Reproducibility Checklist

This paper

- Includes a conceptual outline and/or pseudocode description of AI methods introduced. *Yes*
- Clearly delineates statements that are opinions, hypothesis, and speculation from objective facts and results. *Yes*
- Provides well marked pedagogical references for less familiar readers to gain background necessary to replicate the paper. *Yes*

Does this paper make theoretical contributions? *Yes*

- All assumptions and restrictions are stated clearly and formally. *Yes*
- All novel claims are stated formally (e.g., in theorem statements). *Yes*
- Proofs of all novel claims are included. *Yes*
- Proof sketches or intuitions are given for complex and/or novel results. *Yes*
- Appropriate citations to theoretical tools used are given. *Yes*
- All theoretical claims are demonstrated empirically to hold. *Yes*
- All experimental code used to eliminate or disprove claims is included. *NA*

Does this paper rely on one or more datasets? *No* Does this paper include computational experiments? *Yes*

- Any code required for pre-processing data is included in the appendix. *Yes*
- All source code required for conducting and analyzing the experiments is included in a code appendix. *Yes*
- All source code required for conducting and analyzing the experiments will be made publicly available upon publication of the paper with a license that allows free usage for research purposes. *Yes*
- All source code implementing new methods have comments detailing the implementation, with references to the paper where each step comes from. *Yes*
- If an algorithm depends on randomness, then the method used for setting seeds is described in a way sufficient to allow replication of results. *Yes*
- This paper specifies the computing infrastructure used for running experiments (hardware and software), including GPU/CPU models; amount of memory; operating system; names and versions of relevant software libraries and frameworks. *Yes*
- This paper formally describes evaluation metrics used and explains the motivation for choosing these metrics. (yes/partial/no)
- This paper states the number of algorithm runs used to compute each reported result. *Yes*
- Analysis of experiments goes beyond single-dimensional summaries of performance (e.g., average; median) to include measures of variation, confidence, or other distributional information. *Yes*
- The significance of any improvement or decrease in performance is judged using appropriate statistical tests (e.g., Wilcoxon signed-rank). *NA*
- This paper lists all final (hyper-)parameters used for each model/algorithm in the paper's experiments. *Yes*
- This paper states the number and range of values tried per (hyper-) parameter during development of the paper, along with the criterion used for selecting the final parameter setting. *NA*

Technical Appendix

This appendix contains a literature survey on formal verification methodsa simple example of the problem at hand, the proofs of the lemmas presented in the paper, and implementation details for the experimental section.

An Illustrative Example

Consider the following stochastic differential equation:

$$\mathrm{d}X_t = aX_t\,\mathrm{d}t + \pi(X_t)\,\mathrm{d}W_t. \tag{19}$$

When $u \equiv 0$ there is no stochasticity to the problem, and the resulting deterministic system is unstable for any starting point x_0, since the analytical solution $x(t) = x_0 e^{at}$ to $\dot{x} = ax$ with the initial condition $x(0) = x_0$ tends to infinity.

Now consider the following control policy:

$$\pi_\sigma(x) \equiv \sigma x. \tag{20}$$

In other words, the controller adds a white noise proportional to the state. The resulting system's dynamics are

$$\mathrm{d}X_t = aX_t\,\mathrm{d}t + \sigma X_t\,\mathrm{d}W_t$$

which is a geometric Brownian motion with drift a and volatility σ. It is a well-studied stochastic process and the solution for the initial condition $X_0 = x_0$ is

$$X_t = x_0 e^{\left(a - \frac{\sigma^2}{2}\right)t + \sigma W_t}.$$

If $\sigma > \sqrt{2a}$, surprisingly (and even somewhat counterintuitively), the system becomes stabilized by pure noise injection. This highlights the fundamentally different nature of continuous-time stochastic problems compared to their deterministic counterparts. This phenomenon is illustrated by Figure 5.

Now consider three control policies with $\sigma \in \{1.0, 0.92, 0.84\}$. Let us restric our attention to the set $\mathbb{X} = [-1, 10]$. Suppose we are given the following time-heterogenous unsafe states and time-homogenous targets:

$$\mathfrak{X}_\oslash = \left([2 + 8e^{-0.075t}, 10]\right)_{t\geq 0} \quad \text{and} \quad \mathfrak{X}_\star = \left([-1, 1]\right)_{t\geq 0}.$$

Figure 5 illustrates this problem and shows sample paths of the three systems. The sample path under the policy π_1 satisfies the reach-avoid-stay property at time 10; the sample path under the policy $\pi_{0.84}$ does not satisfy the stay property, but it does satisfy the reach-avoidability. More interestingly, the policy $\pi_{0.92}$ is stable (converges to the equilibrium) so it satisfies the stay property, but not the reach-avoid one.

This example inspects a single sample path for each process. In practice, for each of the processes there exist (a.s.)

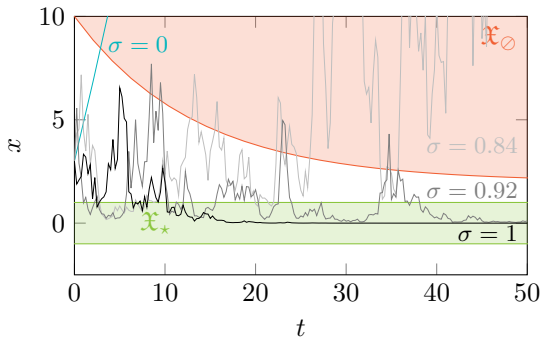

Figure 5: Sample paths of three geometric Brownian motions solving (19) with $a = 0.4$ and policies π_σ given by (20) when σ takes values of 1, 0.92, and 0.84, and 0.

paths failing to satisfy the reach-avoid-stay criteria, especially when they start close to the unsafe set. This is why we aim to verify RAS satisfaction probabilistically and restrict ourselves to some initial set.

The generator of a system driven by (19)–(20) is

$$\mathsf{G} = \frac{\partial}{\partial t} + ax\frac{\partial}{\partial x} + \frac{1}{2}(\sigma x)^2\frac{\partial^2}{\partial x^2}.$$

Analytical Derivatives of a Neural Network

We compute $\frac{\partial V}{\partial x}$ and $\frac{\partial^2 V}{\partial x^2}$ using the following statement.

Proposition 1 (Singla and Feizi (2020), Lemma 1)**.** *Consider an N-layer neural network defined recursively for $i = 0, 1, \ldots, N-1$ as*

$$\mathbf{a}^{(-1)} = \mathbf{x}, \quad \mathbf{z}^{(i)} = \mathbf{W}^{(i)}\mathbf{a}^{(i-1)} + \mathbf{b}^{(i)}, \quad \mathbf{a}^{(i)} = \sigma_i(\mathbf{z}^{(i)}).$$

The j-th row of the Hessian of its output $\mathbf{z}^{(N-1)}$ with respect to the input $\mathbf{x}$ can be computed via

$$\mathsf{H}_{\mathbf{x}}\mathbf{z}_j^{(N-1)} = \sum_{i=0}^{N-2}(\mathbf{B}^{(i)})^\top \operatorname{diag}\big(\mathbf{F}_j^{(N-1,i)} \odot \sigma_i''(\mathbf{z}^{(i)})\big)\mathbf{B}^{(i)};$$

$$\mathbf{B}^{(i)} = \begin{cases} \mathbf{W}^{(0)}, & i = 0, \\ \mathbf{W}^{(i)} \operatorname{diag}\big(\sigma_{i-1}'(\mathbf{z}^{(i-1)})\big)\mathbf{B}^{(i-1)}, & \imath \geq 1; \end{cases}$$

$$\mathbf{F}^{(k,i)} = \begin{cases} \mathbf{W}^{(k)}, & i = k-1, \\ \mathbf{W}^{(k)} \operatorname{diag}\big(\sigma_{k-1}'(\mathbf{z}^{(k-1)})\big) \\ \quad \cdot \mathbf{F}^{(k-1,i)}, & i \leq k-2, \end{cases}$$

and its Jacobian is equal to $\mathbf{B}^{(N-1)}$.

Remark 4. *Our neural network architecture is slightly different, with the final output $\mathbf{a}^{(N)} = \sigma_N(\mathbf{z}^{(N)})$ instead of $\mathbf{z}^{(N)}$. We obtain the formulae for our case by extending the network of Proposition 1 with a final linear layer with $\mathbf{W}^{(N+1)} = [1]$ an $\mathbf{b}^{(N+1)} = 0$ in the calculations.*

Given this proposition, we find the first derivative vector as the transpose of the Jacobian, and the second derivative vector as the diagonal of the Hessian.

Proof of Theorem 1

First, we need to ensure that ψ is indeed a stopping time.

Lemma 2. *Let $\mathfrak{N} = (\mathcal{N}_t)_{t\geq 0}$ be the natural filtration with respect to a state process $(X_{\pi,t}^{0,x_0})_{t\geq 0}$ issuing in some state $x_0 \in \mathbb{X}_0$. Consider a continuous function $V(t,x) : \mathbb{R}_+ \times \mathbb{X} \to \mathbb{R}$. For any constants $\alpha_S, \beta_{RA} \in \mathbb{R}$ such that $\beta \leq \rho$, the random variable ψ given by (9) is an $\mathfrak{N}$-stopping time. Moreover, if the decrease and stay conditions of Definition 2 hold, then $\psi < \infty$ (a.s.).*

In proving Lemma 2, we will employ the following result.

Proposition 2 (recurency citerion, cf. Khasminskii (2011), Theorem 3.9)**.** *A Feller–Dynkin process $(\eta_t)_{t\geq 0}$ with infinitesimal generator G leaves a domain $\mathbb{U}$ in finite time (a.s.) if it is regular (i.e., defined a.s. for all $t \geq 0$) and there exists in $\mathbb{R}_+ \times \mathbb{U}$ a non-negative function $V(t,x)$, twice continuously differentiable with respect to x and continuously differentiable with respect to t, such that $\mathsf{G}V(t,x) \leq -\kappa(t)$ for some non-negative function $\kappa(t)$ satisfying*

$$\lim_{t\to\infty}\int_0^t \kappa(s)\,\mathrm{d}s = \infty.$$

Remark 5. *Note that Proposition 2 requires the process to be regular which under our assumptions follows from continuity (Khasminskii 2011, p. 75).*

Proof of Lemma 2. Note that because V is continuous and therefore preserves Borel-measurability, both of the events $\{V(t, X_{\pi,t}^{0,x_0}) < \alpha_{\mathrm{S}}\}$ and $\{V(t, X_{\pi,t}^{0,x_0}) \geq \beta_{\mathrm{RA}}\}$ are Borel-measurable; therefore, their union is also measurable. Thus, ψ is the first hit time of a measurable set, and by the début theorem it is a stopping time. Consider next a domain $\mathbb{U} = \{x \mid \alpha_{\mathrm{S}} \leq V(t,x) < \beta_{\mathrm{RA}}\}$. The process $(X_{\pi,t}^{0,x_0})_{t\geq 0}$ has a generator G_π which satisfies on $\mathbb{U}$ the condition $\mathsf{G}_\pi V(t,x) \leq -\kappa(t)$ for $\kappa(t) = \zeta(t) \vee \xi(t)$ due to the decrease and stay conditions. Since

$$\int_0^t \kappa(s)\,\mathrm{d}s = \int_0^t \zeta(s) \vee \xi(s)\,\mathrm{d}s \geq \int_0^t \zeta(s)\,\mathrm{d}s,$$

the conditions of Proposition 2 are satisfied and the stopping time ψ is finite (a.s.). □

Next, we use the stopping time ψ to construct the following supermartingale.

Non-negative supermartingales such as ψ-RAS-SM have the following properties useful to us.

Proposition 3 (optional stopping theorem, cf. Le Gall (2016), Theorem 3.25)**.** *Let $(\eta_t)_{t\geq 0}$ be a non-negative supermartingale with right-continuous sample paths with respect to some filtration $(\mathcal{F})_{t\geq 0}$. Let τ_1 and τ_2 be two stopping times such that $\tau_1 \leq \tau_2$. Then, η_{τ_1} and η_{τ_2} are in L_1 and*

$$\eta_{\tau_1} \geq \mathbb{E}[\eta_{\tau_2} \mid \mathcal{F}_{\tau_1}].$$

Corollary 2. *For such a process $(\eta)_{t\geq 0}$, $\mathbb{E}[\eta_{\tau_1}] \geq \mathbb{E}[\eta_{\tau_2}]$.*

Proof. Follows immediately by applying the law of total expectation to the statement of Proposition 3. □

Proposition 4 (Chebyshev's inequality, cf. Stein and Shakarchi (2009), p. 91)**.** *If f is a non-negative function, $(\eta_t)_{t\geq 0}$ is a stochastic process such that $\mathbb{E}\big[f(\eta_t)\big]$ exists, and $r > 0$, then*

$$\mathbb{P}\big\{f(\eta_t) \geq r\big\} \leq \frac{\mathbb{E}\big[f(\eta_t)\big]}{r}.$$

Proposition 5 (Maximal inequality for non-negative supermartingale, cf. Prajna, Jadbabaie, and Pappas (2004), Lemma 6)**.** *Given a filtration $\mathfrak{F} = (\mathcal{F})_{t\geq 0}$, let $(\eta_t)_{t\geq 0}$ be a non-negative $\mathfrak{F}$-supermartingale with right-continuous sample paths. Then for every $r > 0$,*

$$\mathbb{P}\Big\{\sup_{t\geq 0} \eta_t \geq r\Big\} \leq \frac{\mathbb{E}[\eta_0]}{r}.$$

Finally, armed with all the necessary properties of an RAS-SM, we are ready to prove Theorem 1.

Proof of Theorem 1. Consider the supermartingale $(Y_t)_{t\geq 0}$ of Lemma 1. By its construction and the non-negativity condition for RASMs, it is non-negative, and by the assumptions of Theorem 1 it is continuous, since the RAS-C V is twice continuously differentiable with respect to x and continuously differentiable with respect to t, both implying continuity. Therefore, $\mathbb{E}[Y_\tau] \leq \mathbb{E}[Y_0]$ for any stopping time τ by Corollary 2. Note that

$$Y_\tau = V(\tau \wedge \tau, X^{0,x_0}_{\pi,\tau\wedge\tau}) = V(\tau, X^{0,x_0}_{\pi,\tau}), \quad \text{and} \tag{21}$$

$$Y_0 = V(0, X^{0,x_0}_{\pi,0}) = V(0, x_0). \tag{22}$$

The initial condition of Definition 2 implies

$$\mathbb{E}\big[V(0, X^{0,x_0}_{\pi,0})\big] = V(0, x_0) \leq \alpha_{\text{RA}}. \tag{23}$$

Equations (21)–(23) imply

$$\mathbb{E}\big[V(\tau, X^{0,x_0}_{\pi,\tau})\big] \leq \mathbb{E}\big[V(0, X^{0,x_0}_{\pi,0})\big] = V(0, x_0) \leq \alpha_{\text{RA}}.$$

By Proposition 4, this implies

$$\mathbb{P}\big\{V(\tau, X^{0,x_0}_{\pi,\tau}) \geq \beta_{\text{RA}}\big\} \leq \tfrac{1}{\beta_{\text{RA}}}\,\mathbb{E}\big[V(\tau, X^{0,x_0}_{\pi,\tau})\big] \leq \tfrac{\alpha_{\text{RA}}}{\beta_{\text{RA}}}.$$

Now consider the stopping time ψ of Lemma 2. By its construction, either

$$V(\psi, X^{0,x_0}_{\pi,\psi}) \leq \alpha_{\text{S}} \quad \text{or} \quad V(\psi, X^{0,x_0}_{\pi,\psi}) > \beta_{\text{RA}}.$$

Since $\alpha_{\text{S}} \leq \beta_{\text{RA}}$, these events are incompatible, and thus

$$\begin{aligned}\mathbb{P}\big\{V(\psi, X^{0,x_0}_{\pi,\psi}) \leq \alpha_{\text{S}}\big\} &\\ &= 1 - \mathbb{P}\big\{V(\psi, X^{0,x_0}_{\pi,\psi}) \geq \beta_{\text{RA}}\big\}\\ &\geq 1 - \tfrac{\alpha_{\text{RA}}}{\beta_{\text{RA}}} = \varepsilon.\end{aligned}$$

Since $V(t, X^{0,x_0}_{\pi,t}) \leq \beta_{\text{RA}}$ for all $t < \psi$ by the definition of the stopping time ψ, it follows from the safety condition that $X^{0,x_0}_{\pi,t} \neq \mathbb{X}_\oslash$ for all $t < \psi$. Thus,

$$\mathbb{P}\Big\{\big(X^{0,x_0}_{\pi,\psi} \in L^-_{\alpha_{\text{S}}}(V)\big) \wedge \big(\forall t < \psi : X^{0,x_0}_{\pi,t} \notin \mathbb{X}_\oslash\big)\Big\} \geq \varepsilon.$$

Because $L^-_{\alpha_{\text{S}}}(V) \subset L^-_{\beta_{\text{S}}}(V) \subset \mathbb{X}_\star$ by construction and the goal condition, this event is a subset of the reach-avoid event E_{RA} of (3); therefore, the reach-avoid property is satisfied with probability at least ε.

Next, we prove that the stay property is satisfied as well.

Using Proposition 5 and shifting the time index by ψ (which is possible because the process is Markovian),

$$\mathbb{P}\Big\{\sup_{t\geq\psi} V(t, X^{0,x_0}_{\pi,t}) \geq \beta_{\text{S}}\Big\} \leq \tfrac{1}{\beta_{\text{S}}}\,\mathbb{E}\big[V(0, X^{0,x_0}_{\pi,\psi})\big] \leq \tfrac{\alpha_{\text{S}}}{\beta_{\text{S}}}.$$

The opposite event can be written as

$$\begin{aligned}\Big\{\sup_{t\geq\psi} V(t, X^{0,x_0}_{\pi,t}) < \beta_{\text{S}}\Big\} &=\\ &\Big\{\forall t \geq \psi : V(t, X^{0,x_0}_{\pi,t}) < \beta_{\text{S}}\Big\}\\ &\cap \Big\{\nexists \gamma < \beta_{\text{S}} : \forall t \geq \psi : V(t, X^{0,x_0}_{\pi,t}) < \gamma\Big\}.\end{aligned}$$

Because $\mathbb{P}[A] \geq \mathbb{P}[A \cap B]$ for any A and B, this implies

$$\mathbb{P}\Big\{\forall t \geq \psi : V(t, X^{0,x_0}_{\pi,t}) < \beta_{\text{S}}\Big\} \geq 1 - \tfrac{\alpha_{\text{S}}}{\beta_{\text{S}}} = \delta.$$

Again, $L^-_{\alpha_{\text{S}}}(V) \subseteq \mathbb{X}_\star$ due to the goal condition. Thus, this event is a subset of the stay event E_{S} of (4), which means that the stay part of the specification is satisfied with probability at least δ. □

Proof of Lemma 1

Lemma 1 is proven using the following proposition.

Proposition 6 (first exit process is a supermartingale, cf. Khasminskii (2011), Lemma 5.1)**.** *Let $V(t, x)$ be a function twice continuously differentiable with respect to x, continuously differentiable with respect to t on the set $\mathbb{R}_+ \times \mathbb{U}$ for a bounded domain $\mathbb{U}$. Moreover, in this set $\mathsf{G}_\pi V \leq 0$. Let $\tau_{\mathbb{U}}$ be the first exit time from $\mathbb{U}$. Then the process $V(t \wedge \tau_{\mathbb{U}}, X^{0,x_0}_{\pi,t\wedge\tau_{\mathbb{U}}})$ is a supermartingale.*

Proof of Lemma 1. Follows immediately from Proposition 6 and boundedness of $\mathbb{X}$ and therefore any of its subsets. □

Proof Sketch for Corollary 1

Since both proofs are very similar to the proof of Theorem 1, we do not present their full versions, but restrict ourselves to a proof sketch.

Sketch of a proof for Corollary 1. The case of staying is already part of the Theorem 1 proof. The proof for reach-avoidance without staying follows the steps of Theorem 1 proof for a stopping time

$$\psi \triangleq \inf_{t\geq 0}\big\{t \;\big|\; V(t, X^{0,x_0}_{\pi,t}) \notin L^-_{\beta_{\text{RA}}}(V) \setminus \mathsf{int}\,\mathfrak{X}_\star\big\}.$$

It is also similar to the proof of discrete-time reach-avoidance done by Žikelić et al. (2023). □

Computing Infrastructure

We conducted the experiments on MacBook Pro (Model: 14" 2021, CPU: Apple M1 Max, RAM: 32 GB, OS: macOS Sonoma 14.5). The experiments were run using Python 3.11.7. The names and versions of the libraries we used are included in the code appendix in `requirements.txt`.

Hyperparameter Values

For the inverted pendulum, we use a policy pre-trained with `torchRL` for the deterministic version of the problem. The script used for training and the saved policy are both included in the code appendix. The policy network consists of two linear layers with hyperbolic tangent activations, followed by a final linear layer. The hidden layers consist of 64 neurons.

For the certificate network, the architecture is the same, but with the addition of a softplus activation at the end to make the values nonnegative. The hidden layers contain 32 neurons.

The values of the remaining hyperparameters are summarized in Table 2 and can be found in the code appendix.

Table 2: Hyperparameter values.

Parameter		GBM	Pendulum
verification frequency	q	1000	1000
batch size	n	256	256
verifier mesh size	m	200	400
generator threshold	ζ	1.0	1.0
regularizer multiplier	λ	10^{-1}	10^{-1}
verification slack	κ	4	4
maximum verification depth	k	2	5
optimizer		`Adam`	`Adam`
optimizer learning rate		10^{-3}	10^{-3}