $[\![f]\!]^Z$ is not defined. To avoid partiality, we assume the convention $[\![f]\!]^Z := 0$ when Z is not defined.

If f is a characteristic function of a measurable set, then $[\![\langle\alpha\rangle f]\!]^Z$ corresponds to a random variable that reflects the supremal f value that α can reach at least once during its evolution until stopping time $(\!|\alpha|\!)^Z$ when starting in a state corresponding to random variable Z. Then $P([\![\langle\alpha\rangle f]\!]^Z = 1)$ is the probability with which α reaches f at least once and $E([\![\langle\alpha\rangle f]\!]^Z)$ is the expected value, given Z. This includes the special case where Z is a deterministic state $Z(\omega) := x$ for all $\omega \in \Omega$. But first, we prove that these quantities are well-defined probabilities at all. Note that well-definedness of the definition in case 4 uses Theorem 1.

Cases 1–3 of Def. 3 are as in [Koz85] with the notable exception of case 4, which we define as a supremum, not an integral. The reason is that we are interested in probabilistic worst-case verification, not in average-case verification. For discrete programs, it is often sufficient to consider the input-output behavior, so that Kozen did not need to consider the temporal evolution of the program over time, only its final (probabilistic) outcome [Koz85]. In stochastic hybrid systems, the temporal evolution is highly relevant, in addition to the stochastic behavior. When averaging over time, the system state may be very probably good (the integral of the error is small). But, still, it could be very likely that the system exhibits a bug at some state during a run. In this case, we would still want to declare such a system as broken, because, when using it, it will very likely get us into trouble. Stochastic average-case analysis is interesting for performance analysis. But for safety verification, supremal stochastic analysis is more relevant, because a system that is very probably broken at some time, is still too broken to be used safely. We thus consider stochastic dynamics with worst-case temporal behavior, i.e., our semantics performs stochastic averaging (in the sense of probability) among different behaviors, but considers supremal worst-case probability over time. The logic $\mathsf{Sd}\mathcal{L}$ is intended to be used (among other things) to prove bounds on the probability that a system fails at some point at all.

A car that, on average over all times of its use, has a low crash rate, but still has a high probability of crashing at least once during the first ride would not be safe. This is one example where stochastic hybrid systems exhibit new interesting characteristics that we do not see in discrete systems.

We show that the semantics is well-defined. We prove that $[\![f]\!]^Z$ is, indeed, a random variable, i.e., measurable. Without this, probabilistic questions about the value of formulas would not be well-defined, because they are not measurable with respect to the probability space $(\Omega, \mathcal{F}, P)$ and the Borel σ-algebra on $\mathbb{R}$.

Theorem 2 (Measurability) *For any $\mathbb{R}^d$-valued random variable Z, the semantics $[\![f]\!]^Z$ of function term f is a random variable (i.e., $\mathcal{F}$-measurable).*

We give a proof of this theorem in Appendix A.2.

Corollary 1 (Pushforward measure) *For any $\mathbb{R}^d$-valued random variable Z and function term f, probability measure P induces the pushforward measure*

$$S \mapsto P(([\![f]\!]^Z)^{-1}(S)) = P(\{\omega \in \Omega : [\![f]\!]^Z(\omega) \in S\}) = P([\![f]\!]^Z \in S)$$

which defines a probability measure on $\mathbb{R}$. Hence, for each Borel-measurable set S, the probability $P([\![f]\!]^Z \in S)$ is well-defined.

We say that $f \leq g$ is *valid* if it holds for all $\mathbb{R}^d$-valued random variables Z:

$$\vDash f \leq g \quad \text{iff} \quad \text{for all } Z,\ [\![f]\!]^Z \leq [\![g]\!]^Z,\ \text{i.e., } ([\![f]\!]^Z)(\omega) \leq ([\![g]\!]^Z)(\omega) \text{ for all } \omega \in \Omega$$

Validity of $f = g$ is defined accordingly, hence, $\vDash f = g$ iff $\vDash f \leq g$ and $\vDash g \leq f$. As consequence relation on formulas, we use the *(global) consequence relation* that we define as follows (similarly when some of the formulas are $f_i = g_i$):

$$f_1 \leq g_1, \ldots, f_n \leq g_n \vDash f \leq g$$
$$\text{iff } \vDash f_1 \leq g_1, \ldots, \vDash f_n \leq g_n \text{ implies } \vDash f \leq g$$

Also $f_1 \leq g_1, \ldots, f_n \leq g_n \vDash f \leq g$ holds *pathwise* if it holds for each $\omega \in \Omega$.

6 Stochastic Calculus

In this section, we review important results from stochastic calculus [KS91, Øks07, KP10] that we use in our proof calculus. To indicate the probability law of process X starting at $X_0 = x$ a.s., we write P^x instead of P. By E^x we denote the expectation operator for probability law P^x. That is $E^x(f(X_t)) := \int_\Omega f(X_t(\omega))dP^x(\omega)$ for each Borel-measurable function $f : \mathbb{R}^d \to \mathbb{R}$. A very important concept is the infinitesimal generator that captures the average rate of change of a process.

Definition 4 (Infinitesimal generator) *The* (infinitesimal) generator *of an a.s. right continuous strong Markov process (e.g., solution from Def. 1) is the operator A that maps a function $f : \mathbb{R}^d \to \mathbb{R}$ to function $Af : \mathbb{R}^d \to \mathbb{R}$ defined as*

$$Af(x) := \lim_{t \searrow 0} \frac{E^x f(X_t) - f(x)}{t}$$

We say that Af is defined if this limit exists for all $x \in \mathbb{R}^d$. The generator can be used to compute the expected value of a function when following the process until a Markov time without solving the SDE.

Theorem 3 (Dynkin's formula [Øks07, Theorem 7.4.1],[Dyn65, p. 133]) *Let X_t an a.s. right continuous strong Markov process (e.g., solution from Def. 1). If $f \in C^2(\mathbb{R}^d, \mathbb{R})$ has compact support and τ is a Markov time with $E^x\tau < \infty$, then*

$$E^x f(X_\tau) = f(x) + E^x \int_0^\tau Af(X_s)ds$$

Dynkin's formula is very useful, but only if we can compute the generator and its integral. The generator A gives a stochastic expression. It has been shown, however, that it is equal to a deterministic expression called the differential generator under fairly mild assumptions:

Theorem 4 (Differential generator [Øks07, Theorem 7.3.3]) *For a solution X_t from Def. 1, if $f \in C^2(\mathbb{R}^d, \mathbb{R})$ is compactly supported, then Af is defined and*

$$Af(x) = Lf(x) := \sum_i b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j}(\sigma(x)\sigma(x)^*)_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

A stochastic process Y that is adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$ is a *supermartingale* iff $E|Y_t| < \infty$ for all $t \geq 0$ and

$$E(Y_t \mid \mathcal{F}_s) \leq Y_s \quad \text{for all } t \geq s \geq 0$$

Proposition 1 (Doob's maximal martingale inequality [KS91, Theorem I.3.8]) *If $f(X_t)$ is a càdlàg supermartingale with respect to the filtration generated by $(X_t)_{t\geq 0}$ and $f \geq 0$ on the evolution domain of X_t, then for all $\lambda > 0$:*

$$P\left(\sup_{t\geq 0} f(X_t) \geq \lambda \,|\, \mathcal{F}_0\right) \leq \frac{Ef(X_0)}{\lambda}$$

7 Proof Calculus

Now that we have a model, logic, and semantics for stochastic hybrid systems, we investigate reasoning principles that can be used to prove logical properties of stochastic hybrid systems. First we present proof rules that are sound pathwise, i.e., satisfy the global consequence relation pathwise for each $\omega \in \Omega$. By $\sqcup$ we denote the binary maximum operator. It can either be added into the language or approximated conservatively by $+$ as in rule $\langle;\rangle$. Operator $\sqcup$ coincides with $\vee$ for values in $\{0,1\}$, e.g., built using operators $\wedge, \vee, \neg, \langle\alpha\rangle$ from characteristic functions. As a supremum, $\langle\alpha\rangle B$ only takes on values $\{0,1\}$ if B does.

Theorem 5 (Pathwise sound) *The proof rules in Fig. 3 are globally sound pathwise.*

For a proof see Appendix B.1. For $\langle;\rangle'$, β is a.s. continuous at 0 if, on all paths, the first primitive operation that is not a test is a stochastic differential equation, not a (random) assignment. Our rules generalize to the case of probabilistic assumptions. Note that formula $\overline{H} \rightarrow f \leq \lambda$ in *mon′* is equivalent to $\overline{H}f \leq \overline{H}\lambda$ but easier to read. If f is continuous, rule *mon′* is sound when replacing the topological closure $\overline{H}$ (which is computable by quantifier elimination) by H, because the inequality is weak.

Next we show proof rules that do not hold pathwise, but still in distribution.

Theorem 6 (Sound in distribution) *Rule $\langle\oplus\rangle$ is sound in distribution.*

$$P(\langle\lambda\alpha \oplus \nu\beta\rangle f \in S) = \lambda P(\langle\alpha\rangle f \in S) + \nu P(\langle\beta\rangle f \in S) \qquad (\langle\oplus\rangle)$$

$$\langle x := \theta \rangle f = f_x^\theta \quad \text{if admissible substitution replacing } x \text{ with } \theta \qquad (\langle := \rangle)$$
$$\langle ?H \rangle f = Hf \qquad (\langle ? \rangle)$$
$$\langle \alpha; \beta \rangle f \leq \langle \alpha \rangle (f \sqcup \langle \beta \rangle f) \quad \big(\leq \langle \alpha \rangle (f + \langle \beta \rangle f) \text{ if } 0 \leq f \big) \qquad (\langle ; \rangle)$$
$$\langle \alpha; \beta \rangle f \leq \langle \alpha \rangle \langle \beta \rangle f \quad \text{if } \vDash f \leq \langle \beta \rangle f \text{ or } \beta \text{ continuous at 0 a.s.} \qquad (\langle ; \rangle')$$
$$\langle \alpha \rangle (\lambda f) = \lambda \langle \alpha \rangle f \qquad (\langle \rangle \lambda)$$
$$\langle \alpha \rangle (\lambda f + \nu g) \leq \lambda \langle \alpha \rangle f + \nu \langle \alpha \rangle g \qquad (\langle \rangle +)$$
$$0 \leq B = BB \leq 1 \quad \text{if } B \text{ boolean from characteristic functions} \qquad (\mathcal{I})$$
$$0 \leq f \ \vDash \ 0 \leq \langle \alpha \rangle f \qquad (\textit{pos})$$
$$f \leq g \ \vDash \ \langle \alpha \rangle f \leq \langle \alpha \rangle g \qquad (\textit{mon})$$
$$\overline{H} \to f \leq \lambda \ \vDash \ \langle dx = bdt + \sigma dW \,\&\, H \rangle f \leq \lambda \quad (\lambda \in \mathbb{Q}) \qquad (\textit{mon}')$$
$$\langle \alpha \rangle g \leq g \ \vDash \ \langle \alpha^* \rangle g \leq g \qquad (\textit{ind})$$

Figure 3: Pathwise proof rules for $\mathsf{Sd}\mathcal{L}$

For a proof see Appendix B.2. How to prove properties about random assignment $x_i := *$ depends on the distribution for the random assignment. For a uniform distribution in [0,1], e.g., we obtain the following proof rule that is sound in distribution:

$$P(\langle x_i := * \rangle f \in S) = \int_0^1 \mathcal{I}_{\langle x_i := r \rangle f \in S} dr \qquad (\langle * \rangle)$$

The integrand is measurable for measurable S by Corollary 1. The rule is applicable when f has been simplified enough using other proof rules such that the integral can be computed after using $\langle := \rangle$ to simplify the integrand.

Theorem 7 (Soundness for stochastic differential equations) *If function* $\mathsf{f} \in C^2(\mathbb{R}^d, \mathbb{R})$ *has compact support on* H *(which holds for all* $\mathsf{f} \in C^2(\mathbb{R}^d, \mathbb{R})$ *if* H *represents a bounded set), then the proof rule* $\langle' \rangle$ *is sound for* $\lambda > 0, p \geq 0$

$$(\langle' \rangle) \quad \frac{\langle \alpha \rangle (H \to \mathsf{f}) \leq \lambda p \quad H \to \mathsf{f} \geq 0 \quad H \to L\mathsf{f} \leq 0}{P(\langle \alpha \rangle \langle dx = bdt + \sigma dW \,\&\, H \rangle \mathsf{f} \geq \lambda) \leq p}$$

Proof: Since f has compact support on H, it has a $C^2(\mathbb{R}^d, \mathbb{R})$ modification with compact support on $\mathbb{R}^d$ that still satisfies the premises of $\langle' \rangle$, because all properties of f in the premises assume H. To simplify notation, we write $\mathsf{f}(x)$ for $[\![\mathsf{f}]\!]^x$. Let X_t be the stochastic process $[\![dx = bdt + \sigma dW \,\&\, H]\!]^Z$. Let $\check{X}_t$ be X_t restricted to H, i.e., the stopped process $\check{X}_t := X_{t \sqcap (\!| dx = bdt + \sigma dW \,\&\, H |\!)^Z}$, which is stopped at a Markov time by Theorem 1. The stopped process $\check{X}_t$, thus, inherits càdlàg and strong Markov properties from X_t; see, e.g., [Dyn65, Theorem 10.2]. If $A\mathsf{f}$ is defined and continuous and bounded on H [Dyn65, Ch 11.3][Kus67, Ch I.3,I.4], then the infinitesimal generator of $\check{X}_t$ agrees with the generator of X_t on H (and is zero otherwise). This is the case, since $\mathsf{f} \in C^2(\mathbb{R}^d, \mathbb{R})$

has compact support (thus bounded as continuous), because $A\mathsf{f}$ is then defined and $A\mathsf{f} = L\mathsf{f}$ by Theorem 4, hence, $L\mathsf{f}$ is continuous, because b, σ are continuous by Def. 1.

All premises of rule $\langle' \rangle$ still hold when assuming the topological closure $\overline{H}$ instead of H, because the functions f and $L\mathsf{f}$ are continuous and the conditions are weak inequalities, thus, closed. Consider any $x \in \mathbb{R}^d$ and any time $s \geq 0$. The deterministic time s is a (very simple) Markov time with $E^x s = s < \infty$. Since f is compactly supported, Theorem 3 is applicable and implies that

$$E^x \mathsf{f}(\check{X}_s) = \mathsf{f}(x) + E^x \int_0^s A\mathsf{f}(\check{X}_r) dr \tag{2}$$

Now $L\mathsf{f} \leq 0$ on $\overline{H}$ by the third premise. Hence, $A\mathsf{f} \leq 0$ on $\overline{H}$, because $L\mathsf{f} = A\mathsf{f}$ (on $\overline{H}$) by Theorem 4, as $\mathsf{f} \in C^2(\mathbb{R}^d, \mathbb{R})$ has compact support. Because X and $\check{X}$ have a.s. continuous paths *and* are not defined on the event $\{Z \not\models H\}$, we know that $\check{X}_s$ stays in the closure $\overline{H}$ a.s. Thus, $A\mathsf{f}(\check{X}_s) \leq 0$ a.s., hence, $\int_0^s A\mathsf{f}(\check{X}_r) dr \leq 0$ a.s., thus, $E^x \int_0^s A\mathsf{f}(\check{X}_r) dr \leq 0$. Then (2) implies $E^x \mathsf{f}(\check{X}_s) \leq \mathsf{f}(x)$ for all x.

Because the filtration is right-continuous and $\mathsf{f} \in C(\mathbb{R}^d, \mathbb{R})$ is compactly supported (hence bounded), the strong Markov property [KS91, Prop 2.6.7] for $\check{X}_t$ implies for all $t \geq s \geq 0$ that P^x-a.s.: $E^x(\mathsf{f}(\check{X}_t)|\mathcal{F}_s) = E^{\check{X}_s} \mathsf{f}(\check{X}_{t-s}) \leq \mathsf{f}(\check{X}_s)$. The inequality holds, since $E^x \mathsf{f}(\check{X}_s) \leq \mathsf{f}(x)$ for all x, s. Thus, $\mathsf{f}(\check{X}_t)$ is a supermartingale with respect to $\check{X}_t$, because it is adapted to the filtration of $\check{X}_t$ (as $\mathsf{f} \in C^2(\mathbb{R}^d, \mathbb{R})$) and $E^x|\mathsf{f}(\check{X}_t)| < \infty$ for all t since $\mathsf{f} \in C^2(\mathbb{R}^d, \mathbb{R})$ has compact support. Further, $\mathsf{f}(\check{X}_t)$ inherits continuity from $\check{X}_t$ (which follows from X_t), since f is continuous.

Thus, by the second premise, Proposition 1 is applicable. Consider any initial state $Y := [\![\alpha]\!]_t^Z$ for $\check{X}$. Thus, $P\left(\sup_{t \geq 0} \mathsf{f}(\check{X}_t) \geq \lambda \,|\, \mathcal{F}_0\right) \leq \frac{E\mathsf{f}(Y)}{\lambda}$ by Proposition 1 (filtration at $\check{X}_0$ is $\mathcal{F}_0$). On event $\{Y \not\models H\}$, $\check{X}$ is not defined and nothing to show. On $\{Y \models H\}$, $\mathsf{f}(Y) \leq \lambda p$ is *valid* where relevant by the first premise. This implies the conclusion, as $[\![\langle dx = bdt + \sigma dW \,\&\, H\rangle \mathsf{f}]\!]^Y = \sup_{t \geq 0} \mathsf{f}(\check{X}_t)$. □

The implications in the premises can be understood like that in *mon*′. Let H be given by first-order real arithmetic formulas. If f is polynomial and, thus, $\mathsf{f} \in C^2(\mathbb{R}^d, \mathbb{R})$, then the second and third premise of $\langle' \rangle$ are in first-order real arithmetic, hence decidable. Note that our proof rules can be generalized to probabilistic assumptions by the rule of partition and then combined.

The proof shows that it is enough to assume the first premise holds only a.s. From the proof we see that it would be sufficient to replace the third premise of $\langle' \rangle$ with $\int_0^s L\mathsf{f}(X_r) dr \leq 0$. This is a weaker condition, because it does not require $L\mathsf{f} \leq 0$ always, but only "on average". But this condition is computationally more involved, because the integral needs to be computed first. For polynomial expressions, this is not too difficult, but still increases the polynomial degree.

A simple two-dimensional example is the following for $H \equiv x^2 + y^2 < 10$:

$$P(\langle ?x^2{+}y^2{\leq}\frac{1}{3}; x := \frac{x}{2}; dx = \frac{-x}{2}dt - ydW, dy = \frac{-y}{2}dt + xdW \,\&\, H\rangle x^2{+}y^2{\geq}1) \leq \frac{1}{3}$$

which can be proven easily using $\langle ; \rangle'$, $\langle ; \rangle \langle ? \rangle$, $\langle := \rangle$, $\langle' \rangle$, since $\mathsf{f} \equiv x^2 + y^2 \geq 0$ and

$$L\mathsf{f} = \frac{1}{2}\left(-x\frac{\partial \mathsf{f}}{\partial x} - y\frac{\partial \mathsf{f}}{\partial y} + y^2\frac{\partial^2 \mathsf{f}}{\partial x^2} - 2xy\frac{\partial^2 \mathsf{f}}{\partial x \partial y} + x^2\frac{\partial^2 \mathsf{f}}{\partial y^2}\right) \leq 0$$

This implies the second and third premise of $\langle' \rangle$. In order to see why the first premise holds and how the property can be concluded, we first look at a simpler example.

$$P(\langle ?x^2+y^2 \leq \frac{1}{3};\ dx = -\frac{x}{2}dt - ydW, dy = -\frac{y}{2}dt + xdW \,\&\, H\rangle x^2+y^2 \geq 1) \leq \frac{1}{3}$$

The second and third premise of $\langle' \rangle$ continue to hold for this simpler example. We conclude the first premise of $\langle' \rangle$ using $\langle ? \rangle$

$$\langle ?x^2+y^2 \leq \frac{1}{3}\rangle(H \to \mathsf{f}) = \left(H \to x^2+y^2 \leq \frac{1}{3}\right)(x^2+y^2) \leq 1 * \frac{1}{3}$$

Hence, $\langle' \rangle$ is applicable implying the conclusion

$$P(\langle ?x^2+y^2 \leq \frac{1}{3};\ dx = -\frac{x}{2}dt - ydW, dy = -\frac{y}{2}dt + xdW \,\&\, H\rangle x^2+y^2 \geq 1)$$

Using $\langle ; \rangle'$ inside the probability, this expression is $\leq$ the following

$$P(\langle ?x^2+y^2 \leq \frac{1}{3}\rangle\langle dx = -\frac{x}{2}dt - ydW, dy = -\frac{y}{2}dt + xdW \,\&\, H\rangle x^2+y^2 \geq 1) \leq \frac{1}{3}$$

In the same way, we can prove the original property:

$$P(\langle ?x^2+y^2 \leq \frac{1}{3};\ x := \frac{x}{2};\ dx = -\frac{x}{2}dt - ydW, dy = -\frac{y}{2}dt + xdW \,\&\, H\rangle x^2+y^2 \geq 1) \leq \frac{1}{3}$$

The only change is as follows. By $\langle ; \rangle$ we conclude

$$\langle ?x^2+y^2 \leq \frac{1}{3};\ x := \frac{x}{2}\rangle(H \to \mathsf{f}) \leq \langle ?x^2+y^2 \leq \frac{1}{3}\rangle\big((H \to \mathsf{f}) \sqcup \langle x := \frac{x}{2}\rangle(H \to \mathsf{f})\big)$$

which, by $\langle := \rangle$, is $\leq$ the following, because $x := \frac{x}{2}$ makes the f-value drop (and $?x^2+y^2 \leq \frac{1}{3}$ implies H even after $x := \frac{x}{2}$):

$$\langle ?x^2+y^2 \leq \frac{1}{3}\rangle(H \to \mathsf{f}) = \left(H \to x^2+y^2 \leq \frac{1}{3}\right)(x^2+y^2) \leq 1 * \frac{1}{3}$$

The arithmetic is easily decidable by quantifier-elimination in real-closed fields.

8 Related Work

Our approach is partially inspired by the work of Kozen, who studied 3 semantics of programs with random number generators [Koz81] and probabilistic PDL [Koz85]. We generalize from discrete systems to stochastic hybrid systems. To reflect the new challenges, we have departed from probabilistic PDL. Kozen uses a measure semantics. We choose a semantics that is based on stochastic processes, because the temporal behavior of SHPs is more crucial than that of abstract

discrete programs. SdL further uses a supremal semantics that is more interesting for stochastic worst-case verification than the integral semantics assumed in [Koz85].

The comparison to a first-order dynamic logic for deterministic programs with random number generators [FH84] is similar. They axiomatize relative to first-order analysis with arithmetic, enriched with frequencies and random number generators. They do not show how this logic could be handled (incompletely).

Our approach for stochastic differential equations is inspired by barrier certificates [PJP07]. We extend this work by identifying the assumptions that are required for soundness of using Dynkin-type arguments for stochastic differential equations. They propose to use global generators for switching diffusion processes (which cannot reset variables). We use logic and compositional proofs for SHPs.

Probabilities and logic have also been used in AI, e.g., [RD06]. Markov logic networks are a combination of Markov networks and first-order logic and resembles logic programming with weights for probabilities. They are restricted to finite domains, which is not the case in stochastic hybrid systems.

Model checking has been used for discrete probabilistic systems like finite Markov chains, e.g., [YKNP06], and probabilistic timed automata [KNSW07]. Assume-guarantee model checking is a challenge for discrete probabilistic automata, with recent successes for finite automata assumptions [KNPQ10]. We use a compositional proof approach based on logic and consider stochastic hybrid systems.

Statistical model checking has been suggested for validating stochastic hybrid systems [MS06] and later refined for discrete-time hybrid systems with a probabilistic simulation function [ZPC10] based on corresponding discrete probabilistic techniques [YKNP06]. They did not show measurability and do not support stochastic differential equations [ZPC10]. Validation by simulation is generally unsound, but the probability of giving a wrong answer can sometimes be bounded [YKNP06, ZPC10].

Fränzle et al. [FTE10] show first pieces for continuous-time bounded model checking of probabilistic hybrid automata (no stochastic differential equations).

Bujorianu and Lygeros [BL06] show strong Markov and càdlàg properties for a class of systems known as General Stochastic Hybrid Systems. They also study an interesting concatenation operator. For an overview of model checking techniques for various classes of stochastic hybrid systems, we refer to [CL06]. Most verification techniques for stochastic hybrid systems use discretizations, approximations, or assume discrete time, bounded horizon [KR08, CL06, APLS08, HLS00]. We consider the continuous-time behavior and develop compositional logic and theorem proving.

9 Conclusions

We introduce the first verification logic for stochastic hybrid systems along with a compositional model of stochastic hybrid programs. We prove theoretical properties that are important for well-definedness and measurability and we develop a compositional proof calculus. Our logic makes the complexity of stochastic hybrid systems accessible in logic with simple syntactic proof principles.

Our results indicate that $\mathsf{Sd}\mathcal{L}$ is a promising starting point for the study of logic for stochastic hybrid systems. Extensions include nondeterminism.

Acknowledgments. I thank the anonymous referees of the conference version [Pla11] for their good comments. I also want to thank Steve Marcus and Sergio Pulido Niño for helpful discussions.

References

[APLS08] Alessandro Abate, Maria Prandini, John Lygeros, and Shankar Sastry. Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica*, 44(11):2724–2734, 2008.

[BL06] Manuela L. Bujorianu and John Lygeros. Towards a general theory of stochastic hybrid systems. In Henk A. P. Blom and John Lygeros, editors, *Stochastic Hybrid Systems: Theory and Safety Critical Applications*, volume 337 of *Lecture Notes Contr. Inf.*, pages 3–30. Springer, 2006.

[CL06] Christos G. Cassandras and John Lygeros, editors. *Stochastic Hybrid Systems*. CRC, 2006.

[Dut95] Bruno Dutertre. Complete proof systems for first order interval temporal logic. In *LICS*, pages 36–43. IEEE Computer Society, 1995.

[Dyn65] Evgenij Borisovic Dynkin. *Markov Processes*. Springer, 1965.

[FH84] Yishai A. Feldman and David Harel. A probabilistic dynamic logic. *J. Comput. Syst. Sci.*, 28(2):193–215, 1984.

[FTE10] Martin Fränzle, Tino Teige, and Andreas Eggers. Engineering constraint solvers for automatic analysis of probabilistic hybrid automata. *J. Log. Algebr. Program.*, 79(7):436–466, 2010.

[GAM97] Mrinal K. Ghosh, Aristotle Arapostathis, and Steven I. Marcus. Ergodic control of switching diffusions. *SIAM J. Control Optim.*, 35(6):1952–1988, 1997.

[HLS00] Jianghai Hu, John Lygeros, and Shankar Sastry. Towards a theory of stochastic hybrid systems. In Nancy A. Lynch and Bruce H. Krogh, editors, *HSCC*, volume 1790 of *LNCS*, pages 160–173. Springer, 2000.

[KNPQ10] Marta Z. Kwiatkowska, Gethin Norman, David Parker, and Hongyang Qu. Assume-guarantee verification for probabilistic systems. In Javier Esparza and Rupak Majumdar, editors, *TACAS*, volume 6015 of *LNCS*, pages 23–37. Springer, 2010.

[KNSW07] Marta Z. Kwiatkowska, Gethin Norman, Jeremy Sproston, and Fuzhi Wang. Symbolic model checking for probabilistic timed automata. *Inf. Comput.*, 205(7):1027–1077, 2007.

[Koz81] Dexter Kozen. Semantics of probabilistic programs. *J. Comput. Syst. Sci.*, 22(3):328–350, 1981.

[Koz85] Dexter Kozen. A probabilistic PDL. *J. Comput. Syst. Sci.*, 30(2):162–178, 1985.

[KP10] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, New York, 2010.

[KR08] Xenofon D. Koutsoukos and Derek Riley. Computational methods for verification of stochastic hybrid systems. *IEEE T. Syst. Man, Cy. A*, 38(2):385–396, 2008.

[KS91] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1991.

[Kus67] Harold J. Kushner. *Stochastic Stability and Control*. Academic Press, 1967.

[MS06] José Meseguer and Raman Sharykin. Specification and analysis of distributed object-based stochastic hybrid systems. In João P. Hespanha and Ashish Tiwari, editors, *HSCC*, volume 3927 of *LNCS*, pages 460–475. Springer, 2006.

[Øks07] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, 2007.

[PJP07] Stephen Prajna, Ali Jadbabaie, and George J. Pappas. A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE T. Automat. Contr.*, 52(8):1415–1429, 2007.

[Pla10a] André Platzer. Differential-algebraic dynamic logic for differential-algebraic programs. *J. Log. Comput.*, 20(1):309–352, 2010.

[Pla10b] André Platzer. Quantified differential dynamic logic for distributed hybrid systems. In Anuj Dawar and Helmut Veith, editors, *CSL*, volume 6247 of *LNCS*, pages 469–483. Springer, 2010.

[Pla11] André Platzer. Stochastic differential dynamic logic for stochastic hybrid programs. In Nikolaj Bjørner and Viorica Sofronie-Stokkermans, editors, *CADE*, LNCS. Springer, 2011.

[Pra76] Vaughan R. Pratt. Semantical considerations on Floyd-Hoare logic. In *FOCS*, pages 109–121. IEEE, 1976.

[RD06] Matthew Richardson and Pedro Domingos. Markov logic networks. *Machine Learning*, 62(1-2):107–136, 2006.

[Tar51] Alfred Tarski. *A Decision Method for Elementary Algebra and Geometry*. University of California Press, Berkeley, 2nd edition, 1951.

[Wal95] Wolfgang Walter. *Analysis 2*. Springer, 4 edition, 1995.

[YKNP06] Håkan L. S. Younes, Marta Z. Kwiatkowska, Gethin Norman, and David Parker. Numerical vs. statistical probabilistic model checking. *STTT*, 8(3):216–228, 2006.

[ZPC10] Paolo Zuliani, André Platzer, and Edmund M. Clarke. Bayesian statistical model checking with application to Simulink/Stateflow verification. In Karl Henrik Johansson and Wang Yi, editors, *HSCC*, pages 243–252. ACM, 2010.

A Proofs for Semantics

In the appendices, we provide proofs for the results in this paper. In this appendix, we provide proofs for the semantics and its well-definedness.

A.1 Proof of Adaptive Càdlàg Process with Markov Times

Proof(of Theorem 1): We prove càdlàg, adaptedness, and Markov time properties simultaneously by induction on the structure of α. These parts partially depend on each other, so we prove them together not separately. To simplify notation, we shift time so that processes start at time 0.

1–3. Deterministic times $(\!|x_i := \theta|\!)^Z = (\!|x_i := *|\!)^Z = (\!|?H|\!)^Z = 0$ are trivial Markov times. Furthermore, the process $[\![x_i := \theta]\!]^Z$ is adapted to the filtration generated by Z. Process $[\![?H]\!]^Z$ is also adapted if it is defined (otherwise there is nothing to show). Similarly, $[\![x_i := *]\!]^Z$ is adapted to the filtration generated by Z and the u.i.i.d. random variable $([\![x_i := *]\!]^Z_0)_i = U_i$. Process $[\![?H]\!]^Z$ is càdlàg (even constant) if it is defined, otherwise there is no continuity question (can be considered stuck at absorbing state). Processes $[\![x_i := \theta]\!]^Z$ and $[\![x_i := *]\!]^Z$ are trivially càdlàg (even continuous) as the time domain is $\{0\}$.

4. $(\!|dx = bdt + \sigma dW \,\&\, H|\!)^Z = \inf\{t \geq 0 : X_t \notin H\}$ is a Markov time when H is any Borel set [Øks07, Ex 7.2.2][Dyn65, Vol. II, 4.5.C.e], since we complete the filtration to include all null sets. Here X_t is the process $[\![dx = bdt + \sigma dW \,\&\, H]\!]^Z$. More generally, for progressively measurable processes like right-continuous adapted processes, the hitting time of a measurable set is a Markov time by the (deep) début theorem [Øks07]. Solutions of stochastic differential equations are adapted to the filtration generated by $(W_s)_{s\leq t}$ and Z [Øks07, Th 5.2.1][KP10, Ch 4.5] and have almost surely continuous paths by a consequence of Kolmogorov's continuity theorem [Øks07, Th 2.2.3].

5. By induction hypothesis, $(\!|\alpha|\!)^Z$ is a Markov time, hence $\mathcal{I}_{U\leq\lambda}(\!|\alpha|\!)^Z$ is a Markov time, since the filtration includes U and the indicator function only takes on values 0 (where 0 is a stopping time) or 1 (where $1(\!|\alpha|\!)^Z$ is a Markov time). Similarly $\mathcal{I}_{U>\lambda}(\!|\beta|\!)^Z$ is a Markov time. As the sum of two Markov times, $(\!|\lambda\alpha \oplus \nu\beta|\!)^Z$ is a Markov time [KS91, Lem 1.2.9]. Because càdlàg functions form an algebra (Skorokhod space), the linear combination $[\![\lambda\alpha \oplus \nu\beta]\!]^Z$ is càdlàg by induction hypothesis for every outcome of U. This linear combination is adapted, because, by induction hypothesis, the parts are adapted and the choice U generates the filtration.

6. By induction hypothesis, $[\![\alpha]\!]^Z$ is adapted to and $(\!|\alpha|\!)^Z$ a Markov time for the filtration $(\mathcal{F}'_t)_{t\geq 0}$ generated by Z and the constituent Brownian motion and uniform processes during α. Especially, $[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}$ is a random variable. By induction hypothesis, $\left([\![\beta]\!]^{[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}}_t\right)_{t\geq 0}$ is, thus, adapted to and $(\!|\beta|\!)^{[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}}$ a Markov time for the filtration $(\mathcal{F}''_t)_{t\geq 0}$ generated by $[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}$ and the constituent Brownian motion and uniform processes during β. With a time shift

by $-(\!|\alpha|\!)^Z$, $\left(\llbracket\beta\rrbracket^{\llbracket\alpha\rrbracket^Z_{(\!|\alpha|\!)^Z}}_{t-(\!|\alpha|\!)^Z}\right)_{t\geq(\!|\alpha|\!)^Z}$ is then adapted to the filtration $(\mathcal{F}''_{t-(\!|\alpha|\!)^Z})_{t\geq(\!|\alpha|\!)^Z}$. Especially, $(\mathcal{F}_t)_{t\geq 0}$ already includes $(\mathcal{F}'_t)_{t\geq 0}$ and the time-shifted $(\mathcal{F}''_{t-(\!|\alpha|\!)^Z})_{t\geq(\!|\alpha|\!)^Z}$. Note that random variable $\llbracket\alpha\rrbracket^Z_{(\!|\alpha|\!)^z}$ does not contribute to this filtration, because it is already $\mathcal{F}_{(\!|\alpha|\!)^z}$-measurable by induction hypothesis. Consequently, $\llbracket\alpha;\beta\rrbracket^Z$ is adapted to $(\mathcal{F}_t)_{t\geq 0}$, because both of its cases, $\llbracket\alpha\rrbracket^Z$ and $\llbracket\beta\rrbracket^{\llbracket\alpha\rrbracket^Z_{(\!|\alpha|\!)^Z}}_{t-(\!|\alpha|\!)^Z}$, are adapted and the condition which case applies is an event of a Markov time. Similarly, $(\!|\alpha;\beta|\!)^Z = (\!|\alpha|\!)^Z + (\!|\beta|\!)^{\llbracket\alpha\rrbracket^Z_{(\!|\alpha|\!)^Z}}$ is a sum of two Markov times and, thus, a Markov time [KS91, Lem 1.2.9].

   By induction hypothesis, $\llbracket\alpha;\beta\rrbracket^Z$ is càdlàg on $[0,(\!|\alpha|\!)^Z)$ and on $((\!|\alpha|\!)^Z,\infty)$, because the constituent fragments are. At $(\!|\alpha|\!)^Z$, process $\llbracket\alpha;\beta\rrbracket^Z$ is càdlàg, by construction (it is defined in terms of β on the left-closed interval $[(\!|\alpha|\!)^Z,\infty)$, hence càdlàg even if there is a jump before β starts).

7. Because $(\!|\alpha^n|\!)^Z$ are increasing, $(\!|\alpha^*|\!)^Z = \lim_{n\to\infty}(\!|\alpha^n|\!)^Z = \sup_{n\geq 1}(\!|\alpha^n|\!)^Z$, which is a Markov time [KS91, Lem 1.2.11], since, by induction hypothesis, the $(\!|\alpha^n|\!)^Z$ are Markov times. Process $\llbracket\alpha^*\rrbracket^Z$ is adapted, because for each t, the constituent process $\llbracket\alpha^n\rrbracket^Z_t$ is adapted on each event $\{(\!|\alpha^n|\!)^Z > t\}$ by induction hypothesis. Note that $\llbracket\alpha^*\rrbracket^Z$ is not defined if this never happens, i.e., on the event $\{\forall n\,(\!|\alpha^n|\!)^Z \geq t\}$. Since the value $\llbracket\alpha^*\rrbracket^Z_t$ is defined on an n that satisfies the open event $\{(\!|\alpha^n|\!)^Z > t\}$, the process is càdlàg as long as it is defined.

□

A.2 Proof of Measurability

Proof(of Theorem 2): We need to show that $\llbracket f\rrbracket^Z$ is measurable as a function of $\omega\in\Omega$. We prove this by induction on the structure of f.

1. $\llbracket F\rrbracket^Z = F^\ell(Z)$ is a random variable, because Z is measurable and F^ℓ is Borel(!)-measurable. Thus, the composition $F^\ell(Z)$ is measurable (the σ-algebras in the composition are compatible).

2. $\llbracket\lambda f+\nu g\rrbracket^Z = \lambda\llbracket f\rrbracket^Z + \nu\llbracket g\rrbracket^Z$ is a linear combination, hence, measurable by induction hypothesis, because measurable functions form an algebra.

3. $\llbracket Bf\rrbracket^Z = \llbracket B\rrbracket^Z * \llbracket f\rrbracket^Z$ is a product, hence, measurable by induction hypothesis, because measurable functions form an algebra.

4. $\llbracket\langle\alpha\rangle f\rrbracket^Z = \sup\{\llbracket f\rrbracket^{\llbracket\alpha\rrbracket^Z_t} : 0\leq t\leq(\!|\alpha|\!)^Z\}$ is measurable for the following reason. By Theorem 1, $\llbracket\alpha\rrbracket^Z_t$ is measurable (adapted). By induction hypothesis, $\llbracket f\rrbracket^{\llbracket\alpha\rrbracket^Z_t}$ is measurable for each t. We need to show that the supremum is still measurable. Unfortunately, suprema of measurable functions over uncountable sets are generally not measurable. Yet, the (pointwise) supremum of a countable sequence of measurable functions is measurable [Wal95,

§9.9]. Consider a rational mesh $\pi := \{t_1, t_2, \ldots, t_n\} \subset \mathbb{Q}$ with times $0 \leq t_1 \leq \cdots \leq t_n$. By induction hypothesis, $[\![f]\!]^{[\![\alpha]\!]^Z_t}$ is measurable for each $t \in \pi$. Hence, the (finite) countable supremum $\sup\{[\![f]\!]^{[\![\alpha]\!]^Z_t} : t \in \pi, t \leq (\!|\alpha|\!)^Z\}$ is measurable (as a pointwise function of $\omega \in \Omega$). Unlike the set of infinite sequences in $\mathbb{Q}$, the set of finite sequences in $\mathbb{Q}$ is countable. Thus, the countable supremum $\sup\{[\![f]\!]^{[\![\alpha]\!]^Z_t} : t \leq (\!|\alpha|\!)^Z, t \in \pi \text{ for a rational mesh } \pi\}$ is measurable, because the set of rational meshes is countable. In general, however, this latter supremum does not coincide with the supremum defining $[\![\langle\alpha\rangle f]\!]^Z$. But since $[\![\alpha]\!]^Z$ is also càdlàg a.s. by Theorem 1, they do coincide (each path is a.s. right-continuous). Note that either left or right continuity would be sufficient to ensure that there is a convergent sequence of rational meshes whose values converge to the value at each real point in the interval. Note, however, that this only gives us information about the supremum on $0 \leq t < (\!|\alpha|\!)^Z$ for a right continuous process, because $(\!|\alpha|\!)^Z$ could be irrational and no convergent sequence of rational points $t_i \geq (\!|\alpha|\!)^Z$ from the right is in the interval. But, when taking the (binary) pointwise supremum of $[\![f]\!]^{[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}}$ and the above supremum, we obtain the desired equality.

□

B Soundness Proofs

In this appendix, we provide proofs for the soundness theorems.

B.1 Proof of Pathwise Global Soundness

Proof(of Theorem 5): We prove that the rules are globally sound pathwise (which coincides with locally sound if they have no assumptions) by showing that they hold for any $\mathbb{R}^d$-valued random variable Z pathwise, i.e., on every path for every $\omega \in \Omega$.

$\langle:=\rangle$ Soundness of rule $\langle:=\rangle$ is similar to classical dynamic logic [Pra76]. That is, $[\![\langle x := \theta\rangle f]\!]^Z = [\![f]\!]^{[\![x:=\theta]\!]^Z_0} = [\![f^\theta_x]\!]^Z$ deterministically (for all $\omega \in \Omega$). Note that the supremum disappears, because of $(\!|x := \theta|\!)^Z = 0$.

$\langle?\rangle$ $[\![Hf]\!]^Z = [\![H]\!]^Z * [\![f]\!]^Z$ is equal to

$$[\![\langle ?H\rangle f]\!]^Z = \sup\{[\![f]\!]^{[\![?H]\!]^Z_t} : 0 \leq t \leq (\!|?H|\!)^Z\} = \begin{cases} [\![f]\!]^Z & \text{on event } \{Z \models H\} \\ 0 & \text{on event } \{Z \not\models H\} \end{cases}$$

because $(\!|?H|\!)^Z = 0$ (on all events) and our convention evaluates all function terms f to 0 in undefined states (on the event that $?H$ fails by $\{Z \not\models H\}$).

$\langle;\rangle$ $[\![\langle\alpha;\beta\rangle f]\!]^Z = \sup\{[\![f]\!]^{[\![\alpha;\beta]\!]^Z_r} : 0 \leq r \leq (\!|\alpha;\beta|\!)^Z = (\!|\alpha|\!)^Z + (\!|\beta|\!)^{[\![\alpha]\!]^Z_{(\!|\alpha|\!)^Z}}\}$. Also $[\![\langle\alpha\rangle(f \sqcup \langle\beta\rangle f)]\!]^Z = \sup\{[\![f \sqcup \langle\beta\rangle f]\!]^{[\![\alpha]\!]^Z_t} : 0 \leq t \leq (\!|\alpha|\!)^Z\}$. The latter equals $\sup\left\{[\![f]\!]^{[\![\alpha]\!]^Z_t} \sqcup \sup\{[\![f]\!]^{[\![\beta]\!]^{[\![\alpha]\!]^Z_t}_s} : 0 \leq\right.$

$s \leq (\!|\beta|\!)^{[\![\alpha]\!]_t^Z}\} : 0 \leq t \leq (\!|\alpha|\!)^Z\}$. With these expansions, $[\![\langle\alpha;\beta\rangle f]\!]^Z \leq [\![\langle\alpha\rangle(f \sqcup \langle\beta\rangle f)]\!]^Z$ holds as follows. For each path, the values of $[\![f]\!]^{[\![\alpha;\beta]\!]_r^Z}$ on the event $\{r \geq (\!|\alpha|\!)^Z\}$ are included in the nested supremum for $[\![\langle\alpha\rangle(f \sqcup \langle\beta\rangle f)]\!]^Z$ by choosing $t := (\!|\alpha|\!)^Z, s := r - (\!|\alpha|\!)^Z$. The values of $[\![f]\!]^{[\![\alpha;\beta]\!]_r^Z}$ on the event $\{r < (\!|\alpha|\!)^Z\}$ are included in the nested supremum by choosing $t := r$ and the left side of the maximum $[\![f]\!]^{[\![\alpha]\!]_t^Z} \sqcup \ldots$ in the expression. Note that the two sides are generally not equal, because α has to run to completion before β starts in $\langle\alpha;\beta\rangle f$, but α can stop early in $\langle\alpha\rangle(f \sqcup \langle\beta\rangle f)$ and β can then start already.

If, in addition, $\vDash 0 \leq f$, then $\vDash 0 \leq \langle\beta\rangle f$ by *pos*. Hence, *mon* implies by the semantics of $\sqcup$ that $\vDash \langle\alpha;\beta\rangle f \leq \langle\alpha\rangle(f \sqcup \langle\beta\rangle f) \leq \langle\alpha\rangle(f + \langle\beta\rangle f)$.

$\langle;\rangle'$ If $\vDash f \leq \langle\beta\rangle f$, then $\langle;\rangle'$ follows from $\langle;\rangle$ directly. If, instead, $[\![\beta]\!]^Y$ is continuous at 0 a.s., then the proof for $\langle;\rangle$ does not need $f\sqcup$. It can use $t := r, s := 0$ on the event $\{r < (\!|\alpha|\!)^Z\}$, because the process for β a.s. will not change the value of f at time 0 (a.s. continuity). The proof of $\langle;\rangle$ for event $\{r \geq (\!|\alpha|\!)^Z\}$ does not use $f\sqcup$ and carries over to $\langle;\rangle'$ directly.

$\langle\rangle\lambda$ $[\![\langle\alpha\rangle(\lambda f)]\!]^Z = \sup\{[\![\lambda f]\!]^{[\![\alpha]\!]_t^Z} : 0 \leq t \leq (\!|\alpha|\!)^Z\} = \sup\{\lambda[\![f]\!]^{[\![\alpha]\!]_t^Z} : 0 \leq t \leq (\!|\alpha|\!)^Z\} = \lambda[\![\langle\alpha\rangle f]\!]^Z$.

$\langle\rangle+$ $[\![\langle\alpha\rangle(\lambda f + \nu g)]\!]^Z = \sup\{[\![\lambda f + \nu g]\!]^{[\![\alpha]\!]_t^Z} : 0 \leq t \leq (\!|\alpha|\!)^Z\}$. This is equal to $\sup\{\lambda[\![f]\!]^{[\![\alpha]\!]_t^Z} + \nu[\![g]\!]^{[\![\alpha]\!]_t^Z} : 0 \leq t \leq (\!|\alpha|\!)^Z\} \leq \lambda[\![\langle\alpha\rangle f]\!]^Z + \nu[\![\langle\alpha\rangle g]\!]^Z$. The two sides are not equal if the suprema $[\![\langle\alpha\rangle f]\!]^Z$ and $[\![\langle\alpha\rangle g]\!]^Z$ are at different times.

$\mathcal{I}$ B is a Boolean combination of characteristic functions of measurable sets. Characteristic functions only take on the values 0 or 1, for which $\mathcal{I}$ holds. Boolean combinations preserve this property.

pos Rule *pos* is derivable from *mon* and $\langle\rangle\lambda$. By *mon*, $0 \leq f \vDash \langle\alpha\rangle 0 \leq \langle\alpha\rangle f$. By *mon*, $\langle\alpha\rangle 0 = \langle\alpha\rangle(0 * 0) = 0\langle\alpha\rangle$

mon Let $\vDash f \leq g$, i.e., $[\![f]\!]^Y \leq [\![g]\!]^Y$ for all Y. Hence, by Theorem 1, for random variable $Y := [\![\alpha]\!]_t^Z$, we get $[\![f]\!]^{[\![\alpha]\!]_t^Z} \leq [\![g]\!]^{[\![\alpha]\!]_t^Z}$. Since t is arbitrary, this implies

$$[\![\langle\alpha\rangle f]\!]^Z = \sup\{[\![f]\!]^{[\![\alpha]\!]_t^Z} : 0 \leq t \leq (\!|\alpha|\!)^Z\} \leq \sup\{[\![g]\!]^{[\![\alpha]\!]_t^Z} : 0 \leq t \leq (\!|\alpha|\!)^Z\} = [\![\langle\alpha\rangle g]\!]^Z$$

Hence, $[\![\langle\alpha\rangle f]\!]^Z \leq [\![\langle\alpha\rangle g]\!]^Z$, which implies *mon* since Z was arbitrary.

ind Assume $\vDash \langle\alpha\rangle g \leq g$, which implies $\vDash \langle\alpha;\alpha\rangle g \leq \langle\alpha\rangle(g \sqcup \langle\alpha\rangle g) = \langle\alpha\rangle g \leq g$ by $\langle;\rangle$. By induction, $\vDash \langle\alpha^n\rangle g \leq g$. Since $n \in \mathbb{N}$ was arbitrary, we get $\vDash \langle\alpha^*\rangle g \leq g$.

?? Assume $\vDash 0 \leq f$ and $\vDash 0 \leq f + \langle\alpha\rangle g \leq g$. First note that $\vDash 0 \leq f + \langle\alpha\rangle g \leq g$ directly implies $\vDash 0 \leq g$, which implies $0 \leq \langle\alpha\rangle g$ by *pos*, which implies $\vDash f \leq g$ using $\vDash 0 \leq f + \langle\alpha\rangle g \leq g$. Therefore, *mon* implies $\vDash \langle\alpha^*\rangle f \leq \langle\alpha^*\rangle g$. Now, $\vDash 0 \leq f$ and $\vDash 0 \leq f + \langle\alpha\rangle g \leq g$ together imply $\vDash \langle\alpha\rangle g \leq g$. Hence, *ind* implies $\vDash \langle\alpha^*\rangle g \leq g$. Together with $\vDash \langle\alpha^*\rangle f \leq \langle\alpha^*\rangle g$, this implies $\vDash \langle\alpha^*\rangle f \leq g$.

mon′ Assume $\vDash \overline{H} \to f \leq \lambda$. Let X_t be the stochastic process $[\![dx = bdt + \sigma dW \,\&\, H]\!]^Z$. Let $\check{X}_t$ be X_t restricted to H, i.e., $\check{X}_t := X_{t \sqcap (\!| dx=bdt+\sigma dW \,\&\, H |\!)^Z}$, which is stopped at a Markov time by Theorem 1. Because X (and, thus, $\check{X}$) have a.s. continuous paths *and* are not defined on the event $\{Z \not\models H\}$, we know that $\check{X}_s$ stays in the closure $\overline{H}$ a.s. Thus, $\check{X}_t[t] \models \overline{H}$ a.s. for all t. Hence, by assumption $[\![f]\!]^{\check{X}_t} \leq \lambda$ for all t. Then $[\![\langle dx = bdt + \sigma dW \,\&\, H\rangle f]\!]^Z \leq \lambda = [\![\lambda]\!]^Z$.

□

B.2 Proof of Soundness in Distribution

Proof(of Theorem 6): $[\![\langle \lambda\alpha \,\oplus\, \nu\beta\rangle f]\!]^Z = \sup\{[\![f]\!]^{\mathcal{I}_{U\leq\lambda}[\![\alpha]\!]^Z_t + \mathcal{I}_{U>\lambda}[\![\beta]\!]^Z_t} : 0 \leq t \leq (\!|\lambda\alpha \,\oplus\, \nu\beta|\!)^Z\}$, with $(\!|\lambda\alpha \,\oplus\, \nu\beta|\!)^Z = \mathcal{I}_{U\leq\lambda}(\!|\alpha|\!)^Z + \mathcal{I}_{U>\lambda}(\!|\beta|\!)^Z$. This expression splits into two disjoint events, one with $\{U \leq \lambda\}$ and one with $\{U > \lambda\}$. Thus, by additivity for disjoint events:

$$
\begin{aligned}
& P([\![\langle \lambda\alpha \,\oplus\, \nu\beta\rangle f]\!]^Z \in S) \\
=\; & P(U \leq \lambda,\ \sup\{[\![f]\!]^{[\![\alpha]\!]^Z_t} : 0 \leq t \leq (\!|\alpha|\!)^Z\} \in S) \\
& + P(U > \lambda,\ \sup\{[\![f]\!]^{[\![\beta]\!]^Z_t} : 0 \leq t \leq (\!|\beta|\!)^Z\} \in S) && \sigma\text{-additive} \\
=\; & P(U \leq \lambda,\ [\![\langle\alpha\rangle f]\!]^Z \in S) + P(U > \lambda,\ [\![\langle\beta\rangle f]\!]^Z \in S) \\
=\; & P(U \leq \lambda)P([\![\langle\alpha\rangle f]\!]^Z \in S) + P(U > \lambda)P([\![\langle\beta\rangle f]\!]^Z \in S) && \text{independent} \\
=\; & \lambda P([\![\langle\alpha\rangle f]\!]^Z \in S) + \nu P([\![\langle\beta\rangle f]\!]^Z \in S) && \lambda + \mu = 1
\end{aligned}
$$

□