RACE-Nav: Refinable Occupancy Contracts and Prefix-Wide Runtime Assurance for Dynamic Mobile-Robot Navigation

Hari Sunmukeswar Baskaran[1] and **Salem Ameen**[1]

School of Science, Engineering and Environment, University of Salford, Salford, United Kingdom

[1] H.S.Baskaran@edu.salford.ac.uk; S.A.Ameen1@salford.ac.uk

Abstract*—*Modular navigation stacks permit predictors and controllers to be exchanged, but their uncertainty, timing, and braking assumptions rarely form one executable assurance argument. RACE-Nav introduces typed, refinable occupancy contracts whose guaranteed class, hazard scope, frame, horizon, risk allowance, and expiry are checked before command release. A model-conditional radial provider supplies the closed-loop conditional premise used by the risk ledger; marginal split conformal prediction is retained only as a diagnostic. Concrete trajectories are certified by finite sampled tests with Hausdorff-Lipschitz inter-sample margins and a prefix-wide recoverable-stop condition. A horizon-aware lease and downstream watchdog preserve the trusted-gate invariant under cancellation and provider substitution. We prove provider refinement, continuous-time separation, swept-stop safety, finite/anytime risk composition, stale-command exclusion, non-Zeno switching, and conditional release liveness. In a held-out paired-seed benchmark spanning eight dynamic interactions and six controllers (1,920 episodes), RACE-Nav produced 0/320 collisions (95% Wilson interval 0.00%-1.19%) and 97.5% goal success versus 12.2% collisions for the predictive time-to-collision/model predictive control comparator. The decision kernel recorded 0 and modelled 300-ms deadline misses. Stress and ablation results separate guaranteed behaviour from violations of the declared model and actuator bounds.

Index Terms*—*Mobile robot navigation, occupancy contracts, predictive safety, risk composition, ROS 2, runtime assurance.

Figure 1 summarizes the end-to-end data, certification, command-ownership, and audit path.

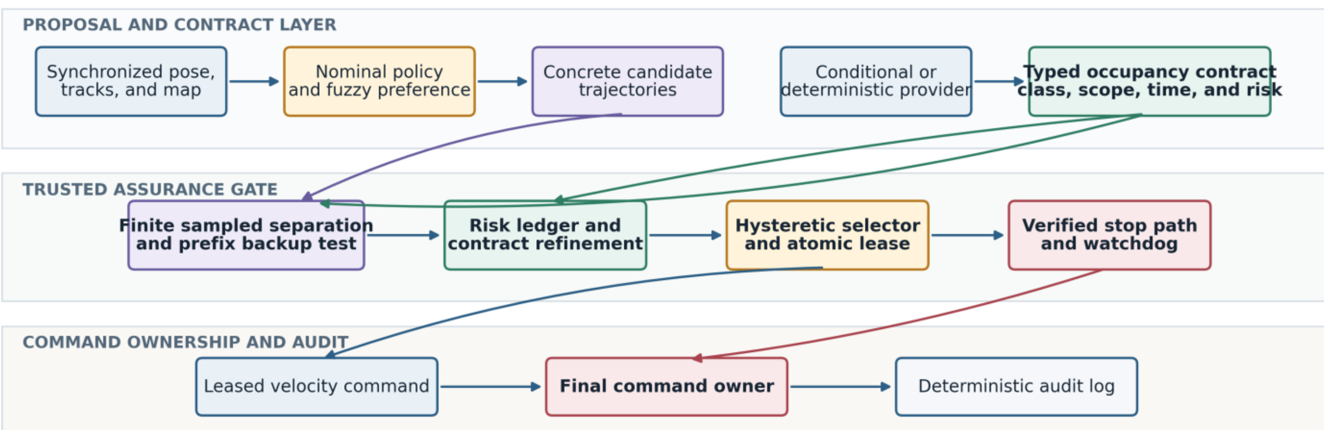

Fig. 1. RACE-Nav data and command path. Nominal preferences and predictor outputs remain outside the trusted gate. Only a concrete certified trajectory receives a finite lease; invalidity, expiry, or a deadline miss invokes the verified stop path downstream of command smoothing.

I. INTRODUCTION

Robot Operating System 2 (ROS 2) and Nav2 separate perception, localization, planning, control, behavior management, and velocity processing into replaceable components [1], [2]. That modularity improves engineering reuse but does not make their safety assumptions compositional. A predictor may offer only marginal coverage, a controller may test collision at discrete samples, and an action cancellation may remain pending while a smoothed velocity continues downstream. The missing object is an executable interface that preserves statistical semantics, continuous-time geometry, stopping recoverability, and command lifetime together.

RACE-Nav (Risk-Aware Contract Enforcement for Navigation) places that interface between arbitrary nominal navigation modules and the final actuator command. Predictors expose typed occupancy contracts; the assurance gate accepts only guarantee classes compatible with its mission theorem. Candidate labels such as Continue, Slow, and Replan have no assumed safety order. Every concrete rollout is checked independently, while the nominal or fuzzy output contributes only a preference cost after certification.

The statistical statement is deliberately narrower than total collision risk. The ledger budgets occupancy-contract miscoverage for represented hazards under declared estimation, acceleration, timing, tracking, and braking bounds. Missed agents, invalid association, stale transforms, unsupported guarantee classes, or an exhausted budget invalidate the motion contract and invoke the verified stop. This separation prevents a finite-sample prediction statement from being reported as an unconditional system-safety probability.

Contributions: 1) a typed contract algebra with guarantee classes, hazard scope, refinement, provider substitution, and multi-provider composition; 2) an instantiated model-conditional provider that supplies the conditional premise used by adaptive risk spending; 3) finite, implementable continuous-time separation and prefix-wide swept-stop tests with horizon-aware command leases; 4) a trusted-gate proof deriving contact-implies-miscoverage rather than assuming it, together with non-Zeno switching and conditional liveness; and 5) an executable reference kernel evaluated against five comparators, six ablations, nine fault/shift conditions, and measured tail latency.

II. RELATED WORK AND DISTINCTION

A. Modular Navigation and Runtime Assurance

Nav2 organizes global and local planning, controller plugins, recoveries, and action interfaces [2]. The dynamic-window family samples short-horizon velocity trajectories [3], while runtime assurance separates an advanced controller from a trusted monitor, switch, and backup [4]. SOTER demonstrates this separation in ROS [5]. RACE-Nav adds a predictor-facing contract and risk ledger, and it binds every release to the exact contract identifier and stopping horizon used by the certificate.

B. Predictive and Conformal Safety

Predictive safety filters establish recursively feasible constrained backups [6]. Conformal regions have been integrated with safety filtering and planning in dynamic environments [7], [8], including adaptive motion planning [9], mission-joint shrinking-horizon model-predictive control (MPC) [10], egocentric navigation [11], crowd navigation [12], time-aware planning [13], and conformal policy learning under shift [14]. These methods define important providers or comparators; RACE-Nav's distinct object is the refinement-compatible boundary between a provider and a command-owning assurance gate.

Exact nontrivial distribution-free conditional coverage is impossible in general over rich conditioning information [15]. Online conformal methods can control long-run behavior under specified shift criteria [16], but neither ordinary split conformal nor adaptive average coverage implies the per-update conditional bound needed for arbitrary predictable risk spending. We therefore instantiate an explicit conditional observation model for the reported benchmark and classify split conformal output as marginal only. Mission-joint conformal calibration [10] is an alternative compatible class for fixed complete missions.

C. Geometry, Braking, and Hybrid Execution

Chance constraints [17], scenario optimization [18], control barrier functions [19], and inevitable-collision states [20] provide complementary tools. RACE-Nav combines occupancy tubes with a finite sampled Hausdorff margin and a recoverable-stop invariant at every executable-prefix point. Its command lease covers prediction expiry, the supplied horizon, and the full stopping guard. This closes the gap between a mathematical path certificate and asynchronous command execution.

III. TYPED OCCUPANCY CONTRACTS

A. Dynamics, History, and Scope

Let k index every provider attempt, including rejected attempts, and let wall-clock time remain separate. A planar acceleration-limited rollout has

$$\mathbf{x} = \left(p_x, p_y, \theta, v, \omega\right)^{\top}, \mathbf{u} = (a, \gamma)^{\top} \tag{1}$$

$$\dot{\mathbf{p}} = v(\cos\theta, \sin\theta)^{\top}, \dot{\theta} = \omega, \dot{v} = a, \dot{\omega} = \gamma \tag{2}$$

The body-frame footprint B is compact, and r_B is its circumscribed radius. The benchmark and main proof use the disc $B_2(0,r_B)$, so rotation does not enlarge the occupied set; Appendix A gives the noncircular swept-rotation term. The true represented hazard union is O(t). A new birth, unmatched detection, identity alarm, or a hazard outside the declared scope invalidates the contract instead of inheriting coverage.

The filtration F_k is the immutable proposal/certification snapshot: it contains observations, calibration information, robot state, prior decisions, and ledger entries available before the covered future occupancy is revealed. The contract, allowance, proposed release, and activation predicate must be F_k-measurable. Theorem eligibility therefore requires the later atomic commit outcome to be determined by that snapshot under the declared deterministic plant and timing bounds. Any intervening obstacle observation, uncertain callback survival, anchor mismatch, or deadline mismatch invalidates the proposal and requires a stop or fresh certification; it cannot be used to select a favorable contract. In the nominal benchmark, quantized release and the exact anchor are snapshot-predictable; overload and plant-mismatch stresses are outside this premise.

B. Contract Fields and Guarantee Classes

A provider returns

$$\begin{gathered} \mathcal{C}_k = (\iota_k, I_k, f_k, \mathbf{U}_k, \mathcal{H}_k, \mathcal{A}_k, \\ \Gamma_k, \alpha_k, \beta_k, t_{e,k}, \nu_k, \chi_k). \end{gathered} \tag{3}$$

Here ι_k is immutable; I_k is the support interval of length $H_k \Delta_k$; f_k is the frame; U_k is the sampled tube; H_k is the represented-hazard scope; $\mathcal{A}_k$ is the declared assumption predicate; Γ_k is the guarantee class; α_k and β_k are miscoverage and outer-confidence allowances; the next field is expiry; ν_k identifies the model/calibration version; and χ_k records observable health. Frame conversion, time normalization, complete h=0 support, version, freshness, association, horizon, and cancellation capability are acceptance predicates, not statistical evidence.

$$\begin{aligned} t_{\sup,k} &= \min\{t_{e,k}, \sup I_k\}, \\ I_k^{+} &= I_k \cap \left[t_{r,k}, t_{\sup,k}\right], \\ E_k &= \{\mathcal{O}(t) \subseteq \overline{\mathcal{U}}_k(t), \forall t \in I_k^{+}\}, \\ \mathbb{P}(E_k^c \mid \mathcal{F}_k) &\leq \alpha_k \text{a.s.} \end{aligned} \tag{4}$$

$$\begin{aligned} A_k &\in \{0,1\}, \sigma(A_k, \alpha_k, \beta_k) \subseteq \mathcal{F}_k, \\ \textstyle\sum_{k\geq 0} A_k\,(\alpha_k + \beta_k) &\leq \delta \text{a.s.} \end{aligned} \tag{5}$$

Here A_k=1 exactly when the snapshot-predictable proposal completes the atomic activation transaction; rejected or canceled attempts have A_k=0 and spend no allowance. The effective interval I_k^+ begins at the proposed release, so already elapsed contract support is never charged as future coverage.

Table I defines the consumer-visible classes. A marginal provider is eligible only for a fixed evaluation whose joint or per-epoch guarantee is independently justified; it cannot enter the adaptive ledger. A contract with a probably approximately correct conditional guarantee (PAC-conditional contract) enters the adaptive ledger only when its outer-validity event has a predictable bound that remains valid after acceptance; its β_k is then charged with α_k. Otherwise it is restricted to a fixed study.

TABLE I
GUARANTEE CLASSES EXPOSED AT THE CONTRACT BOUNDARY

Class	Provider obligation	Adaptive ledger
Det.	E_k holds under declared reachable-set bounds; α=0	Yes
Cond.	Conditional failure probability is at most α_k	Yes
PAC-cond.	Conditional statement plus selection-valid outer failure β_k	Only with α+β budget
Mission	One joint event covers the complete fixed mission	Fixed mission
Marg.	Marginal failure probability is at most α_k	No

C. Instantiated Conditional Provider

A generic conditional provider reports an F_k-measurable tail envelope. Missing representation sets the scene-horizon score to +infinity:

$$\begin{aligned} \Phi_{k,i,h}(q) &= \{\mathcal{O}_{k,i,h} \subseteq \hat{\mathcal{O}}_{k,i,h} \oplus B_2(0,\gamma_h q)\}, \\ S_k &= \inf\{q \geq 0 : \Phi_{k,i,h}(q), \forall i,h\}. \end{aligned} \tag{6}$$

$$\begin{aligned} \overline{G}_k(q) &\geq \mathbb{P}(S_k > q \mid \mathcal{F}_k), \\ q_k &= \min\{q \in \mathcal{Q}_k : \overline{G}_k(q) \leq \alpha_k\}. \end{aligned} \tag{7}$$

If the feasible quantile set is empty, the provider abstains. Otherwise, the inflated union of predicted footprints satisfies the required conditional event by construction. In the benchmark, the scenario supplies a deterministic reference motion with acceleration bounded by $\bar{a}_O$. Each agent receives one persistent latent offset $z_i=(\delta p_x,\delta p_y,\delta v_x,\delta v_y)$ drawn from $N(0,P_0)$ before the episode; $H_p(t)=[I,tI]$ maps it into position. Position and velocity measurements equal the latent-perturbed physical state plus independent, declared Gaussian sensor noise. Linear-Gaussian conditioning therefore gives the exact history-conditional posterior $z_i|F_k$ distributed as $N(\mu_{i,k},P_{i,k})$. A union allocation over M_k agent-support pairs gives

$$\begin{aligned} \hat{p}_{i,h} &= \bar{p}_i(s_k+t_h) + H_p(s_k+t_h)\mu_{i,k}, \\ \Sigma_{i,h} &= H_p(s_k+t_h)P_{i,k}H_p(s_k+t_h)^\top, \\ \kappa_k &= \sqrt{2\log(M_k/\alpha_k)}, \\ r_{i,h} &= r_i + \bar{a}_O\Delta_k^2/8 + \kappa_k\sqrt{\lambda_{\max}(\Sigma_{i,h})}. \end{aligned} \tag{8}$$

where t_h=h Δ, $\bar{p}_i$ is the declared scripted mean, and r_i is the physical obstacle radius. The largest-eigenvalue factor bounds the Euclidean norm of the anisotropic Gaussian residual. The scripted mean is C^1 with $\bar{a}_O$-Lipschitz velocity (equivalently, a weak second derivative bounded almost everywhere by $\bar{a}_O$), so the acceleration term is the sharp chord-interpolation remainder on a support sampled every Δ_k. Each pair has conditional tail probability at most α_k/M_k, so the union bound yields (4) without independence across agents or times. Missing detections, an association alarm, stale data, or unsupported noise/acceleration parameters invalidate the contract. The distribution-shift stress intentionally violates the declared acceleration bound.

D. Refinement, Substitution, and Composition

After verified frame/time normalization, C_n refines C_o when it covers no fewer hazards for no shorter interval, uses no stronger assumptions, supplies a no-weaker guarantee class and no larger allowance, and its tube is contained in the old tube:

$$\begin{aligned} \mathcal{C}_n \sqsubseteq \mathcal{C}_o &\Leftrightarrow I_n^+ \supseteq I_o^+, \mathcal{H}_n \supseteq \mathcal{H}_o, \\ \mathcal{A}_o &\Rightarrow \mathcal{A}_n, \\ \mathcal{U}_n(t) &\subseteq \mathcal{U}_o(t)(t \in I_o^+), \\ \Gamma_n &\succcurlyeq \Gamma_o, (\alpha_n,\beta_n) \leq (\alpha_o,\beta_o). \end{aligned} \tag{9}$$

The guarantee order is semantic: under the same assumption predicate and scope, deterministic containment is stronger than a conditional bound, and a conditional bound is stronger than a merely marginal bound. A PAC-conditional contract with positive outer allowance and an ordinary conditional contract are incomparable under (9); refinement is possible only when the PAC outer allowance is zero and every other displayed relation holds. Mission-joint and per-update classes are otherwise incomparable unless their events and intervals coincide. Expiry, version, frame-transform, and identifier checks remain mandatory even when the displayed set inclusions hold; the allowance-pair inequality in (9) is componentwise.

Proposition 1 (provider substitution). *For an obstacle-monotone certifier, replacing C_o by C_n satisfying (9), binding or recomputing the certificate to the new immutable identifier, preserves the deterministic certificate and cannot increase the ledger charge.*

Proof. On the new coverage event, the true occupancy lies in the new tube and therefore in the old tube. The old monotone clearance proof remains valid, while the new conditional failure allowance is no greater. Different frames, grids, identifiers, or nonnested tubes require atomic re-certification; an old lease is never extended.

Compatible simultaneous providers compose without an independence assumption. For the same hazards, a union tube fails only if every provider fails, whereas an intersection tube can fail when any provider fails:

$$\begin{aligned} \zeta_r &= \alpha_r + \beta_r, \\ \mathcal{U}_h^{\cup} &= \bigcup_r \mathcal{U}_h^r, \zeta^{\cup} = \min_r \zeta_r, \\ \mathcal{U}_h^{\cap} &= \bigcap_r \mathcal{U}_h^r, \zeta^{\cap} = \min(1, \textstyle\sum_r \zeta_r). \end{aligned} \tag{10}$$

For disjoint hazard scopes, the aggregate tube is the spatial union, and its conditional allowance is the sum. These operators require common time/frame semantics and explicit hazard identity; otherwise, composition is rejected. Figure 2 visualizes refinement and the mandatory re-certification boundary. The executable benchmark instantiates the conditional provider in Section III-C and unions its per-agent sets; (9)-(10) specify the provider interface and are not reported as a measured generic hot-swap API.

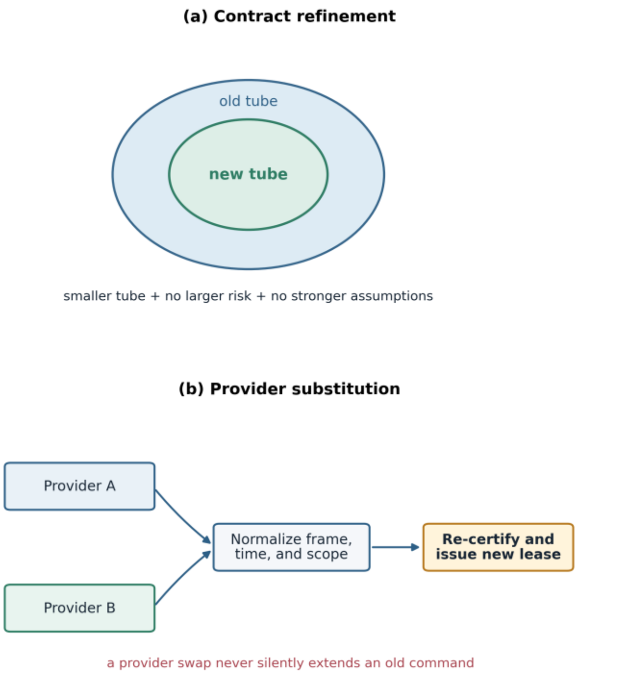

Fig. 2. Contract refinement and substitution. A valid refinement uses a contained tube with no larger allowances and no stronger assumptions. Every provider replacement is normalized, rebound to a new identifier, re-certified, and issued a new lease.

IV. PREFIX-WIDE RUNTIME CERTIFICATION

A. Finite Continuous-Time Separation Test

The predicted body placement combines candidate position, heading, and the compact body set. Under verified bounds on translational and angular speed, this footprint is Hausdorff-Lipschitz:

$$d_H\left(\hat{\mathcal{R}}(t),\hat{\mathcal{R}}(t')\right) \le L_R|t-t'|, L_R = \bar{v}_R + r_B\bar{\omega}_R \quad (11)$$

A contract sample cannot be treated as constant until the next sample. Between adjacent supports, the gate linearly interpolates each disc center and radius, then applies the verified frame transform and its declared set-valued error bound:

$$\begin{aligned} t &\in [t_{k,h}, t_{k,h+1}], \lambda = (t - t_{k,h})/\Delta_k, \\ \bar{c}_i(t) &= (1-\lambda)c_{i,h} + \lambda c_{i,h+1}, \\ \bar{r}_i(t) &= (1-\lambda)r_{i,h} + \lambda r_{i,h+1} + \varepsilon_{\mathrm{tf}}, \\ \overline{\mathcal{U}}_k(t) &= T_{f_k \to f_0} \bigcup_i B_2\left(\bar{c}_i(t), \bar{r}_i(t)\right). \end{aligned} \quad (12)$$

Here the transform contribution $\varepsilon_{\mathrm{tf}}$ is included in the normalized radius. Equation (13) gives a certified Hausdorff speed of the serialized piecewise-linear disc boundary; a time-varying transform contributes its own verified rate to L_U. Let a finite query set cover an executable prefix, let d_Q be its fill distance, let $\varepsilon_R(t)$ be a Lipschitz pose/tracking error envelope, and let the distance value at each query be a conservative numerical lower bound from the predicted footprint to the static-map union and normalized occupancy tube. The implemented finite test is

$$\begin{aligned} L_U &= \max_{i,h} \frac{\|c_{i,h+1} - c_{i,h}\| + |r_{i,h+1} - r_{i,h}|}{\Delta_k}, \\ \underline{d}_\ell - \varepsilon_R(q_\ell) &\ge (L_R + L_U + L_\varepsilon)d_Q + \eta, \ell = 0{:}M. \end{aligned} \quad (13)$$

Theorem 1 (sampled-to-continuous separation). *If (13) holds at every query point, the transform, tracking, and set-motion bounds hold, and E_k occurs, then the realized footprint remains at least η from represented dynamic and static hazards throughout the complete executable prefix.*

Proof. For any prefix time t, choose a query q_l with $|t-q_l|<=d_Q$. The Hausdorff perturbation inequality bounds the loss of set separation by the footprint, interpolated tube, and tracking rates in (13); the tracking envelope accounts for realization error. Subtracting this quantity from (13) leaves η. On E_k, replacing the contract tube by true represented occupancy can only preserve the lower bound. Figure 3 shows the two finite margins used by the gate.

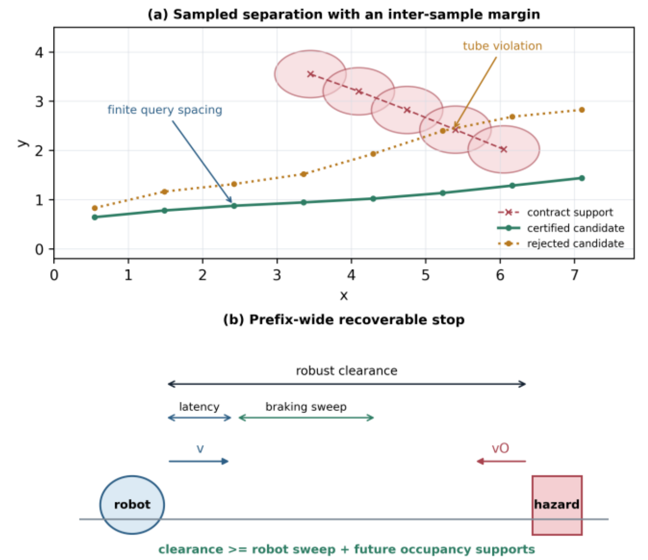

Fig. 3. Finite prefix certification. Sample checks reserve a Lipschitz inter-sample margin. At every possible interruption point, robust clearance must also contain the complete reaction-and-braking sweep against all future contract supports.

B. Swept Stop and Horizon-Aware Lease

A prefix is executable only if braking is recoverable from every point, not merely from its endpoint. Let τ bound sensing, gate reaction, command cancellation, and brake engagement. During τ, translational and angular accelerations are bounded by $\bar{a}_{\mathrm{lat}}$ and $\bar{\gamma}_{\mathrm{lat}}$; once engaged, braking magnitudes are at least a_b and γ_b. With $w=|\omega|$, define

$$\begin{aligned} v_\tau &= v + \bar{a}_{\mathrm{lat}}\tau, w_\tau = |\omega| + \bar{\gamma}_{\mathrm{lat}}\tau, \\ T_B &= \tau + \max\{v_\tau/a_b, w_\tau/\gamma_b\}, \\ \Theta_B &= |\omega|\tau + 1/2\,\bar{\gamma}_{\mathrm{lat}}\tau^2 + w_\tau^2/(2\gamma_b), \\ R_\theta &= 2r_B\sin(\min\{\Theta_B, \pi\}/2). \end{aligned} \quad (14)$$

The center displacement is bounded by D_B in (15). For a noncircular body, R_θ bounds the chord swept under the cumulative reaction-and-braking angle. The verified circumscribed disc used in the implementation has $R_\theta=0$, while angular braking still determines T_B. Rather than adding an independent closing-speed term, the gate compares the stopping ball with the union of the contract's time-indexed occupancy supports over the complete stopping interval:

$$
\begin{aligned}
D_B(v) &= v\tau + 1/2\,\bar{a}_{\mathrm{lat}}\tau^2 + v_\tau^2/(2a_b),\\
t_\ell &= t_{r,k} + q_\ell,\\
t_{\mathrm{sup}} &= \min\{t_{e,k}, s_k + H_k\Delta_k\},\\
t_{B,\ell} &= \min\{t_\ell + T_B(t_\ell), t_{\mathrm{sup}}\},\\
K_k &= \{s_k, s_k+\Delta_k, \dots, s_k + H_k\Delta_k\},\\
G_{k,\ell} &= \{t_\ell, t_{B,\ell}, t_{\mathrm{sup}}\} \cup \big(K_k \cap [t_\ell, t_{\mathrm{sup}}]\big).
\end{aligned}
\tag{15}
$$

The benchmark strengthens this test by extending the union to the contract support end, so the stopped ball also remains clear for the remaining support. It evaluates the exact stop start, stop end, support end, and interior future knots; the static corridor is tested against the same stopping ball. For prefix grid Q, d_Q is its fill distance. The future audit grid in (15) has fill distance d_U. Let L_D bound the rate of D_B along an admissible speed profile and let $L_{pre}=L_R+L_D+L_{R\theta}+L_\varepsilon$, where $L_{R\theta}$ bounds rotational-sweep radius variation. The finite certificate is

$$
\begin{aligned}
\mathcal{H}_k(u) &= \mathcal{S}_{\mathrm{static}} \cup \overline{\mathcal{U}}_k(u),\\
\rho_\ell &= r_B + D_B\big(v(t_\ell)\big) + R_\theta\big(\omega(t_\ell)\big) + \varepsilon_R(t_\ell),\\
b_\ell &= \min_{u\in G_{k,\ell}} \mathrm{dist}\big(\hat{p}(t_\ell), \mathcal{H}_k(u)\big) - \rho_\ell,\\
b_\ell &\geq L_{\mathrm{pre}} d_Q + L_U d_U + \eta_B.
\end{aligned}
\tag{16}
$$

Theorem 2 (prefix-wide swept-stop separation). *If (16) holds on the prefix grid and the declared reaction, acceleration, braking, tracking, and obstacle-motion bounds hold, then a stop initiated at any prefix time remains separated from represented occupancy until translational and angular rest; the stronger support-end audit preserves the stopped footprint until the activated contract support ends.*

Proof. The acceleration-aware displacement contains every center trajectory during reaction and braking; the rotational term contains every body orientation. The future-set union contains every represented obstacle support at its corresponding absolute time, so its motion is included once. The L_{pre} d_Q and L_U d_U residuals transfer each finite inequality to continuous interruption and future times. Contract coverage then contains the true represented occupancy for the complete stop interval.

For the reported disc-footprint benchmark, rotational invariance sharpens (11) to $L_R=\bar{v}_R$, $L_{R\theta}=0$, and $\varepsilon_R=\varepsilon_{tf}=0$ under the matched plant and common frame; activation requires numerical equality to the certified anchor. With $\bar{a}_R=\max\{a_{max,ab}$, the implemented analytic bound is $L_D=\bar{a}_R[\tau+(\bar{v}_R+\bar{a}_{lat\tau})/a_b]$; it uses d_Q=0.10 s, d_U=0.10 s, and the strictly positive margins $\eta=\eta_B$=0.06 m. Deployment values must instead include measured tracking and transform envelopes.

At release $t_{r,k}$, candidate j has executable duration $T_{c,k}{}^{j}$ and worst stopping time $T_{guard}{}^{j}$ over that prefix. Its atomic lease is admitted only when

$$
t_{r,k} + T_{c,k}^{j} + T_{\mathrm{guard}}^{j} \leq t_{\mathrm{sup}} \tag{17}
$$

Thus expiry alone is insufficient: the executable prefix and its complete stop must lie inside both freshness expiry and prediction support. A later provider cannot extend this lease. The old command is closed, the new contract identifier is bound, and the candidate is re-certified before publication.

C. Candidate Arbitration and Hysteresis

The gate certifies every concrete rollout from Continue, Slow, Replan, or other proposal sources. These labels are not assumed to be safety-ordered because their feasible sets are nonnested. After binary contract, separation, stop, lease, deadline, and platform checks, selection is restricted to

$$
\begin{aligned}
\mathcal{A}_k &= \{c : \mathrm{Cert}_k(c) = 1\},\\
\mathcal{J} &= \{g, \mathrm{clr}, \mathrm{turn}, \mathrm{mgr}, \mathrm{fuzzy}, \sigma\},\\
J_k(c) &= \textstyle\sum_{z\in\mathcal{J}} J_z(c),\\
c_k &= \arg\min_{c\in\mathcal{A}_k} J_k(c).
\end{aligned}
\tag{18}
$$

Here c=(m,j) pairs a behavior label with a concrete rollout. J_g contains terminal goal distance, heading, and effort; $J_{clr}=0.72[\max(0,1.35-c_{min})]^2$; $J_{mgr}=1.25|v_c-v_{\mu k}|$; J_{fuzzy} contains a 0.38 speed-deviation term and a 1.15 turn-away term scaled by fuzzy severity; and $J_\sigma=0.18/\max\{0.05,\sigma_k(c)+0.05\}$. The fixed J_{turn} also contains the deterministic detour and dense-flow hold costs. The residual term favors larger certified margin. All terms are evaluated only inside the certified set. Fuzzy inference follows Zadeh and Mamdani [21], [22] and contributes a soft preference; it never removes a certified candidate or admits an uncertified one. Empty feasibility invokes the verified stop path.

For candidate j, define residual safety margin $\sigma_k(j)$ as the minimum remaining separation and backup slack. The global intervention latch is driven by the residual of the exact certified candidate selected at each epoch; candidate identity may change across epochs. Escalation or Stop is immediate after invalidation or $\sigma_k(j)\leq h_{on}$. A one-level recovery is allowed only after selected margins satisfy $\sigma_k(j)\geq h_{off}>h_{on}$ for three consecutive accepted epochs and the two-tick dwell count. No safety order is assigned to behavior names.

D. Deterministic Assurance Update

Algorithm 1. RACE-Nav assurance update

1. Snapshot monotonic time, robot state, transforms, map version, tracks, identities, and health into one immutable record.
2. Select a predictable allowance, request the fixed-class conditional contract, and validate provider health, age, support, expiry, and ledger capacity.
3. Generate and time-anchor 25 differential-drive candidates at the proposed release; an anchor or deadline mismatch requires a fresh proposal.
4. Evaluate fuzzy preference but retain every concrete candidate for independent certification.
5. Evaluate (13), (16), (17), kinematic limits, and decision deadline; store margins and every rejection cause.
6. Rank only certified candidates. Atomically commit contract ID, candidate ID, risk charge, and watchdog deadline.
7. Publish through the sole command owner, or issue the verified stop on empty feasibility, invalidity, or deadline miss.
8. Atomically close the old lease on renewal, cancellation, watchdog expiry, or any provider substitution.

V. SYSTEM-LEVEL GUARANTEES

A. Trusted-Gate Invariant

Theorem 3 (trusted-gate execution safety). *Assume complete representation within the declared hazard scope, faithful sole-owner gate and watchdog execution, the deterministic bounds of Section IV, and E_k for every accepted contract. Then no robot-caused contact with represented static or dynamic occupancy occurs during any finite sequence of certified prefixes and their stopping intervals.*

Proof. Induct over releases. The first command is a subset of a prefix covered by Theorem 1. A replacement starts only after its synchronized certificate and anchor checks pass. An atomic compare-and-swap then replaces the old lease directly with the new certified lease; a failed or absent replacement selects the verified stop. Without replacement, invalidity or lease end triggers a stop from a point covered by Theorem 2, and (17) keeps the complete stop inside support. Sole command ownership prevents a nominal command from crossing a watchdog boundary. Every executed time is therefore inside a separated prefix or its separated stop tube.

The conclusion is passive only until rest. A dynamic obstacle may subsequently strike a stationary robot; that event is not converted into provider miscoverage. Likewise, an unrepresented birth or identity failure is a deterministic invalidation, not an event charged to α_k.

B. Conditional Mission Composition

Theorem 4 (finite/anytime mission risk). *Let C_T denote robot-caused contact before rest over the certified mission. If each accepted contract satisfies (4), or supplies a selection-valid PAC outer event whose combined conditional failure is bounded by $\alpha_k+\beta_k$, A_k is the atomic activation indicator, A_k, α_k, and β_k are predictable, the ledger charges only successful activations and enforces (5), and Theorem 3's deterministic premises hold, then*

$$\begin{aligned} C_T &\subseteq \bigcup_{k\geq 0}\{A_k=1, E_k^c\}, \\ \mathbb{P}(C_T) &\leq \mathbb{E}\left[\textstyle\sum_{k\geq 0} A_k\,(\alpha_k+\beta_k)\right] \leq \delta. \end{aligned} \tag{19}$$

Proof. Theorem 3 derives the event inclusion: contact on the joint coverage event is impossible. Applying the union bound and tower property to the admitted conditional bounds yields (19). No cross-epoch independence is used. A PAC-conditional provider consumes its selection-valid outer β_k allowance; deterministic contracts charge zero.

For an unknown number of attempts, the following predictable schedule is summable. The provider must support each requested tail level; otherwise, the attempt is rejected. A PAC provider splits the same summable charge between α_k and β_k.

$$\alpha_k = \frac{6\delta}{\pi^2(k+1)^2} \Rightarrow \textstyle\sum_{k=0}^{\infty} \alpha_k = \delta \tag{20}$$

Ordinary split conformal prediction supplies marginal finite-sample coverage for a nonadaptive exchangeable test item, not (4). Its exact order statistics and minimum finite calibration size are

$$\begin{aligned} r &= \lceil (n+1)(1-\alpha) \rceil, \\ q_\alpha &= S_{(r)} \text{or } +\infty, n \geq \alpha^{-1}-1. \end{aligned} \tag{21}$$

At α=0.0005, at least 1,999 calibration scenes are needed for a nontrivial split quantile; the archive uses 49,999 only as a single-update marginal diagnostic. It is never consumed by adaptive ledger.

C. Switching, Stale Commands, and Liveness

Corollary 1 (stale-command exclusion). *If the watchdog is armed before publication and owns the final command channel downstream of smoothing, no command issued under contract k remains active after its prefix deadline, invalidation, or support boundary.*

Proof. Each publication carries the immutable lease identifier and deadline. One atomic compare-and-swap closes the old label and installs either a fully certified successor or the verified stop; an overdue or canceled proposal cannot acquire ownership.

Assurance ticks and watchdog callbacks have positive periods and are locally finite. Each callback is atomic and causes at most one transition; Stop remains latched until later accepted epochs. Because both recovery counters are discrete and reset after every one-level recovery, successive recoveries are separated by

$$\Delta t_{\text{recover}} \geq \max\{N_{\text{clear}}, N_{\text{dwell}}\} T_{\text{tick}} > 0 \tag{22}$$

Here T_{tick} is the positive assurance-tick period and $\tau_d=N_{\text{dwell}} T_{\text{tick}}$ is the dwell contribution used in the release bound below; the reported configuration uses $N_{\text{clear}}=3$ and $N_{\text{dwell}}=2$.

Theorem 5 (non-Zeno switching). *Under the locally finite callback and latch conditions, every finite interval contains finitely many mode switches. Equation (22) additionally lower-bounds every pair of one-level recovery switches. Immediate escalations do not threaten non-Zeno behavior because they occur only on the same locally finite callbacks.*

Theorem 6 (conditional release liveness). *Suppose from t_0 a concrete moving candidate remains available with margin at least h_{off}, accepted contract gaps are at most Δ_{max}, each committed moving interval from activation to atomic replacement or stop lasts at least T_{min}, the ledger has sufficient balance, planner time is at most τ_p, no invalidation occurs, and the ranking rule selects a progress candidate whenever one exists. Then a moving command is released no later than $t_0+\tau_p+\tau_d+(N_{clear}+1)\Delta_{max}$. If V is nonnegative and vanishes exactly on the goal set, and each actual committed moving interval outside that set decreases V by ε_V, the goal is reached within ceil($V(x_0)/\varepsilon_V$) such intervals.*

Proof. The candidate remains in the certified set. Planner completion, dwell, and the required clear observations therefore finish within the stated intervals, with one further accepted-update gap for atomic release. Here T_{min} is the committed duration required by the stated progress certificate. The actual command intervals partition execution at atomic replacement times, so the progress premise is not stated on overlapping candidate horizons. Thereafter every committed progress interval decreases V; more than the stated number would make V negative. No unconditional progress claim is made for blockage, an exhausted ledger, or adversarial crowds.

VI. IMPLEMENTATION AND EXPERIMENTAL METHOD

A. Executable Assurance Kernel and Nav2 Boundary

The executable package separates tracking/provider logic, fixed-schema conditional-contract validity checking, trajectory rollout, finite-grid certification, mission ledger, hysteretic arbitration, final command ownership, watchdog, and audit logging. The full controller evaluates 25 unicycle commands every 0.3 s over a 1.2-s prefix. It uses the conditional posterior radial-union provider of Section III-C, acceleration-limited release-anchored rollouts, a circumscribed-disc footprint, the inter-sample residual, prefix-wide future-set backup closure, the exact horizon-and-stop lease, and a finite ledger. Every decision stores timestamps, anchor state, immutable contract and lease identifiers, candidate index, proposal and activation events, expiry, health, margins, activation charge, command, close reason, and elapsed time. The instantiated provider is conditional with β=0; proposal creation reserves no risk, and the ledger changes only inside the successful activation transaction.

For a ROS 2/Nav2 deployment, navigation output is an untrusted proposal. Certification must occur inside an instrumented controller or adapter that exposes the complete timed rollout before computeVelocityCommands returns. The assurance gate publishes the only base command downstream of velocity smoothing; action cancellation alone is not treated as proof of rest. Deadline/lifespan QoS, monotonic-clock age checks, pinned map/transform versions, and a separately implemented downstream watchdog are deployment obligations. The reported measurements are from the standalone deterministic kernel; no ROS callback or hardware claim is inferred. Table II maps each executable module to its auditable evidence.

TABLE II
IMPLEMENTATION MODULES AND AUDITABLE OUTPUTS

Module	Input	Output / rejection evidence
Provider	time-stamped tracks, requested α	map-frame disc tube, IDs, class, support, expiry
Validity gate	packet, age, association, IDs, occlusion	valid contract or explicit reason
Rollout	state and proposal sources	25 timed differential-drive candidates
Certifier	candidate, tube, brake bounds	prefix/backup margins and binary certificate
Ledger	attempt index and allowance	activation charge, balance, or rejection
Gate/watchdog	ranked certified set and lease	sole velocity command or verified stop
Audit	immutable decision snapshot	replayable record and platform latency

B. Scenarios, Parameters, and Hold-Out Design

Eight pre-specified interactions cover crossing, head-on approach, overtaking, sudden entry, stop-go motion, blockage, mapped occlusion, and dense bidirectional flow. Obstacle scripts are independent of the controller; every method receives the same seeded trajectory and observation stream. Seeds 0-1 were confined to smoke checks. Final core seeds 10000-10039, stress seeds 20000-20011, nominal-ablation seeds 30000-30019, and targeted-ablation seeds 40000-40019 were held out from development. Figure 4 depicts the scenario families, and Table III records the fixed configuration.

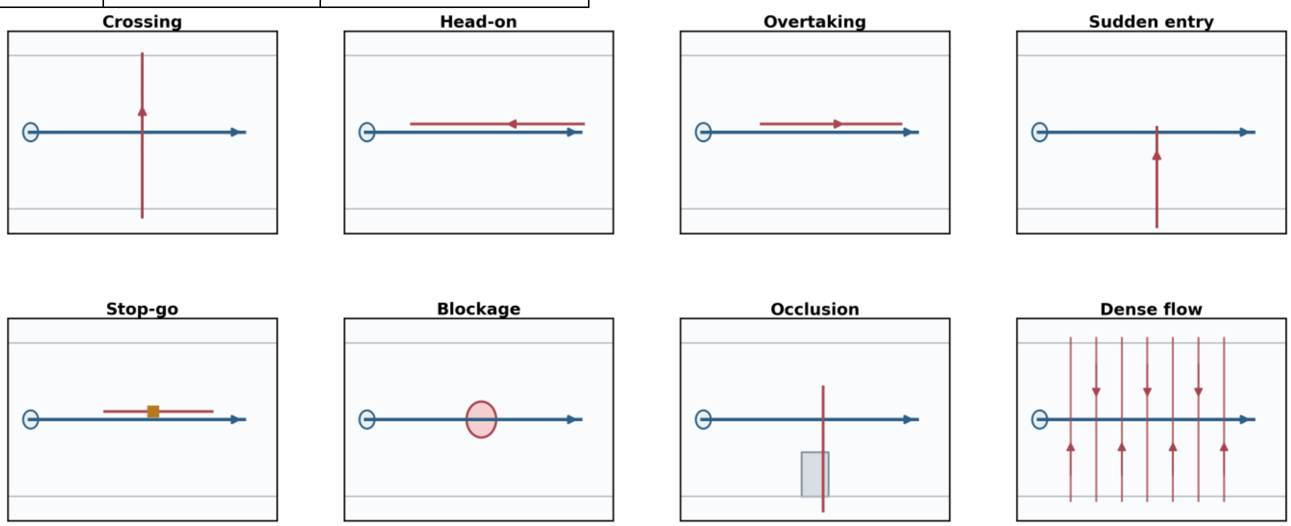

Fig. 4. The eight controlled dynamic-interaction families. Blue denotes the robot route; red denotes dynamic-agent motion. All controllers share the same scenario instance and held-out seed.

TABLE III
FIXED BENCHMARK CONFIGURATION

Quantity	Value	Quantity	Value	Quantity	Value
Integration / decision	0.1 / 0.3 s	Robot radius	0.31 m	Max speed	1.00 m/s
Prefix / horizon	1.2 / 3.0 s	Cert. / pred. grid	0.2 / 0.2 s	Freshness	0.28 s
Reaction latency	0.20 s	Accel. / brake	0.85 / 0.95 m/s²	Obstacle accel.	0.30 m/s²
Angular accel. / brake	2.80 / 2.80 rad/s²	Latency accel. lin. / ang.	0.15 / 0.20	Contract lease	3.00 s
Positive η / η_B	0.06 / 0.06 m	Corridor half-width	3.2 m	Anchor gates	all 1e-10 in SI units
Latent prior σ_p / σ_v	0.180 / 0.065	Meas. σ_p / σ_v	0.025 / 0.035	Transform / tracking	0 / 0 (matched)
Hysteresis on / off	0.16 / 0.38 m	Clear / dwell cycles	3 / 2	Max angular speed	1.25 rad/s
Mission budget	0.047	Per-contract α	0.0005	Episode limit	30 s

C. Comparators, Outcomes, and Statistics

Six controlled implementations are evaluated: nominal goal seeking; hard distance/time-to-collision (TTC) thresholds; a fuzzy supervisor without certification; short-horizon predictive TTC/model-predictive control (MPC) informed by uncertainty-aware planning practice; a discrete point-prediction safety filter; and full RACE-Nav. The first five are benchmark implementations, not embedded copies of production Nav2, MPPI, or third-party solvers. They share dynamics, limits, observation streams, and candidate resolution.

Collision is footprint overlap at any simulator step; goal success requires reaching the 0.38-m goal region within 30 s; clearance is signed footprint-to-obstacle distance. Core testing uses 40 paired seeds per scenario-method cell (320 episodes/method). Proportions use Wilson 95% intervals [23]. Paired collision differences use deterministic 10,000-resample bootstrap intervals and exact McNemar tests [24] with Holm correction. Stress cells pool 12 seeds over eight scenarios (96 episodes). Ablation cells use 20 seeds per scenario or targeted pair. Wall latency is descriptive; a deterministic modeled-latency channel drives faults. The predictive comparator's uncertainty-aware planning structure follows [25].

VII. RESULTS

A. Core Paired Comparison

Table IV gives the paired core outcomes; Figures 5 and 6 show collision uncertainty and the full clearance distributions.

TABLE IV
HELD-OUT CORE RESULTS OVER 320 PAIRED EPISODES PER CONTROLLER

Controller	Collisions	Rate [95% Wilson]	Success [95% Wilson]	Median clr. (m)	5th pct. clr. (m)
Nominal goal seeking	259/320	80.94% [76.28%-84.86%]	19.06% [15.14%-23.72%]	-0.035	-0.101
Hard threshold	93/320	29.06% [24.36%-34.26%]	61.56% [56.13%-66.73%]	0.327	-0.050
Fuzzy supervisor	264/320	82.50% [77.96%-86.27%]	17.50% [13.73%-22.04%]	-0.032	-0.101
Predictive TTC/MPC	39/320	12.19% [9.04%-16.23%]	84.06% [79.65%-87.67%]	0.473	-0.054
Point safety filter	27/320	8.44% [5.86%-12.00%]	90.62% [86.93%-93.35%]	0.075	-0.018
RACE-Nav	0/320	0.00% [0.00%-1.19%]	97.50% [95.15%-98.73%]	1.113	0.525

RACE-Nav recorded 0/320 collisions; its 95% Wilson interval is [0.00%, 1.19%], so the observed count is not reported as a population risk. Goal success was 97.50% and median minimum clearance was 1.113 m. The predictive TTC/MPC comparator recorded 39/320 (12.19%). The paired risk difference was -12.19 percentage points with 95% interval [-15.94, -8.75]. Every exact paired comparison remained significant after Holm correction (largest adjusted p=1.49e-08).

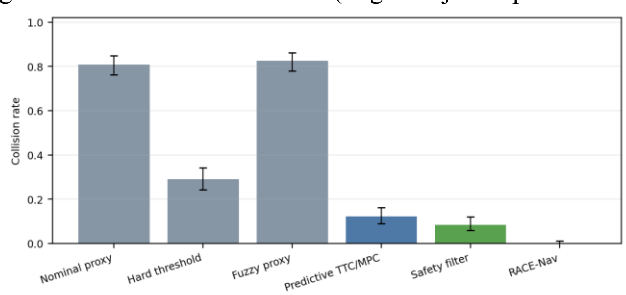

Fig. 5. Held-out core collision rates with Wilson 95% intervals. Counts are printed above each estimate; all methods use the same 320 paired scenario-seed episodes.

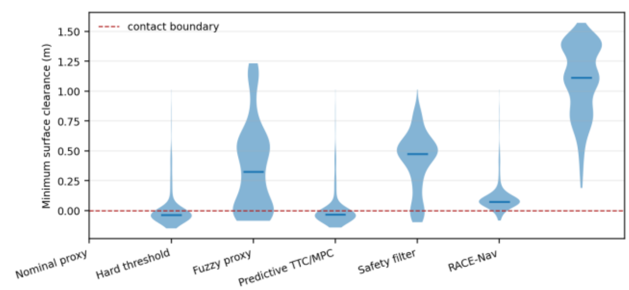

Fig. 6. Distribution of minimum signed clearance in the paired core benchmark. Negative values indicate overlap; embedded boxes show the median and interquartile range for every controller.

RACE-Nav recorded 8 timeouts, distributed as dense flow (1), head-on (5), overtaking (1), and stop-go (1). This is an observed conservative liveness cost, not recoded success. Among 16,885 valid core contract checks, 0 coverage failures occurred (0.00e+00); that empirical rate is diagnostic and does not replace the model-conditional argument.

B. Assumption-Violation Stress Tests

Table V reports the full controller under each declared-assumption stress; Figure 7 compares raw contact rates across all controllers.

TABLE V
FULL-CONTROLLER STRESS BEHAVIOR; 96 HELD-OUT EPISODES PER CONDITION

Stress	Raw contacts	Contact rate	Success	Mechanism diagnostic
Brake degradation	0/96	0.00%	95.83%	brake bound violated
Compute delay	4/96	4.17%	0.00%	9,484 modeled deadlines missed
Delayed state	4/96	4.17%	0.00%	9,422 stale-contract rejections
Distribution shift	16/96	16.67%	71.88%	4,026 coverage failures
Detection dropout	3/96	3.12%	3.12%	3,390 validity rejections
Identity switch	1/96	1.04%	84.38%	631 identity rejections
Increased noise	0/96	0.00%	94.79%	larger robust sets
Packet loss	2/96	2.08%	4.17%	2,627 validity rejections
Wheel slip	3/96	3.12%	0.00%	coverage failure; slip violates plant model

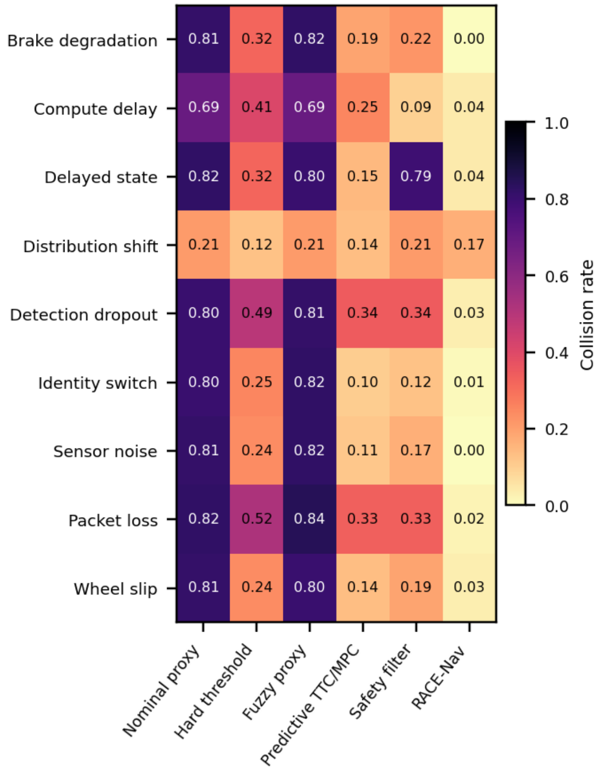

Fig. 7. Raw contact rate under nine stresses. Distribution shift, degraded braking, delayed state, and slip can violate provider, plant, or timing premises; these rows are outside the nominal mission guarantee.

Raw contact deliberately counts post-stop impacts. Delayed state produced 4/96 contacts and 0/96 successes after stale information triggered 9,422 validity rejections. Distribution shift intentionally violated the obstacle-acceleration envelope, producing 4,026 coverage failures and 16/96 contacts. These outcomes expose the theorem boundary rather than being folded into its δ.

C. Mechanism Ablations

Table VI ties each removal to a mechanism-specific diagnostic, while Figure 8 summarizes nominal collision and success outcomes.

TABLE VI
ABLATION OUTCOMES TIED TO EACH MECHANISM

Removed mechanism	Observed effect
Inter-sample margin	1/160 nominal contacts; 280 counterfactual rejects
Expiry + horizon	2,020 unsupported releases under compute delay
Prefix backup	352 commands rejected by the full backup test
Mission ledger	aggregate excess 0.4820 in low-budget crossing
Hysteresis	manager switches 3.65 → 18.45
Fuzzy preference	overtaking success 95% → 100% when removed

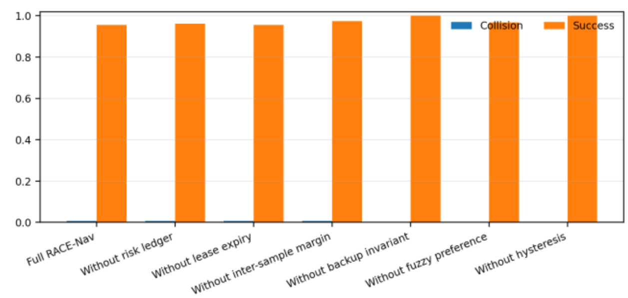

Fig. 8. Nominal all-scenario collision and goal-success rates for the full controller and six ablations. Mechanism-specific unsupported-release, counterfactual-rejection, ledger, and switching diagnostics are reported in Table VI.

The no-inter-sample-margin ablation recorded 1/160 nominal contacts. Expiry removal recorded 2,020 unsupported releases, and backup removal selected 352 commands that the full certificate rejected; the targeted backup trials recorded 0/20 contacts. Thus, absence of observed contact is not treated as proof. Removing fuzzy preference changed targeted overtaking success from 95% to 100% and manager switches from 1.85 to 3.40; fuzzy logic remains an efficiency preference whose measured effect is reported rather than assumed.

D. Decision Latency and Reproducibility

Across 18,125 core decisions, wall latency was 19.32 ms mean, 68.66 ms p95, 93.02 ms p99, 125.38 ms p99.9, and 134.35 ms maximum. No modeled 300-ms core decision deadline was missed. These platform measurements are not a hard real-time certificate; the compute-delay stress uses the deterministic modeled-latency channel. Figure 9 gives the complete empirical latency distributions.

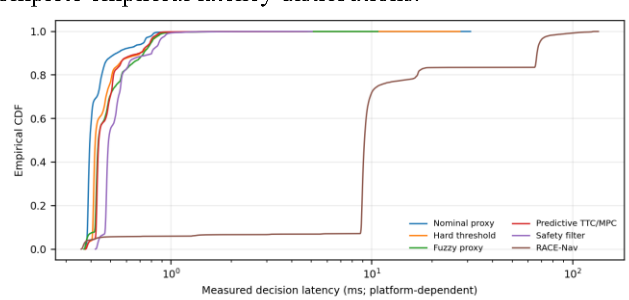

Fig. 9. Empirical decision-latency distributions in the core benchmark. The full controller maximum is 134.35 ms on the recorded platform; the 300-ms command period lies beyond the displayed range.

The final suite contains 8,464 episodes and 375,054 decision records. Deterministic outcome tables exclude wall time, and a SHA-256 manifest covers packaged results. Re-execution with the recorded environment and seeds passed 31 unit/mechanism tests plus byte-stable outcome verification.

VIII. DISCUSSION

A. What the Probability Statement Covers

The finite/anytime δ in Theorem 4 is a bound on robot-caused contact attributable to accepted occupancy-contract failures while deterministic representation, localization/tracking, transformation, timing, kinematic, obstacle-motion, and braking assumptions hold. It is not a certificate for total collision probability. Births, missed detections, stale state, software faults, localization excursions, actuator degradation, and post-stop impacts require separate evidence or invalidate motion. Reporting both raw contact and contract diagnostics prevents those distinct events from being collapsed into one favorable number.

B. Safety-Efficiency Tradeoff

The core benchmark combines 0 observed RACE-Nav contacts with 97.50% success and 8 transparent timeouts. The ledger, prefix backup, and robust tube can each stop motion when statistical or deterministic support is insufficient. This is the intended fail-safe semantics, but it does not solve crowd deadlock. The conditional liveness theorem identifies the additional availability and progress premises needed for goal completion; the overtaking and stop-go results show where those premises were absent.

C. Deployment Evidence Boundary

The standalone kernel establishes algorithmic execution and simulation behavior, not physical braking, sensor completeness, separate stop circuitry, or ROS 2 callback ordering. A deployment claiming Theorem 3 must identify worst-case reaction/braking envelopes on the actual platform, verify the footprint and transform error, place the watchdog downstream of every command modifier, demonstrate sole command ownership, and repeat timing and dynamic-agent tests under the pinned middleware configuration. The contract fields make these obligations auditable without changing the theorem.

IX. CONCLUSION

RACE-Nav turns replaceable prediction and navigation components into a typed, time-bounded command contract. Its model-conditional provider matches the ledger premise; finite sampled margins establish continuous-time separation; prefix-wide stopping and horizon-aware leases close the interruption gap; and the trusted-gate invariant derives the mission event inclusion. Held-out comparisons, mechanism ablations, fault/shift tests, and complete decision traces quantify both its collision reduction and conservative liveness cost. The result is a reproducible assurance architecture whose probability claim, deterministic obligations, and command authority are explicit at the same interface.

MULTIMEDIA MATERIALS

The accompanying Multimedia Materials archive contains the executable benchmark, complete held-out raw trial and decision data, deterministic outcome tables, statistical summaries, vector/PNG figures, fixed configuration and seed policy, unit tests, an integrity verifier, and a SHA-256 manifest. The paper's claims, methods, equations, and numerical conclusions are self-contained; the archive enables independent replay.

APPENDIX A. PROOF DETAILS AND GENERAL FOOTPRINTS

A. Conditional Provider

Condition on the pre-revelation history. The linear-Gaussian posterior makes each future latent position residual Gaussian with covariance $\Sigma_{i,h}$. Writing it as Lg for standard two-dimensional Gaussian g gives $||Lg|| \leq \mathrm{sqrt}(\lambda_{max}(\Sigma_{i,h}))||g||$. The Rayleigh tail of $||g||$ and the multiplier in (8) limit each pair's conditional failure probability to its equal share of the requested allowance. On the joint endpoint event, the latent residual between adjacent knots is affine in time; convexity of its norm makes its interior value no larger than the linear interpolation of the two endpoint radii. For the scripted mean, the interpolation-error identity extends to C^1 curves with $\bar{a}_O$-Lipschitz velocity and gives the chord remainder $\bar{a}_{O\Delta k}^2/8$; piecewise acceleration may jump on a measure-zero set. Adding that remainder and the physical radius therefore contains the complete continuous-time footprint between knots. A union bound over the finite agent-knot index establishes (4), regardless of temporal correlation. Missing representation assigns infinite score and cannot be hidden inside a finite radius.

B. General Rotating Footprint

For compact body B with circumscribed radius r_B, rotation through angle φ moves any body point by at most the corresponding chord length. The cumulative angle bound in (14) contains reaction and angular braking. Minkowski addition of its rotational-sweep ball and the center-displacement ball therefore contains the entire noncircular swept footprint. The disc implementation is invariant under rotation and sets this geometric term to zero, while retaining angular stopping time in the lease and future-occupancy sweep.

C. Provider Composition

For a union tube over the same scope, failure implies simultaneous failure of all provider events, giving the minimum individual allowance without an independence assumption. For an intersection tube, success requires every event, so the union bound gives the sum. Disjoint scopes use a spatial union and summed allowance. Measurable adaptive composition remains valid only when each admitted provider supplies its declared history-conditional class at the requested allowance.

REFERENCES

[1] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, and W. Woodall, "Robot Operating System 2: Design, architecture, and uses in the wild," Sci. Robot., vol. 7, no. 66, Art. no. eabm6074, 2022, doi: 10.1126/scirobotics.abm6074.

[2] S. Macenski, F. Martin, R. White, and J. Gines Clavero, "The Marathon 2: A navigation system," in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), 2020, pp. 2718-2725, doi: 10.1109/IROS45743.2020.9341207.

[3] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," IEEE Robot. Autom. Mag., vol. 4, no. 1, pp. 23-33, 1997, doi: 10.1109/100.580977.

[4] K. L. Hobbs, M. L. Mote, M. C. L. Abate, S. D. Coogan, and E. M. Feron, "Runtime assurance for safety-critical systems: An introduction to safety filtering approaches for complex control systems," IEEE Control Syst. Mag., vol. 43, no. 2, pp. 28-65, 2023, doi: 10.1109/MCS.2023.3234380.

[5] S. Shivakumar, H. Torfah, A. Desai, and S. A. Seshia, "SOTER on ROS: A run-time assurance framework on the Robot Operating System," in Runtime Verification, LNCS 12399, 2020, pp. 184-194, doi: 10.1007/978-3-030-60508-7_10.

[6] K. P. Wabersich and M. N. Zeilinger, "A predictive safety filter for learning-based control of constrained nonlinear dynamical systems," Automatica, vol. 129, Art. no. 109597, 2021, doi: 10.1016/j.automatica.2021.109597.

[7] K. J. Strawn, N. Ayanian, and L. Lindemann, "Conformal predictive safety filter for RL controllers in dynamic environments," IEEE Robot. Autom. Lett., vol. 8, no. 11, pp. 7833-7840, 2023, doi: 10.1109/LRA.2023.3322644.

[8] L. Lindemann, M. Cleaveland, G. Shim, and G. J. Pappas, "Safe planning in dynamic environments using conformal prediction," IEEE Robot. Autom. Lett., vol. 8, no. 8, pp. 5116-5123, 2023, doi: 10.1109/LRA.2023.3292071.

[9] A. Dixit, L. Lindemann, S. X. Wei, M. Cleaveland, G. J. Pappas, and J. W. Burdick, "Adaptive conformal prediction for motion planning among dynamic agents," Proc. Mach. Learn. Res., vol. 211, pp. 300-314, 2023.

[10] C. Stamouli, L. Lindemann, and G. J. Pappas, "Recursively feasible shrinking-horizon MPC in dynamic environments with conformal prediction guarantees," Proc. Mach. Learn. Res., vol. 242, pp. 1330-1342, 2024.

[11] J. Shin, J. Lee, and I. Yang, "Egocentric conformal prediction for safe and efficient navigation in dynamic cluttered environments," in Proc. 64th IEEE Conf. Decis. Control (CDC), 2025, pp. 4910-4917, doi: 10.1109/CDC57313.2025.11312645.

[12] J. Yao, X. Zhang, Y. Xia, Z. Wang, A. Roy-Chowdhury, and J. Li, "Towards generalizable safety in crowd navigation via conformal uncertainty handling," Proc. Mach. Learn. Res., vol. 305, pp. 4206-4225, 2025.

[13] K. Liang, L. Luo, Y. Wang, M. Cai, and C. I. Vasile, "Time-aware motion planning in dynamic environments with conformal prediction," Proc. Mach. Learn. Res., vol. 331, pp. 183-195, 2026.

[14] H. Huang, S. Sharma, A. Loquercio, A. N. Angelopoulos, K. Goldberg, and J. Malik, "Conformal policy learning for sensorimotor control under distribution shifts," in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), 2024, pp. 16285-16291, doi: 10.1109/ICRA57147.2024.10610223.

[15] R. F. Barber, E. J. Candès, A. Ramdas, and R. J. Tibshirani, "The limits of distribution-free conditional predictive inference," Inf. Inference, vol. 10, no. 2, pp. 455-482, 2021, doi: 10.1093/imaiai/iaaa017.

[16] I. Gibbs and E. J. Candès, "Conformal inference for online prediction with arbitrary distribution shifts," J. Mach. Learn. Res., vol. 25, no. 162, pp. 1-36, 2024.

[17] L. Blackmore, M. Ono, and B. C. Williams, "Chance-constrained optimal path planning with obstacles," IEEE Trans. Robot., vol. 27, no. 6, pp. 1080-1094, 2011, doi: 10.1109/TRO.2011.2161160.

[18] O. de Groot, B. Brito, L. Ferranti, D. M. Gavrila, and J. Alonso-Mora, "Scenario-based trajectory optimization in uncertain dynamic environments," IEEE Robot. Autom. Lett., vol. 6, no. 3, pp. 5389-5396, 2021, doi: 10.1109/LRA.2021.3074866.

[19] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," IEEE Trans. Autom. Control, vol. 62, no. 8, pp. 3861-3876, 2017, doi: 10.1109/TAC.2016.2638961.

[20] T. Fraichard and H. Asama, "Inevitable collision states-A step towards safer robots?" Adv. Robot., vol. 18, no. 10, pp. 1001-1024, 2004, doi: 10.1163/1568553042674662.

[21] L. A. Zadeh, "Fuzzy sets," Inf. Control, vol. 8, no. 3, pp. 338-353, 1965, doi: 10.1016/S0019-9958(65)90241-X.

[22] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," Int. J. Man-Mach. Stud., vol. 7, no. 1, pp. 1-13, 1975, doi: 10.1016/S0020-7373(75)80002-2.

[23] E. B. Wilson, "Probable inference, the law of succession, and statistical inference," J. Amer. Stat. Assoc., vol. 22, no. 158, pp. 209-212, 1927, doi: 10.1080/01621459.1927.10502953.

[24] Q. McNemar, "Note on the sampling error of the difference between correlated proportions or percentages," Psychometrika, vol. 12, pp. 153-157, 1947, doi: 10.1007/BF02295996.

[25] E. Schöneberg, M. Schröder, D. Görges, and H. D. Schotten, "Trajectory planning with model predictive control for obstacle avoidance considering prediction uncertainty," IFAC-PapersOnLine, vol. 59, no. 18, pp. 349-354, 2025, doi: 10.1016/j.ifacol.2025.10.245.