FASTFUNNELS: Appendix

Anonymous Author(s)

I. OVERVIEW

This appendix gives the full definitions behind Sections III–VI of the main text: the single-track vehicle model and its actuator servos (App. II), the leader MDP (App. III), the follower DMPC problem with its information set, funnel prediction, slot assignment, separating planes and cost terms (App. IV), the PPO training setup (App. V), the 14 benchmark scenes and the evaluation protocol (App. VI–VII), the baseline and ablation configurations (App. VIII), and the hardware deployment (App. IX). Equation numbers prefixed with "A" refer to this appendix; unprefixed numbers refer to the main text. Alg. 1 summarises both phases.

Two time steps appear throughout: the simulator and control tick $\Delta t = 0.01\,\mathrm{s}$ (100 Hz), at which the leader acts and every follower re-solves its DMPC problem, and the DMPC prediction step $\Delta\tau = 0.025\,\mathrm{s}$, with horizon $H = 40$ steps (1.0 s).

II. VEHICLE DYNAMICS AND ACTUATION

A. Frenet frame

Both phases project a world-frame position onto the track centerline $\mathcal{C}$,

$$\Pi_{\mathcal{C}} : (x_t, y_t) \mapsto (\sigma_t, e_{y,t}), \tag{A1}$$

where $\sigma_t \in [0, L]$ is the arc-length progress along $\mathcal{C}$, L is the centerline length, and $e_{y,t}$ is the signed perpendicular distance to $\mathcal{C}$ (positive on the left). $\psi_{\mathcal{C}}(\sigma)$ and $\kappa_{\mathcal{C}}(\sigma)$ denote the tangent heading and signed curvature of $\mathcal{C}$ at σ. Observations built from $(\sigma_t, e_{y,t})$ rather than (x_t, y_t) do not depend on the global layout of the track.

B. Single-track dynamic model

The patch car and the followers use the single-track dynamic model of the F1TENTH gym. With the state and input of Eqs. (1)–(2) of the main text,

$$\xi_t = [x_t\ y_t\ \delta_t\ v_t\ \psi_t\ \dot\psi_t\ \beta_t]^\top, \qquad u_t = [\alpha_t\ \dot\delta_t]^\top, \tag{A2}$$

the dynamics $\dot\xi_t = f(\xi_t, u_t)$ of Eq. (3) of the main text are

$$\dot\xi = \big[\, v\cos(\psi{+}\beta),\ v\sin(\psi{+}\beta),\ \dot\delta,\ \alpha,\ \dot\psi,\ \ddot\psi(\xi,\alpha),\ \dot\beta(\xi,\alpha)\,\big]^\top, \tag{A3}$$

with

$$\begin{aligned}\ddot\psi &= \frac{\mu m}{I_z(l_f+l_r)}\Big[\, l_f C_{S,f} F_f\,\delta + \big(l_r C_{S,r} F_r - l_f C_{S,f} F_f\big)\beta \\ &\qquad - \big(l_f^2 C_{S,f} F_f + l_r^2 C_{S,r} F_r\big)\frac{\dot\psi}{v}\Big], \\ \dot\beta &= \frac{\mu}{v(l_f+l_r)}\Big[\, C_{S,f} F_f\,\delta - \big(C_{S,r} F_r + C_{S,f} F_f\big)\beta \\ &\qquad + \big(C_{S,r} F_r l_r - C_{S,f} F_f l_f\big)\frac{\dot\psi}{v}\Big] - \dot\psi,\end{aligned} \tag{A4}$$

where $F_f = g l_r - \alpha h$ and $F_r = g l_f + \alpha h$ model the longitudinal load transfer. $\ddot\psi$ and $\dot\beta$ therefore depend on α, so the model is not control-affine. The equations are singular as $v \to 0$; for $v < 0.5\,\mathrm{m/s}$ the gym switches to the

Algorithm 1: FASTFUNNELS: leader training and online follower control

Input: map pool $\mathcal{P}$; E parallel environments; rollout length n per environment; total steps T_0; PPO hyper-parameters $(\varepsilon, c_1, c_2, \gamma, \lambda)$; DMPC horizon H; Gauss–Seidel sweeps S; damping η

Output: leader policy π_{θ^*}; follower inputs $u_t^i(0)$, $i = 1, \dots, N$, at every tick

`// Phase 1: leader (funnel) training, offline`

Initialise π_θ, the observation/reward normaliser, and E environments;

for $t = 1$ **to** T_0/E **do**

  foreach *environment* $e = 1, \dots, E$ **do**

    if *episode ended* **then** reset on a map $M \sim \mathcal{P}$ (held for 4 episodes);

    $o_t \leftarrow Z^{\mathrm{patch}}(s_t)$; `// Frenet, vehicle/funnel, LiDAR (Eq. (A7))`

    $u_t^{\mathrm{patch}} = (\bar\delta_t, \bar v_t, a_t^{\mathrm{cmd}}, b_t^{\mathrm{cmd}}) \leftarrow \pi_\theta(o_t)$ decoded by Eqs. (A9)–(A10);

    Step the single-track car with the servos of Eqs. (A5)–(A6);

    Update (a_t, b_t) by the rate law, Eq. (A12);

    Collect r_t^{patch} (Eq. (A14)); update the normaliser statistics;

  if $t \bmod n = 0$ **then** $\theta \leftarrow \theta + \alpha\,\nabla_\theta L^{\mathrm{patch}}(\theta)$ (Eq. (12) of the main text);

$\theta^* \leftarrow \theta$;

`// Phase 2: online follower control, no learning`

for $t = 1, 2, \dots$ **do**

  $u_t^{\mathrm{patch}} \leftarrow \pi_{\theta^*}(o_t)$ (mean action) ; `// funnel command`

  Predict the funnel $\{\hat c_t(k), \hat\varphi_t(k), \hat v_t(k), \hat a_t(k), \hat b_t(k)\}_{k=0}^{H}$ (Eq. (A16));

  Compute slot references $q_t^{i,\mathrm{slot}}(k)$ (App. IV-C);

  for $s = 1$ **to** S **do**

    for $i = 1$ **to** N **do** `// Gauss--Seidel: i sees the new plans of 1,...,i−1`

      Build the separating planes from $\hat q_t^j(0{:}H)$, $j \neq i$, and the predicted centre $\hat q_t^0(k) = \hat c_t(k)$;

      Solve the DMPC problem (Eqs. (A17)–(A21)) $\rightarrow$ new plan $\tilde q_t^i(0{:}H)$;

      $\hat q_t^i \leftarrow (1-\eta)\,\hat q_t^i + \eta\,\tilde q_t^i$ and broadcast;

  Apply u_t^{patch} and each follower's first input $u_t^i(0) = (\alpha_t^i, \dot\delta_t^i)$; advance the simulator by Δt;

TABLE AI
SINGLE-TRACK PARAMETERS (F1TENTH GYM DEFAULTS).

Symbol	Description	Value
m	mass	3.74 kg
I_z	yaw inertia	0.04712 kg m^2
$l_f,\ l_r$	CoM to front / rear axle	0.15875, 0.17145 m
h	CoM height	0.074 m
μ	friction coeff. (patch / followers)	1.0 / 1.0489
$C_{S,f},\ C_{S,r}$	cornering-stiffness coefficients	4.718, 5.4562
$\delta_{\max}$	steering-angle limit	0.4189 rad
$\dot{\delta}_{\max}$	steering-rate limit	3.2 rad/s
$a_{\max}$	acceleration limit	9.51 m/s^2
v_{sw}	switching speed	7.319 m/s
$[v_{\min}, v_{\max}]$	speed limits	$[-5, 20]$ m/s
—	body width × length	0.31 m × 0.58 m
Δt	integration step (RK4)	0.01 s

kinematic single-track model with $\beta = \arctan\big(\frac{l_r}{l_f+l_r}\tan\delta\big)$ and $\dot{\psi} = v\cos\beta\tan\delta/(l_f+l_r)$. The model is integrated with fixed-step RK4 at Δt, with u_t held constant over each step. Table AI lists the parameters.

C. Actuator servos

The leader issues setpoints $(\bar{\delta}_t, \bar{v}_t)$ (Eq. (5) of the main text); the gym's servos convert them into the model inputs $u_t = (\alpha_t, \dot{\delta}_t)$. The steering servo is bang-bang at the rate limit,

$$\dot{\delta}_t = \dot{\delta}_{\max}\,\mathrm{sgn}(\bar{\delta}_t - \delta_t) \quad \text{if } |\bar{\delta}_t - \delta_t| > 10^{-4},\ \text{else } 0. \tag{A5}$$

The speed servo is proportional, $\alpha_t = k_p(\bar{v}_t - v_t)$, with $k_p = 10\,a_{\max}/v_{\max} = 4.755$ when accelerating and $k_p = 10\,a_{\max}/|v_{\min}| = 19.02$ when braking. It is limited to $-a_{\max} \le \alpha_t \le \bar{a}(v_t)$,

$$\bar{a}(v) = \begin{cases} a_{\max}, & v \le v_{sw}, \\ a_{\max}\,v_{sw}/v, & v > v_{sw}, \end{cases} \tag{A6}$$

which models the motor's power limit above the switching speed v_{sw}. The followers do not use these servos: their DMPC solution is applied directly as $u_t^i(0)$ (App. IV).

III. LEADER MDP: ADDITIONAL DETAILS

The leader state is $s_t = (\xi_t^{\text{patch}}, a_t, b_t)$, where a_t and b_t are the current semi-axes of the elliptical funnel. The funnel is centred at the patch car's position $c_t = (x_t^{\text{patch}}, y_t^{\text{patch}})$ and oriented along its heading φ_t. We write $(a_t^{\text{cmd}}, b_t^{\text{cmd}})$ for the commanded semi-axes in u_t^{patch} and (a_t, b_t) for the realised ones; they differ because of the rate law (A12) and the wall cap (A11).

Two policy interfaces: The results use two versions of the leader interface (Table AV). The *absolute* interface ($o_t \in \mathbb{R}^{115}$) outputs the setpoints of Eq. (5) of the main text directly. The *rate* interface ($o_t \in \mathbb{R}^{122}$) outputs a steering increment that is integrated under a lateral-acceleration envelope, and a fraction of the locally available lateral room for b_t^{cmd}. This gives an $11\times$ smaller mean per-step steering change at higher speed. Both produce the same command tuple $u_t^{\text{patch}} = (\bar{\delta}_t, \bar{v}_t, a_t^{\text{cmd}}, b_t^{\text{cmd}})$, and everything downstream of u_t^{patch} (servos, rate law, followers) is shared.

A. Observation

$$o_t = \big[\,\tilde{\sigma}_t,\ e_{y,t},\ v_t,\ \psi_{\text{err},t},\ a_t,\ b_t,\ \kappa_{\mathcal{C}}(\sigma_t),\ (\chi_t),\ \ell_t\,\big], \tag{A7}$$

- $\tilde{\sigma}_t = \sigma_t/L \in [0,1]$: normalised progress.
- $e_{y,t}$: signed cross-track error.
- v_t: longitudinal speed.
- $\psi_{\text{err},t} = (\psi_t - \psi_{\mathcal{C}}(\sigma_t))_{\text{wrap}} \in (-\pi, \pi]$: heading error to the local tangent.
- $a_t,\ b_t$: current semi-axes.
- $\kappa_{\mathcal{C}}(\sigma_t)$: centerline curvature.
- $\ell_t \in [0, 30]^{108}$: every tenth beam of the 1080-beam, 4.7 rad field-of-view LiDAR (Gaussian range noise, $\sigma = 0.01$ m), clipped to 30 m.

The scan is ray-cast against the static occupancy map only: the patch car is simulated alone, so follower bodies never appear in it. The 7-D block $\chi_t = [\,\delta_t/\delta_{\max},\ \nu_{t-1},\ b_t^{\text{room}}/b_{\max},\ \dot{\psi}_t\,]$ appears only in the rate interface. It contains the steering state, the previous raw action $\nu_{t-1} \in [-1,1]^4$, the normalised lateral room of Eq. (A10), and the yaw rate. The two interfaces therefore have $7 + 108 = 115$ and $7 + 7 + 108 = 122$ dimensions. Observations are normalised by running mean and standard deviation (VecNormalize, clipped at ± 10); the statistics are frozen at deployment.

B. Action space and decoding

Absolute interface. The policy output is clipped to

$$\begin{aligned}\mathcal{A}^{\text{patch}} = [-0.4189, 0.4189]\,\text{rad} \times [2, 10]\,\text{m/s} \\ \times\,[1.5, 3.0]\,\text{m} \times [1.0, b_{\max}]\,\text{m},\end{aligned} \tag{A8}$$

giving $(\bar{\delta}_t, \bar{v}_t, a_t^{\text{cmd}}, b_t^{\text{cmd}})$ directly, with $b_{\max} = 3.0$ m, or 4.6 m for the clutter checkpoint.

Rate interface. The raw action is $\nu_t \in [-1,1]^4$. The steering setpoint is integrated under a lateral-acceleration envelope,

$$\begin{aligned}\bar{\delta}_t &= \mathrm{clip}\big(\bar{\delta}_{t-1} + 0.03\,\nu_{t,1},\ -\delta^{\text{env}}(v_t),\ \delta^{\text{env}}(v_t)\big), \\ \delta^{\text{env}}(v) &= \min\Big(\delta_{\max},\ \arctan\frac{\ell\,a_{\text{lat}}^{\max}}{v^2}\Big) \quad (v > 0.3\,\text{m/s}),\end{aligned} \tag{A9}$$

with $\ell = 0.33$ m and $a_{\text{lat}}^{\max} = 6$ m/s^2. The steering setpoint changes by at most 0.03 rad per tick, and δ^{env} is 0.42, 0.055 and 0.02 rad at 2, 6 and 10 m/s. The remaining components are decoded as $\bar{v}_t = \mathrm{clip}(6 + 4\nu_{t,2},\ 2,\ 10)$ and $a_t^{\text{cmd}} = 2.25 + 0.75\,\nu_{t,3}$. The semi-minor command is a fraction of the lateral room seen by the LiDAR,

$$b_t^{\text{cmd}} = b_{\min} + \tfrac{1}{2}(\nu_{t,4} + 1)\,\big(b_t^{\text{room}} - b_{\min}\big), \tag{A10}$$

where $b_t^{\text{room}} = \mathrm{clip}\,\big(0.85\,b_t^{\star},\ b_{\min},\ b_{\max}\big)$. Here $b_t^{\star}$ is the largest b for which an ellipse with semi-major axis a_t^{cmd}, centred on the car and shifted ahead by 0 and by $0.4\,v_t$ m, contains no LiDAR return. The bounds are $b_{\min} = 1.0$ m and $b_{\max} = 4.6$ m.

C. Funnel shape transition

The semi-axes are not part of the physical dynamics. Before integration, the commanded semi-minor axis is capped by the track half-width $w_h(\sigma_t)$ (App. III-D),

$$\begin{aligned} \bar{b}_t &= \max\big(\min(b_{\max},\ 0.85\,\rho_t),\ b_{\min}\big), \\ \rho_t &= \max\big(w_h(\sigma_t) - |e_{y,t}|,\ 0.05\big), \end{aligned} \tag{A11}$$

and both axes then follow a rate-limited tracking integrator,

$$\begin{aligned} a_{t+1} &= a_t + \mathrm{clip}\big(\min(a_t^{\mathrm{cmd}}, a_{\max}) - a_t,\ \pm\dot{a}_{\max}\Delta t\big), \\ b_{t+1} &= b_t + \mathrm{clip}\big(\min(b_t^{\mathrm{cmd}}, \bar{b}_t) - b_t,\ -\dot{b}^{-}_{\max}\Delta t,\ \dot{b}^{+}_{\max}\Delta t\big), \end{aligned} \tag{A12}$$

with $a_{\max} = 3.0\,\mathrm{m}$ and $\dot{a}_{\max} = \dot{b}^{-}_{\max} = 8\,\mathrm{m/s}$. Growth is limited to $\dot{b}^{+}_{\max} = 2\,\mathrm{m/s}$ in the rate interface and $8\,\mathrm{m/s}$ in the absolute one. The finite rates keep the safe region continuous between ticks and give the followers a bounded reaction time. Episodes start with $(a_0, b_0) = (2.0, 1.5)\,\mathrm{m}$, the car at the first centerline point and its heading aligned with the local tangent.

D. Track-width lookup

The half-width $w_h(\sigma)$ is precomputed at environment initialisation at $n_w = \max(200, \lceil L/0.15 \rceil)$ uniformly spaced arc-length samples $\{\sigma_j\}$, each the mean of a left and a right ray cast (step $0.1\,\mathrm{m}$, at most $5\,\mathrm{m}$ per side) on the occupancy map. At run time it is read from the next sample,

$$w_h(\sigma) \approx w_h(\sigma_{j^*}), \qquad j^* = \min\{\, j : \sigma_j \geq \sigma \,\}, \tag{A13}$$

found by binary search ($O(\log n_w)$) instead of a per-step ray cast. This more than triples the single-CPU rollout throughput. w_h is wall-based and is used only in the cap (A11) and in r_t^{shape}; obstacle avoidance comes from the LiDAR.

E. Reward

The per-step reward of Eq. (6) of the main text, $r_t^{\mathrm{patch}} = r_t^{\mathrm{prog}} - r_t^{\mathrm{track}} - r_t^{\mathrm{ctrl}} + r_t^{\mathrm{shape}} - r_t^{\mathrm{safe}}$, has the terms

$$\begin{aligned} r_t^{\mathrm{prog}} &= w_{\mathrm{prog}}\,\Delta\sigma_t + w_v\,v_t\max(\Delta\sigma_t, 0), \\ r_t^{\mathrm{track}} &= w_{e_y}\,|e_{y,t}|, \\ r_t^{\mathrm{ctrl}} &= w_{v\delta}\,v_t|\bar{\delta}_t| + w_{\dot{\delta}}\,\frac{|\bar{\delta}_t - \bar{\delta}_{t-1}|}{\Delta t} \\ &\quad + w_\omega \max\big(0, |\dot{\psi}_t| - \omega_{\mathrm{thr}}\big) \\ &\quad + w_{\mathrm{rev}}\max(-\Delta\sigma_t, 0) + w_{\mathrm{ut}}\,\mathbb{1}_{\mathrm{ut},t} + w_\tau\Delta t, \\ r_t^{\mathrm{shape}} &= w_{\mathrm{fill}}\ \mathrm{clip}\big(b_t/w_h(\sigma_t),\, 0,\, 1\big), \\ r_t^{\mathrm{safe}} &= C_{\mathrm{wall}}\mathbb{1}_{\mathrm{wall},t} + C_{\mathrm{stuck}}\mathbb{1}_{\mathrm{stuck},t}. \end{aligned} \tag{A14}$$

- $\Delta\sigma_t = \sigma_t - \sigma_{t-1}$ is the signed per-step progress.
- The speed bonus counts only forward progress.
- $\mathbb{1}_{\mathrm{ut},t}$ flags reversing at speed ($\Delta\sigma_t < 0$ and $v_t > 1\,\mathrm{m/s}$).
- w_τ is a constant time penalty that rewards finishing quickly.
- $\mathbb{1}_{\mathrm{wall},t}$ fires when at least 2 of 32 points sampled on the *ellipse boundary* lie in an occupied or off-map cell. The patch car's own body contact is not penalised separately, because the car lies inside the ellipse.
- $\mathbb{1}_{\mathrm{stuck},t}$ fires when $|\Delta\sigma_t| < 10^{-3}\,\mathrm{m}$ for 100 consecutive steps (1 s).

TABLE AII
LEADER REWARD WEIGHTS (PER STEP, $\Delta t = 0.01\,\mathrm{s}$).

Symbol	Rate	Absolute	Role
w_{prog}	40	40	arc-length progress
w_v	8.0	2.0	speed × forward progress
w_{e_y}	1.0	1.0	cross-track error
$w_{v\delta}$	0.1	0.3	speed × steering
$w_{\dot{\delta}}$	0.1	0	steering-setpoint rate
$w_\omega,\ \omega_{\mathrm{thr}}$	0.15, 7 rad/s	0.15, 7 rad/s	yaw-rate (anti-spin)
w_{rev}	120	120†	reverse progress
w_{ut}	25	25†	reversing at speed
w_τ	$15\,\mathrm{s}^{-1}$	$15\,\mathrm{s}^{-1}$	time penalty
w_{fill}	2.0	2.0	funnel fills the corridor
C_{wall}	1000	1000	ellipse touches a wall
C_{stuck}	10	10	no progress for 1 s

† Clutter checkpoint only; the two May checkpoints have no w_{rev}, w_{ut} terms.

- There is no terminal or lap bonus.

Table AII lists the weights. The yaw-rate term w_ω is the penalty added to suppress the spinning mode described in Sec. V-A.

F. Termination

An episode ends when the car arrives ($\sigma_t \geq L$), when the funnel touches a wall ($\mathbb{1}_{\mathrm{wall},t}$), when the stuck condition fires, or after 10^5 steps.

IV. Follower DMPC: Additional Details

Nothing on the follower side is learned. Every follower runs the same solver κ (Eq. (10) of the main text); only its information set $\mathcal{I}_t^i$ differs.

A. Information set

At tick t, follower i uses

$$\begin{aligned} \mathcal{I}_t^i = \Big\{ &\xi_t^i,\ i,\ \big(c_t, \varphi_t, v_t^{\mathrm{patch}}, \dot{\psi}_t^{\mathrm{patch}}, a_t, b_t\big), \\ &u_t^{\mathrm{patch}},\ \mathcal{C},\ \{\hat{q}_t^j(0{:}H)\}_{j\neq i} \Big\}, \end{aligned} \tag{A15}$$

- ξ_t^i: its own measured single-track state.
- i: its slot index.
- $(c_t, \varphi_t, v_t^{\mathrm{patch}}, \dot{\psi}_t^{\mathrm{patch}}, a_t, b_t)$: the current patch pose, speed, yaw rate and semi-axes.
- u_t^{patch}: the current patch command.
- $\mathcal{C}$: the known centerline, giving (σ, e_y), $\kappa_{\mathcal{C}}$ and w_h.
- $\{\hat{q}_t^j(0{:}H)\}_{j\neq i}$: the latest broadcast plans of the other followers.

It does not use the other followers' measured states, future patch commands, a map of obstacles, or its own LiDAR.

B. Funnel prediction

Holding u_t^{patch} constant over the horizon, each follower predicts the funnel centre $\hat{c}_t(k)$, heading $\hat{\varphi}_t(k)$, speed $\hat{v}_t(k)$ and semi-axes $(\hat{a}_t(k), \hat{b}_t(k))$ for $k = 0, \ldots, H$. It starts

from $\hat{c}_t(0) = c_t$, $\hat{\varphi}_t(0) = \varphi_t$, $\hat{v}_t(0) = \max(v_t^{\text{patch}}, 0.5)$, $(\hat{a}_t(0), \hat{b}_t(0)) = (a_t, b_t)$ and $\hat{\sigma}_t(0) = \sigma_t$, and propagates

$$\begin{aligned}
\hat{\omega}_t(k) &= w_k^{\text{cmd}}\Big[w_k^{\text{act}}\,\dot{\psi}_t^{\text{patch}} + (1 - w_k^{\text{act}})\,\frac{\hat{v}_t(k)\tan\bar{\delta}_t}{\ell}\Big] \\
&\quad + (1 - w_k^{\text{cmd}})\,\hat{v}_t(k)\,\kappa_{\mathcal{C}}\big(\hat{\sigma}_t(k)\big), \\
\hat{\varphi}_t(k{+}1) &= \hat{\varphi}_t(k) + \hat{\omega}_t(k)\,\Delta\tau, \\
\hat{v}_t(k{+}1) &= \text{clip}\Big(\hat{v}_t(k) + \Delta\hat{v}_t(k),\ 0.5,\ 10\Big), \\
\Delta\hat{v}_t(k) &= \text{clip}\,\big(\bar{v}_t - \hat{v}_t(k),\ \pm a_{\max}\Delta\tau\big), \\
\hat{c}_t(k{+}1) &= \hat{c}_t(k) + \hat{v}_t(k)\,\mathbf{e}\big(\hat{\varphi}_t(k)\big)\,\Delta\tau, \\
\hat{\sigma}_t(k{+}1) &= \hat{\sigma}_t(k) + \hat{v}_t(k)\,\Delta\tau,
\end{aligned} \tag{A16}$$

where $\mathbf{e}(\varphi) = [\cos\varphi,\ \sin\varphi]^\top$, $w_k^{\text{act}} = e^{-k\Delta\tau/0.05}$, $w_k^{\text{cmd}} = e^{-k\Delta\tau/0.3}$, and $\dot{\psi}_t^{\text{patch}}$ clipped to ± 2.5 rad/s. The predicted turn rate blends the measured yaw rate (first $\approx$0.05 s), the held steering command (first $\approx$0.3 s) and the centerline curvature (beyond). The semi-axes follow the same law as the environment, Eqs. (A11)–(A12), with Δt replaced by $\Delta\tau$, the cap evaluated at $\hat{\sigma}_t(k)$ and $e_{y,t}$ held. The predicted funnel velocity is $\hat{\mathbf{v}}_t(k) = \hat{v}_t(k)\,\mathbf{e}(\hat{\varphi}_t(k))$.

C. Slot assignment

Slots are offsets in the funnel frame, recomputed every tick from the current shape.

1) With $\tilde{a} = \max(a{-}0.15, 0.5)$ and $\tilde{b} = \max(b{-}0.15, 0.5)$, N points are packed inside the scaled ellipse $(\rho\tilde{a}, \rho\tilde{b})$ by 140 iterations of projected repulsion. The repulsion keeps each pair, and each point and the patch car at the origin, at least 0.82 m apart. Three deterministic seeds are tried.
2) $\rho \in [0.25, 1]$ is the smallest feasible scale (largest margin), found by 11 bisection steps.
3) The packing is rotated in 10° steps to place the formation as far astern as containment allows.
4) The offsets are scaled by $\text{clip}\,\big(r_{\max}/\max_i \|p_i\|, 0.35, 1\big)$, where $r_{\max} = (0.8 \cdot 12 - v_t^{\text{patch}})/|\dot{\psi}_t^{\text{patch}}|$. This shrinks the formation when turning at speed would demand more than the followers can achieve.
5) Slots are ordered by bearing.

Offsets are computed for the predicted shapes at $k = 0$ and $k = H$ and interpolated linearly in k, giving $q_t^{i,\text{slot}}(k) = \hat{c}_t(k) + R(\hat{\varphi}_t(k))\,p_i(k)$.

D. Optimal control problem

At every tick, follower i solves

$$\begin{aligned}
\min_{\{\xi_t^i(k)\},\{u_t^i(k)\},\{s_t^{ij}(k)\}} \ & \sum_{k=0}^{H}\Big(J_t^{i,\text{trk}}(k) + J_t^{i,\text{con}}(k) + J_t^{i,\text{int}}(k)\Big) \\
& + \sum_{k=0}^{H-1} J_t^{i,\text{sm}}(k)
\end{aligned} \tag{A17}$$

subject to

$$\xi_t^i(0) = \xi_t^i, \qquad \xi_t^i(k{+}1) = f_{\Delta\tau}\big(\xi_t^i(k), u_t^i(k)\big), \tag{A18}$$

$$\xi_t^i(k) \in \mathcal{X}, \qquad u_t^i(k) \in \mathcal{U}, \tag{A19}$$

$$n_t^{ij}(k)^\top q_t^i(k) \geq d_t^{ij}(k) - s_t^{ij}(k), \quad s_t^{ij}(k) \geq 0, \quad j \neq i, \tag{A20}$$

where $f_{\Delta\tau}$ is one RK4 step of Eq. (3) of the main text over $\Delta\tau$ and $q_t^i(k) = (x_t^i(k), y_t^i(k))$ is the planned position. The hard constraints are the dynamics and the box sets

$$\begin{aligned}
\mathcal{X} &= \{\xi : 0.5 \leq v \leq 12\ \text{m/s},\ |\delta| \leq \delta_{\max}\}, \\
\mathcal{U} &= \{u : |\alpha| \leq 9.5\ \text{m/s}^2,\ |\dot{\delta}| \leq \dot{\delta}_{\max}\}.
\end{aligned} \tag{A21}$$

Containment and separation are soft, through the costs and the slacks $s_t^{ij}(k)$, as stated in assumption (v) of the main text. Only $u_t^i(0)$ is applied, directly to the follower's single-track plant with no low-level servo, and the problem is re-solved at the next tick.

E. Separating planes

After each solve, follower j broadcasts the positions $\hat{q}_t^j(0{:}H)$ its plan visits. The patch car enters as neighbour $j = 0$ with the predicted centre $\hat{q}_t^0(k) = \hat{c}_t(k)$ as its plan. For every neighbour and every $k = 1, \ldots, H$, the plane normal points from the neighbour's plan to follower i's own latest plan,

$$n_t^{ij}(k) = \frac{\hat{q}_t^i(k) - \hat{q}_t^j(k)}{\|\hat{q}_t^i(k) - \hat{q}_t^j(k)\|}, \tag{A22}$$

falling back to the follower's travel direction rotated by 90° when the two plans coincide. The offset is

$$d_t^{ij}(k) = \begin{cases} n_t^{ij}(k)^\top \frac{1}{2}\big(\hat{q}_t^i(k) + \hat{q}_t^j(k)\big) + \frac{r}{2}, & j \geq 1, \\ n_t^{i0}(k)^\top \hat{q}_t^0(k) + r, & j = 0, \end{cases} \tag{A23}$$

with clearance $r = 0.8$ m. Between two followers, j receives the same plane with the opposite normal, so each concedes $r/2$. The patch car does not yield, so its plane is placed the full clearance from its path. No plane is imposed at $k = 0$.

F. Cost terms

Throughout, $\mathbf{v}_t^i(k) = v_t^i(k)[\cos(\psi_t^i(k) + \beta_t^i(k)),\ \sin(\psi_t^i(k) + \beta_t^i(k))]^\top$ is the planned velocity vector and $[\,\cdot\,]_+ = \max\{0, \cdot\}$.

Tracking. The follower moves with the funnel and holds its slot,

$$\begin{aligned}
J_t^{i,\text{trk}}(k) = w_v\,\big\|\mathbf{v}_t^i(k) - \hat{\mathbf{v}}_t(k)\big\|^2 \\
+ w_q\,\big\|q_t^i(k) - q_t^{i,\text{slot}}(k)\big\|^2.
\end{aligned} \tag{A24}$$

The velocity term matches a vector rather than a speed set point, so moving with the patch is the default and the slot term corrects only the residual offset.

Containment. Let $(x_r, y_r)^\top = R(\hat{\varphi}_t(k))^\top\big(q_t^i(k) - q_t^{i,\text{slot}}(k)\big)$, with semi-axes reduced by the margin m and

floored at 0.5 m, $\tilde{a} = \max(\hat{a}_t(k) - m, 0.5)$ and $\tilde{b} = \max(\hat{b}_t(k) - m, 0.5)$. Then

$$\begin{aligned} g_t^i(k) &= \left(\frac{x_r}{\tilde{a}}\right)^2 + \left(\frac{y_r}{\tilde{b}}\right)^2 - 1, \\ J_t^{i,\mathrm{con}}(k) &= w_c\Big(1 + 0.8\,\frac{k}{H}\Big)\,\phi\big([g_t^i(k)]_+\big), \\ \phi(h) &= 2g_0^2\Big(\sqrt{1 + (h/g_0)^2} - 1\Big). \end{aligned} \tag{A25}$$

ϕ is a pseudo-Huber hinge: quadratic for small excursions and linear for large ones, so a follower pushed far outside is pulled back firmly but with a bounded gradient. The predicted-shape ellipse is anchored at the follower's slot reference, so its size and orientation track the predicted funnel.

Interaction. The follower keeps clear of the patch car and of every other follower,

$$\begin{aligned} J_t^{i,\mathrm{int}}(k) = \sum_{j \neq i} \Big(& w_d\big[r^2 - \|q_t^i(k) - \hat{q}_t^j(k)\|^2\big]_+ \\ & + w_s\big(s_t^{ij}(k) + 10\,s_t^{ij}(k)^2\big)\Big). \end{aligned} \tag{A26}$$

The slack price w_s makes the plane nearly hard: it is crossed only when the alternative is no plan. The distance term is a smooth backup for pairs whose plans coincide and therefore have no well-defined separating direction.

Smoothness and effort.

$$\begin{aligned} J_t^{i,\mathrm{sm}}(k) = & \big(\alpha_t^i(k{+}1) - \alpha_t^i(k)\big)^2 + \big(\delta_t^i(k{+}1) - \delta_t^i(k)\big)^2 \\ & + 0.1\,\alpha_t^i(k)^2 + 0.1\,\delta_t^i(k)^2 + w_{\mathrm{lat}}\,\big(S_t^i(k)\big)^2, \end{aligned} \tag{A27}$$

where the last term softly bounds the lateral acceleration, $|v_t^i(k)\,\dot{\psi}_t^i(k)| \le a_{\mathrm{lat}}^{\max} + S_t^i(k)$, $S_t^i(k) \ge 0$. It is active for the rate-interface leader ($a_{\mathrm{lat}}^{\max} = 6\,\mathrm{m/s^2}$) and disabled otherwise.

G. Coordination and solver

- *Coordination.* Each tick runs $S = 2$ Gauss–Seidel sweeps. In each sweep every follower in turn re-solves against the latest plans of the others and broadcasts a damped update, $\hat{q}_t^i \leftarrow (1-\eta)\hat{q}_t^i + \eta\,\tilde{q}_t^i$ with $\eta = 0.75$.
- *Solver.* Problems are transcribed with CasADi and solved by IPOPT (at most 200 iterations, tolerance $2 \cdot 10^{-3}$, acceptable tolerance 10^{-2}).
- *Warm start.* Each solve is warm-started from the previous solution; a failed warm solve is retried once from a straight-line guess.
- *Failure fallback.* After three consecutive failures, the follower executes its last plan open-loop and then re-seeds.

Table AIII collects the parameters.

TABLE AIII
FOLLOWER DMPC PARAMETERS.

Symbol	Description	Value
$H,\ \Delta\tau$	horizon steps, step length	40, 0.025 s (1.0 s)
—	re-solve rate	every tick (100 Hz)
$w_v,\ w_q$	velocity, slot tracking	40, 90
$w_c,\ m,\ g_0$	containment weight, margin, hinge	1500, 0.15 m, 1.0
$w_d,\ r$	keep-out weight, clearance	200, 0.8 m
w_s	plane-slack price	6000
$w_{\mathrm{lat}},\ a_{\mathrm{lat}}^{\max}$	lat.-accel. slack, bound	2000, 6 m/s^2/off
$S,\ \eta$	Gauss–Seidel sweeps, damping	2, 0.75
$\mathcal{X}$	speed, steering bounds	$[0.5, 12]$ m/s, ± 0.4189 rad
$\mathcal{U}$	accel., steering-rate bounds	± 9.5 m/s^2, ± 3.2 rad/s
—	slot packing clearance	0.82 m

V. Leader Training

A. PPO

The leader is trained with PPO (Stable-Baselines3). Policy and value function are separate MLPs with two hidden layers of 256 units (tanh) and a state-independent Gaussian log-std. Rollouts are collected from E parallel environments, each reset with seed $42 + e$. Observations and rewards are normalised online (VecNormalize, clip 10 and 5), and the statistics are frozen at deployment, where the mean action is used. Table AIV lists the hyper-parameters.

TABLE AIV
PPO HYPER-PARAMETERS (ALL CHECKPOINTS).

Parameter	Value
learning rate	linear $3 \cdot 10^{-4} \to 10^{-5}$
rollout per update ($E \cdot n$)	65,536 steps
minibatch size	8192
epochs per update	10
discount γ, GAE λ	0.99, 0.95
clip ε	0.2
value coeff. c_1, entropy coeff. c_2	0.5, 10^{-3}
max. gradient norm	0.5
networks (policy / value)	[256, 256] / [256, 256], tanh

B. Checkpoints and training maps

A leader policy depends only on the map family it was trained for, not on N. The results use the best checkpoint per map family (Table AV). Training maps are sampled uniformly from a pool, so repeated entries act as weights. The vehicle parameters of Table AI are fixed; there is no plant or LiDAR randomisation beyond the simulator's range noise.

- *Open corridors (rate interface).* 450 entries over 303 distinct maps, each kept for 4 consecutive episodes:
  - the 3-block corridor of the benchmark ($\times 60$);
  - a second 3-block layout ($\times 40$);
  - the obstacle-free corridor ($\times 50$);
  - 200 variants with 2–7 random blocks (rectangles, triangles, discs);
  - 100 variants with 4–8 wall-attached blocks.
- *Escalating neck (rate interface).* 510 entries over 404 maps:
  - the benchmark map ($\times 80$);
  - three denser versions with 26, 33 and 40 blocks ($\times 10$ each);
  - 200 variants with 8–15 centre blocks;
  - 200 heavy variants with 20–38 blocks.
- *Clutter (absolute interface).* 320 maps with 18–41 wall-attached and centre obstacles, resampled every episode.

TABLE AV
LEADER CHECKPOINT USED PER BENCHMARK MAP.

Benchmark map	Interface	$\dim o_t$	E	Steps
open_narrow (obs1–3)	rate	122	13	40.0M
neck_escalate_easy (obs1–7, tail)	rate	122	13	40.0M
clutter	absolute	115	12	28.5M
neck_escalate (obs1)	absolute	115	16	17.2M
slalom_longbars45	absolute	115	16	10.0M

- *Narrow neck and slalom (absolute interface).* Trained on a single corridor each, without map randomisation.

VI. BENCHMARK SCENES

The 14 scenes are cut from five corridor maps whose wall-to-wall width varies between 3.2 and 11.1 m (Table AVI).

- `open_narrow` (118.9 m): a corridor narrowing from 9.27 m to a 3.19 m section ($\sigma \in [47.2, 97.5]$ m), with three large blocks.
- `clutter` (360 m, 25 obstacles): the scene is the first group of three blocks, leaving a 3.05 m free passage.
- `neck_escalate_easy` (360 m): seven single 2.3×2.55 m blocks that alternately close the port and starboard side and leave a 3.2–4.1 m lane, followed by a tail of eight 1.2×1.3 m blocks spaced $\approx$2 m apart.
- `neck_escalate`: the same layout with larger 3.4×3.8 m blocks, leaving a 3.1 m lane; one scene.
- `slalom_longbars45` (189 m): 14 bars at 45° protruding alternately from the two walls about 12 m apart, with 4.4–4.5 m gates; one scene covering all gates.

Each scene spans one feature: an obstacle group (blocks closer than 10.5 m along the track form one group), or the whole slalom.

- *Spawn line:* 7 m before the feature start.
- *Finish line:* 4.5 m after the feature end, clipped to 0.5 m before the map end.
- *Moved finish lines:* where the standard finish lies within 2 m of the point at which the patch car stops at the end of its course, the finish is moved to $\max(\sigma_{\text{end}}, \min(\sigma_{\text{end}} + 1, \sigma_{\text{stop}}^{\text{patch}} - 0.5))$ for every method. This applies to `open_narrow`:obs3 and `neck_escalate_easy`:tail.

VII. EVALUATION PROTOCOL AND METRICS

A. Protocol

- *Team size.* $N \in \{3, 4, 5, 6\}$ is the number of followers; the patch car is an additional vehicle.
- *Our method.* The followers start at the slots of the initial funnel. For the three absolute-interface checkpoints, they are re-placed at the slots of the current funnel when the patch car reaches the scene's spawn line, and scoring starts there. The rate-interface checkpoints drive their whole map once, and each scene is scored on its segment of the trace.
- *Baselines.* They spawn in a wedge (rows 1.2 m apart, ± 0.62 m lateral) at the spawn line, with a 5 m/s target speed and a planned lane through the scene. A car that makes contact stops, and the others continue.
- *Sensing.* Every method receives the centerline and its own LiDAR scan. DEFORM, which has no LiDAR-only mode, receives the map as a point cloud.
- *Spawn jitter.* Each trial perturbs every car's spawn pose by $\mathcal{U}(\pm 0.15)$ m along and across its heading and $\mathcal{U}(\pm 3°)$ in heading.

TABLE AVI
THE 14 BENCHMARK SCENES. ARC LENGTHS IN M; WIDTHS ARE WALL-TO-WALL OVER THE SCENE; "FREE" IS THE NARROWEST OBSTACLE-FREE PASSAGE.

Scene	Feature σ	Length	Width	Free	Obstacles
ON obs1	19.5–23.3	15.3	4.9–9.3	4.9	1 block
ON obs2	100.8–104.2	14.9	3.2–9.3	6.1	1 block
ON obs3	113.3–116.9	10.9	5.2–9.3	5.2	1 block
CL group1	27.3–40.9	25.0	3.7–9.3	3.1	3 blocks
NEE obs1	39.9–42.4	14.1	3.3–9.3	4.0	1 block
NEE obs2	102.1–104.7	14.1	3.3–9.3	3.6	1 block
NEE obs3	157.3–159.9	14.1	4.0–9.3	3.6	1 block
NEE obs4	221.6–224.1	14.1	3.3–9.3	3.9	1 block
NEE obs5	239.7–242.2	14.1	9.3	3.3	1 block
NEE obs6	257.7–260.3	14.1	9.2–9.4	3.5	1 block
NEE obs7	275.8–278.3	14.1	4.1–9.3	4.1	1 block
NEE tail	339.6–354.0	22.4	3.2–9.3	3.8	8 blocks
NE obs1	39.3–43.1	15.3	3.3–9.3	3.6	1 block
SL slalom	15.6–176.4	172.3	4.5–11.1	4.5	14 bars

ON = open_narrow, CL = clutter, NEE = neck_escalate_easy, NE = neck_escalate, SL = slalom_longbars45.

B. Metrics

A car is in *contact* when its clearance to a wall or obstacle falls below 0.15 m, or its centre comes within 0.35 m of another car's centre (two 0.31 m-wide cars side by side). The patch car counts as a car. Contacts are counted only within the scene, from 0.5 m before the spawn line to 0.5 m after the finish, on poses sampled every 0.05 s.

- *Completion rate* C_r: the share of agents that cross 95 % of the spawn-to-finish distance without contact, pooled over scenes. Collisions and deadlocks (timeouts) count as failures.
- *Average velocity* $\hat{V}$: for our method, the patch car's mean speed between spawn and finish; for the baselines, the mean speed of all agents inside the scene.
- $V_{\min}$: the lowest speed reached inside the feature, masked from 0.2 s before a team's first contact. Speeds are finite differences smoothed over 0.2 s.
- T_{clear}: the time from the first agent crossing the spawn line to the last agent crossing the finish, averaged over runs in which every agent crossed.

VIII. BASELINES AND ABLATIONS

All baselines run in the same F1TENTH simulator with the vehicle of Table AI, except DEFORM on TurtleBot3, and are scored by the same code.

ORCA [1]:

- *Collision avoidance.* RVO2 with agent radius 0.33 m, 2 s agent and obstacle time horizons, and a 12 m neighbour range (at most 10 neighbours).
- *Preferred velocity.* Points to the lane with magnitude $5/(1+6\,\kappa_{\max})$, where $\kappa_{\max}$ is the maximum lane curvature over the next 20 waypoints.
- *Vehicle control.* The resulting velocity is tracked by pure pursuit (1.5 m look-ahead) at 100 Hz.
- *Sensing.* Obstacles are rebuilt from the cars' scans, and the lane is re-planned every 0.25 s over a 14 m horizon.

GCBF+ [2]: The authors' pre-trained DubinsCar checkpoint, not retrained. The scene is scaled so that a car maps to the training agent's footprint (factor 0.172), and walls and obstacles are given as rectangles. Each agent's goal is a carrot one waypoint ahead on the lane. The output (ω, a) is converted to steering and speed commands, with speed limited to 8 m/s.

DEFORM [3]: The authors' code at its published commit, with fixes needed to run it: velocity caps exposed as parameters, the A* pool size, the grid origin, and the robot-height type. It is run in two ways:

- *F1TENTH bridge.* Velocity caps 3.0 m/s and 2.5 rad/s, 300 s timeout; the map is given as a point cloud.
- *Native Gazebo.* TurtleBot3 Burger with its stock 0.22 m/s cap and DEFORM's own perception, 900 s timeout, scored with TurtleBot3-sized contact thresholds (0.12 m wall, 0.25 m pair).

Multi-agent RL (IPPO): Followers share one PPO policy (parameter sharing), trained with the frozen leader. Each follower is rewarded only for staying inside the funnel. Training converged for $N = 1$ and plateaued at $N = 2$; $N \geq 3$ was not trained.

DMPC (no funnel): Each car runs a lane-following NMPC with the follower coordination of App. IV, but with no patch and no containment cost. Its components are:

- *Model and rate.* Single-track model, 1.0 s horizon at 20 Hz.
- *Tracking.* The lane reference and a 5 m/s speed reference.
- *Corridor.* Slacked corridor half-planes.
- *Collision avoidance.* A slacked 0.66 m keep-out and a soft 0.9 m keep-out against neighbours.
- *Coordination.* Plan exchange with separating planes at $r = 0.8$ m, slack price 6000, $S = 2$, $\eta = 0.75$.

FASTFUNNELS + *fixed patch:* The full method with the funnel shape frozen at $(a, b) = (2.4, 1.6)$ m, the mean adaptive shape. The leader's speed and steering remain adaptive, and the wall cap (A11) and the rate law still apply.

Leader–follower (fixed formation):

- *Leader.* Follows the lane by pure pursuit at 5 m/s (1.5 m look-ahead).
- *Followers.* Hold a fixed wedge behind the leader with PD gap keeping (desired gap 1.5 m, gains 1.5 and 0.4, 1.2 m look-ahead).
- *Slowdown.* The team slows to 2 m/s within 8 m of obstacles and narrow sections.

IX. HARDWARE DEPLOYMENT

Vehicle:

- *Platform.* Traxxas UDR, 1/7 scale: wheelbase 0.42 m ($l_f = l_r = 0.21$ m), 6.1 kg, 0.345×0.694 m body.
- *Actuation.* VESC drive, steering lock ± 0.418 rad (minimum turning radius $\approx$0.95 m).
- *Sensing.* 2D LiDAR at 40 Hz (1081 beams).
- *Simulated plant.* Uses the measured mass, wheelbase and body size, with I_z, h and the cornering stiffnesses scaled from the F1TENTH values.

Leader retraining:

- *Map.* The leader is retrained with the UDR plant on the SLAM map of the experiment corridor.
- *Map pool.* 350 maps: 250 variants with the obstacles jittered by up to ± 0.3 m (about half driven in the reverse direction), and 100 copies of the scanned map in both directions, with and without added clutter.
- *Domain randomisation:*
  - steering lock $\mathcal{U}[0.36, 0.418]$ rad and steering rate $\mathcal{U}[2.0, 3.2]$ rad/s;
  - per-episode LiDAR noise $\sigma \sim \mathcal{U}[0, 0.03]$ m;
  - pose noise 0.02 m / 0.02 rad per step;
  - collision checks on the car body inflated by 0.10 m.
- *Commands and observation.* Speed is capped at 2.0 m/s, the funnel semi-axes are scaled to the car ($a \in [0.45, 1.5]$ m, $b \in [0.27, 1.2]$ m), and the LiDAR is clipped at 6 m. The policy uses the absolute interface (115-D).

Onboard stack:

- *Leader node.* The policy is exported to NumPy and runs in a single ROS 2 node on the car, from the LiDAR scan and VESC odometry.
- *Localisation.* The pose comes from a scan matcher against the SLAM map (distance-field search, 1.4 cm mean error in simulation), fused with odometry.
- *Latency.* An actuator latency of 80 ms is compensated.
- *Speed command.* A minimum speed command overcomes the VESC dead-band.
- *Followers.* Each follower runs the DMPC of App. IV in its own node (CasADi/IPOPT), receives the published funnel state, and exchanges plans with the other followers.

REFERENCES

[1] J. van den Berg, S. Guy, M. Lin, and D. Manocha, "Reciprocal n-body collision avoidance," 2011.

[2] S. Zhang, O. So, K. Garg, and C. Fan, "GCBF+: A neural graph control barrier function framework for distributed safe multiagent control," *IEEE T-RO*, 2025.

[3] J. Li, Y. Xu, X. Shi, and L. Li, "DEFORM: Adaptive formation reconfiguration of multi-robot systems in confined environments," *IEEE RA-L*, 2025.