RVSpec: Cyber-Physical Interplay Graphs for Formal Specification of Robotic Vehicle Control Software

Chaoqi Zhang*, Minhyun Cho†, Inseok Hwang†, Hyungsub Kim*

*Indiana University Bloomington

†Purdue University

I. The Problem

Previous works on automatically generating formal specifications overlook *cyber-physical interplay*, which is often absent from the documentation but is essential for capturing intended behaviors, e.g., altitude changes caused by air pressure and servo lag.

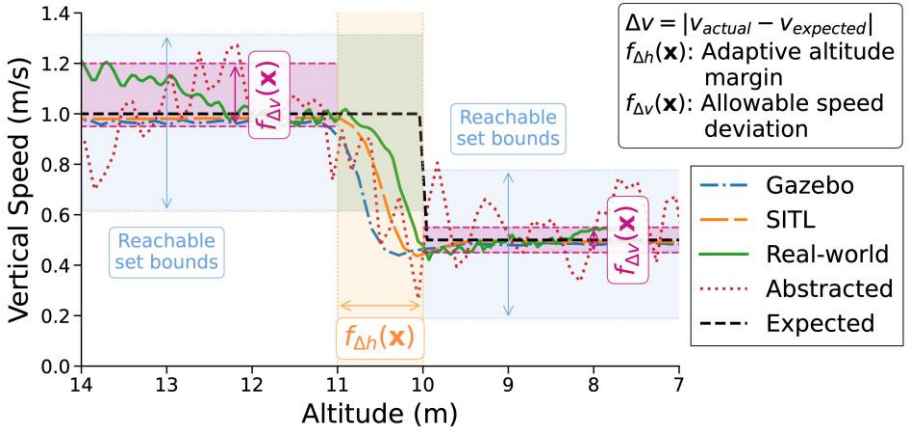

Fig. 1. The actual vertical speed deviates from the documented "expected" vertical speed by margin $f_{\Delta v}(x)$. Strict equality specifications consider these normal physical dynamics as bugs.

II. Our Approach

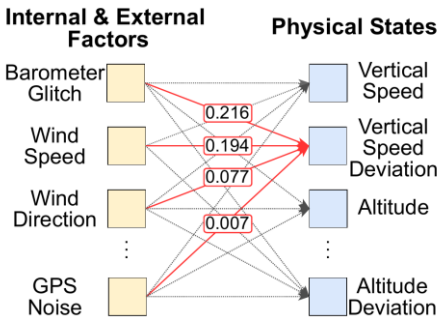

Fig. 2. CPG: red edges = influence weights.

Cyber-Physical Interplay Graph (CPG)

A directed, weighted graph that quantifies how much *internal* and *external factors* influence an RV's physical states.

HOW THE CPG WORKS

- **Identifies** internal & external factors (e.g., wind, GPS noise) from docs + source code
- **Quantifies** their influence on physical states via uncertainty-aware modeling
- **Derives** tolerance functions $f_{\Delta v}(x)$ that relax strict equalities in MTL formulas.

BASELINE (PGFuzz-previous work, strict) ✗ 288 false alarms / hour

$$\square\{(\text{Mode} = \text{LAND}) \wedge (\text{ALT} \geq 10) \rightarrow (\text{Speed\_vert} = \text{WPNAV\_SPEED\_DN})\}$$

Demands exact vertical speed match — flags normal landing physics as bugs.

RVSPEC (CPG-aware, tolerance) ✓ 2 false alarms / hour

$$\square\{(\text{Mode} = \text{LAND}) \wedge (\text{ALT} \geq 10) \rightarrow (|\text{Speed\_vert} - \text{WPNAV\_SPEED\_DN}| \leq f_{\Delta v}(x))\}$$

Allows a CPG-derived margin — tolerates real-world variance, still catches real bugs.

III. Workflow

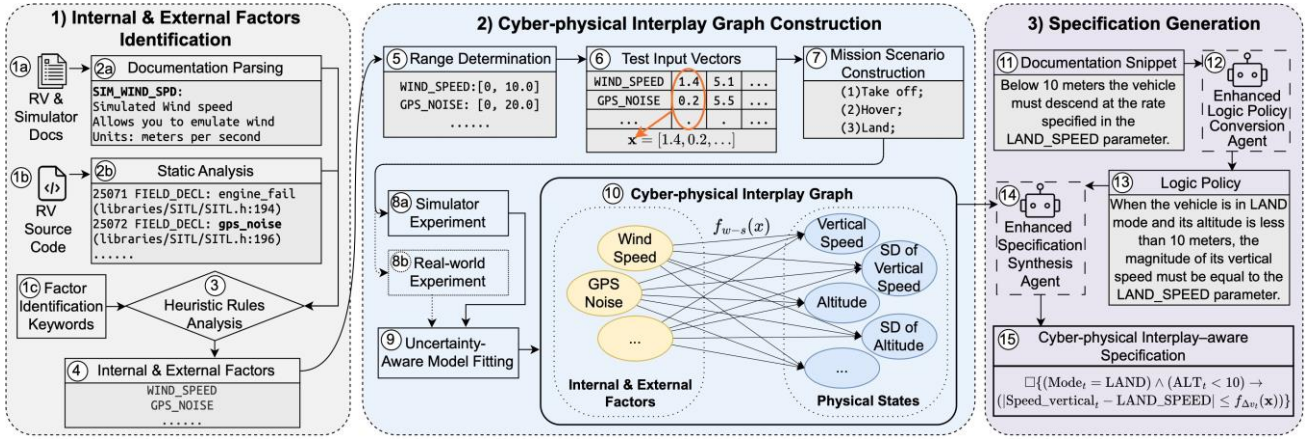

Fig. 3. RVSpec workflow: (1) Factor identification → (2) CPG construction (uncertainty-aware modeling) → (3) Specification generation (two LLM agents).

IV. Results

Software	Syntactic Validity	Semantic Accuracy	Cyber-physical Consistency
ArduPilot	200/200 (100%)	182/200 (91.0%)	164/200 (82.0%)
PX4	200/200 (100%)	181/200 (90.5%)	163/200 (81.5%)
openpilot	44/44 (100%)	39/44 (88.6%)	33/44 (75.0%)
cFS	44/44 (100%)	39/44 (88.6%)	34/44 (77.3%)
Overall	**488/488 (100.0%)**	**441/488 (90.4%)**	**394/488 (80.7%)**

Tab.1. Spec. correctness across 4 RV platforms

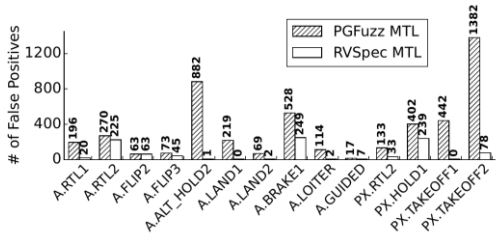

Fig.4. False positives: PGFuzz vs RVSpec MTL

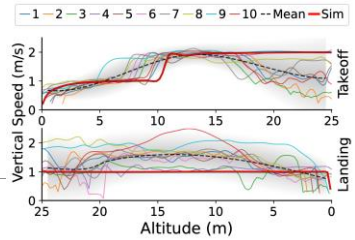

Fig.5. Vertical speed in takeoff and landing for 10 real-world tests and simulation

Spec. Accuracy **80.7%** *vs. 51.6% LLM-only baseline*

FP Reduction **79.9%** *4,790 → 964 false alarms*

KEY INSIGHT

Accurate specifications require understanding the interplay between *cyberspace* and *physical space*. To address this, we propose *cyber-physical interplay graphs (CPGs)* that quantify how much *internal* and *external factors* (e.g., scheduling delay, wind) affect a vehicle's physical states (e.g., roll, pitch, yaw).

Code github.com/KimHyungSub/RVSpec