

Pour wine into a glass — environment dossier

GAR environment wine-pouring (Newton 1.5 + MuJoCo rigid world, a GPU DFSPH liquid written for this package, Isaac RTX rendering). Action group: docs/research/pour_wine_glass (25 actions, 23 robot and 14 physics capabilities). Research and design: artifacts/creations/wine-pouring-devin-20261001/research.md, cited as **R§n**; its measurement scripts M1-M5 are in research/measurements/. Every number below comes from a script in this package, and every evidence file is listed with its SHA-256 in evidence/MANIFEST.json. The typeset version of this document is report/wine_pouring_dossier.pdf (python -m wine_pouring.report_pdf); its figures are drawn only from evidence JSON.

What is ASSISTED: a scripted operator (IK joint targets streamed to the robot's physical servos, with privileged inventory feedback), or a phantom fixture (a spring-damper wrench on one body, or a force ramp). Both are labelled in every video and are not reachable from the policy interface. Nothing in this package solves a task autonomously.

Contents

1. The real job
2. Scene at real scale
3. The liquid
4. Liquid in moving containers: coupling
5. The bottle's headspace
6. Robots, controllers, grasp and protective stops
7. Rendering and cameras
8. Parameter ledger
9. Tasks and evaluation
10. Physics validation against theory and measurement
11. Assisted demonstrations
12. Stability at rest and under legal abuse
13. Runtime
14. Interfaces
15. Failures found and corrected during the build
16. Limitations
17. Reproduction and evidence custody
18. Sources

1. The real job

The group (docs/research/pour_wine_glass; R§1) is gravity transfer of a liquid from a held vessel into a receiving one, ending at a target quantity or level: wine from a bottle into a glass (rows 008, 022-024), water or another beverage from a pitcher or carafe into glasses and cups (004, 010-012, 016, 017, 021), plus dispensing by lever (001, 005, 009), condiments and cream (013-015), hot liquids (002, 003), a mixed drink (007) and five granular rows (006, 018-020, 025). What makes the job physical rather than a lookup (R§2.6-2.7):

- **The pour window is narrow and depends on how full the bottle is.** The reconstructed LEONOR starts to pour with its axis 16.5° above horizontal when full (750 mL), 9.5° at 658 mL, 7.75° at 600 mL, 4.25° at 450 mL (section 10.5): an operator finds the onset by tilting, not at a fixed angle.
- **A sealed headspace throttles the outflow.** Once wine fills the neck, the air above it is cut off; letting 1 mL out of a full bottle's 19 mL headspace drops its pressure by $\gamma p \Delta V / V = 7.5 \text{ kPa}$, three times the column's head (R M3). A real bottle held mouth-down empties at 85-95 g/s (Kenton, Neufeld & Huppert 2012 [S33]), not at the ~300 g/s of a vented orifice.
- **The payload changes while it is held.** A full 75 cl bottle weighs 1.25 kg; it gets 150 g lighter per glass and its centre of mass moves; the Franka Hand's 70 N holds it by the body, not by the neck (R M5).
- **Level is judged by eye.** 150 mL fills the PURE Bordeaux glass to 152 mm above its foot — just under the widest part of the bowl, the sommelier's fill [S1, S3]; the ISO tasting glass's 50 mL sits at 80 mm; the UK lined goblet's 175 mL line is at 133 mm [S5].
- **Spills and wet bottles.** Wine left running down the bottle or spilt on the bar is a service failure [S1, S4].

Robotic bars do this today with two arms behind a counter and illuminated delivery spots (Makr Shagr Toni [S20], Richtech ADAM [S21]); this environment is that setting.

Coverage of the group (built tasks, section 9):

rows	activity	built
008, 022, 023, 024	wine from an opened bottle into a glass, to a desired amount	pour-glass, taste-pour, split-bottle, refill-glass, tasting-flight, lined-measure
004, 010, 011, 012, 016, 017	water / beverage from a pitcher or carafe into glasses	water-tumbler, decant-carafe
021	pour to a fill line	lined-measure (fired 175 mL line)
007	drink presented in a serving glass; carrying it	carry-glass
(service practice)	dumping a tasting sample before the next wine	dump-and-pour
001, 005, 009	lever dispenser	not built (section 16)
002, 003	hot liquid	not built separately (density/viscosity only differ)
013-015	condiments, cream	not built
006, 018-020, 025	granular	not built

2. Scene at real scale



Figure 1. service/pour-glass seed 0, room camera (diagnostic). Rendered by probes/render_views.py.

2.1 Physical geometry (collision, mass, friction)

World frame: x along the bar to the robot's right, y from the robot toward the guests, z up; SI units throughout.

object	geometry	mass	source
bar top	3.00 × 0.66 × 0.03 m at 1.090 m (finished height 42–45 in, 26 in deep)	static	[S9, S10]
bar die, back panel	walnut boxes under the top	static	authored
robot pedestal	steel column, top plate at 1.00 m; arm bases at (±0.25, −0.18) m facing +y	static	authored
LEONOR 75 cl bottle	310 mm, base Ø72, shoulder Ø77, EN 12726 bore 18.5 mm, fill 63 mm below top; 750.0 mL at the fill line, 768.9 mL brimful	498 g	[S11, S15]
PURE Bordeaux glass	680 mL, bowl Ø94, 267 mm	210 g	[S16, S17]
ISO 3591 tasting glass	155 mm, bowl Ø65, rim Ø46, 215 mL	94.5 g (computed)	[S6–S8]
UK lined goblet	175 mL fired line at 132.7 mm, 381 mL brimful	computed	[S5] (line from the interior)
1 L carafe, 500 mL carafe	100/80 mm body, 50/43 mm neck	computed	estimate
tumbler	80 mm rim, 66 mm base, 105 mm	computed	estimate
dump bucket	stainless, 120 mm rim, 150 mm	300 g	estimate
bottle coaster	stainless dish Ø105, 15 mm rim	150 g	estimate

Each vessel is one closed (r, z) polygon of its wall (geometry.py). From that one polygon come its volumes (exact Green integrals; e.g. LEONOR fill line 750.0 mL), its mass properties (glass density scaled to the datasheet mass), the liquid's walls (signed distance and density map, section 3.2), the rigid collision pieces (solid frusta where nothing enters; hollow bands split into convex sectors where a finger or neck can enter) and the render mesh (section 7). Contact friction (params.MU): every glass contact — glass on glass, on the quartz aggregate of the bar top, on the stainless coaster — 0.4, and pad on glass 0.8. MuJoCo combines two equal-priority geoms by the maximum, so these are the values its contacts actually use (read back from `mj_data.contact.friction`; the pads carry geom priority 1 and impose theirs). Basis: Engineering ToolBox's friction table gives glass–glass 0.9–1.0 static and 0.4 sliding when clean and dry, 0.1–0.6 static when greased, glass–metal 0.5–0.7 clean and 0.2–0.3 greased [S41]; a bar in service is wiped and often wet, and MuJoCo has one coefficient, so 0.4 (the dry sliding value, inside the wet static range). The pad value is barcode_scanning's qualified Franka Hand pad (rubber on smooth dry surfaces 0.5–0.85 in the same table); a wet pad (rubber–wet concrete 0.45–0.75) is not modelled. The first ledger listed glass–quartz 0.35 and bottle–coaster 0.3, values MuJoCo never applied (qualitative audit, round 1).

2.2 Visual dressing (no physics)

A wine bar built around the working bar: walnut die and footrail, four bar stools (Poly Haven bar_chair_round_01, 0.75 m seat, 0.48 × 0.49 m) 0.65 m apart centre to centre (bar seating is laid out at 0.6–0.75 m per stool; the first layout put them 0.34 m apart and neighbours interpenetrated in every station-camera frame), a back bar with a walnut counter and two shelves of this package's own bottles (green and flint glass, labels, wine inside) and inverted Bordeaux glasses, a wine barrel, plant and picture, brick back wall, plank floor (CC0 Poly Haven assets pinned by SHA-256 in visual_assets.json; the brick texture's long axis runs along the wall — the first build stood the bricks on end). The bar top is a generated quartz-speckle texture with a satin (honed) finish, roughness 0.30: polished at 0.15, the centre pendant's bulb reflected as a saturated highlight in the middle of the policy camera, where the bottle and coaster stand. Three industrial pendants hang 1.0 m above the bar (3500 K sphere lights sized analytically for 500 lx on the bar top — EN 12464-1's food-preparation value; its bar value is 300 lx [S40]), a back-bar wash and a weak ambient dome. Dressing never collides; everything a robot or the liquid touches is built in scene.py.

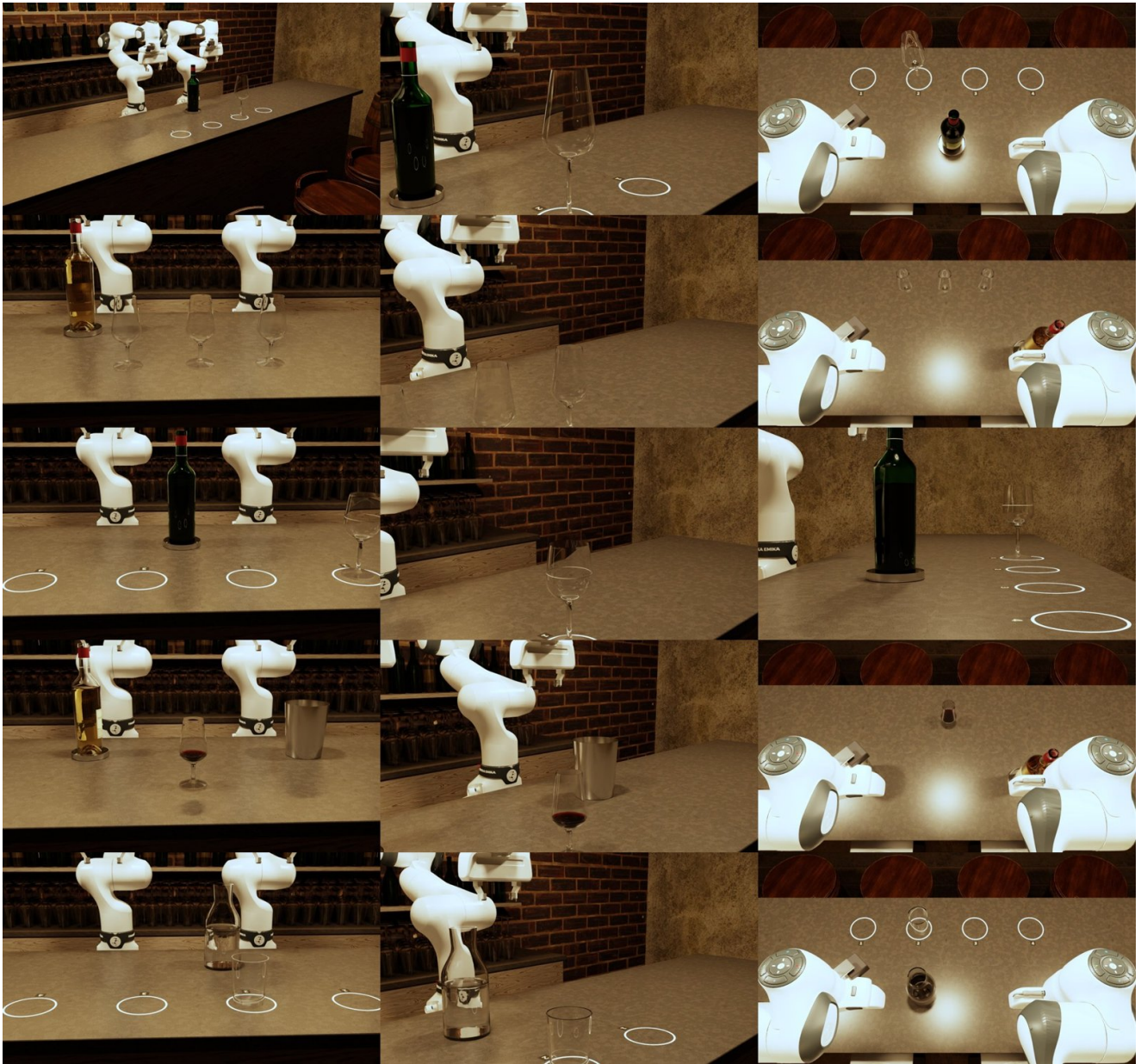


Figure 2. Rows: service/pour-glass (room, closeup, station_rgb), tasting/tasting-flight with white wine (guest, closeup, station_rgb), lined/lined-measure (guest, closeup, side), tasting/dump-and-pour (guest, closeup, station_rgb), service/water-tumbler with the 1 L carafe (guest, closeup, station_rgb). Delivery spots are lit in the presets whose tasks name them.

3. The liquid

3.1 Divergence-free SPH (liquid.py)

Particles of volume $V = dx^3$ ($dx = 2$ mm, $8\ \mu\text{L}$ each; a full bottle is 93,750 particles) carry a liquid id (red, white, rosé wine, water: density, viscosity, optical absorbance in `params.LIQUIDS`). Volume-fraction formulation $s_i = \sum_j V W_{ij} + \Phi_{\text{wall}}(x_i)$, cubic spline with support $2\ dx$ normalised so that a dx lattice sums to exactly 1. Each substep (CFL $0.4\ dx$, 2–24 substeps per 1 ms tick):

1. density s_i and DFSPH factor $\alpha_i = 1 / \max(|\sum_j V \nabla W_{ij} + \nabla \Phi|^2 + \sum_j |V \nabla W_{ij}|^2, \frac{1}{4} \alpha_{\text{lattice}}^{-1})$;
2. divergence projection (velocity field divergence-free);
3. non-pressure forces: gravity; Morris viscosity with each liquid's ν ; wall viscosity and no-slip shear;
4. density projection (predicted density error);
5. per-substep speed cap 8 m/s (section 3.4), then the velocity-level wall guard: no particle centre may end the substep closer than $0.3\ dx$ to a solid — container walls, bar top, floor, finger pads and the robot's links and hand (3.2). Each wall is tested over the distance the particle can travel in the substep; whatever approach velocity would cross is removed exactly; a particle already too close is pushed out at ≤ 2 m/s; the reaction goes to the body. A particle found inside a container wall is moved back out (position repair, counted);
6. symplectic Euler; XSPH velocity filter $\epsilon = 0.02$.

The first guard tested walls only within $4\ \text{m/s} \times \Delta t$ and added at most 2 m/s per substep, and the cap ran after the integration: a full bottle dropped 5 cm onto its coaster drove ~ 120 particles through its 9.6 mm base and up to 112 sat inside the glass at once (evidence audit, round 1). With the swept test, exact removal and the cap first, the same drop puts none there (section 12.2).

Both projections are projected relaxed Jacobi iterations (Bender & Koschier 2017 with signed increments): $\kappa_i \leftarrow \max(\kappa_i + \omega (s_i^* - 1) \alpha_i / \Delta t^2, 0)$, $\omega = 0.5$, velocity update $\Delta v_i = -\Delta t \sum_j V (\Delta \kappa_i + \Delta \kappa_j) \nabla W_{ij} - \Delta t \Delta \kappa_i \nabla \Phi_i$, warm-started from the previous substep's total, to a mean density error $\leq 1 \cdot 10^{-4}$ and a maximum $\leq 1 \cdot 10^{-2}$ (divergence $\leq 1 \cdot 10^{-3}$). The total stays ≥ 0 (free-surface condition) while increments are signed, so an overshooting warm start is taken back.

3.2 Walls: exact-geometry density maps

A container wall acts on the liquid through its density map (Koschier & Bender 2017): $\Phi(r, z)$ is the kernel integral over the part of a particle's support hidden behind the wall — inside the glass, or reached only through it — computed once per container on a 0.5 mm grid in its own (r, z) frame with line-of-sight occlusion (`geometry.density_map`, cached) and scaled so a flat wall at the resting distance $dx/2$ gives the value of the lattice it replaces. Pressure on a near-wall particle is $-\kappa_i \nabla \Phi$; its reaction, the viscous and the guard reactions go to the container as a wrench. Thin walls work from either side (wine inside a glass, a glass standing in spilt wine). The bar top (above its 30 mm slab) and the floor (unbounded) are planes with the exact planar table. The robot is a boundary too: the finger pads as oriented boxes, and every link and the hand as a convex polytope circumscribing its collision hull (the hull's own face normals merged within 14° , completed with fixed directions, each plane at the hull's support distance — tight on faces, a few millimetres loose on rounded parts; `world.robot_hulls`). Liquid that falls on the hand or forearm lands on it and runs off, with the reaction on the link (section 12.2). The first build guarded the fingers only: the links' axis-aligned boxes were centimetres loose and reached through a gripped bottle's glass into its wine, and without them liquid fell straight through the arm and hand in free fall (qualitative audit, round 1).

3.3 Initial conditions

A fill places particles on wall-conforming rings dx apart (a cubic lattice cut by a curved wall leaves a $dx/2$ gap: a 750 mL fill did not fit the 769 mL bottle), inner rings $\geq 0.75\ dx$ from the centre particle. Preparation: 40 steps of zero-gravity packing relaxation (velocities zeroed), hydrostatic pressure $\kappa = g \cdot \text{depth}$, then 150 ms of damped dynamics (speeds $\times 0.9$ per substep) so that the pressure field, not the velocity zeroing, carries the weight. Measured in a full bottle: after a quasi-static settle the first free step pressed 1.6 N on the 6.3 N load and the column rang ± 5 N; after the damped settle it pressed 6.28 N and stayed below 17 mm/s rms. Section 10.1 measures the result in every container.

3.4 Surface tension, wetting and spills; numerical limiters

On by default since the repair round (`params.SURFACE_FORCES`). Three terms, each with its basis and its measured discrete behaviour (V4, `probes/v_surface.py`, `evidence/v_surface/surface.json`; V2b, section 10.3):

- **Cohesion** (surface tension): the Akinci, Akinci & Teschner (2013) pair force $F_{ij} = -\gamma m_i m_j C(r) \hat{r}$ with the coefficient normalised so the pair potential's planar surface energy equals each liquid's σ , $\gamma = 8\sigma / (\pi \rho_0^2 \int_0^h C(s) s^4 ds)$ (titration's `cohesion2013.py` derives the same), between pairs that include a surface-layer particle (volume fraction $s < 0.90$; $s = 1$ in the bulk). Applied everywhere, the bulk pair forces — which cancel on average — gave the particle lattice a shear stiffness at its rest spacing, which equals the kernel's attraction/repulsion crossover $h/2$: the tumbler's slosh frequency rose 9 % above the cylinder mode (section

10.3). Akinci's curvature term is not used: with the σ -normalised γ it is dimensionally inconsistent (one mass factor fewer than the cohesion term), and with it meniscus particles were thrown out of a moving bottle's neck at 3–7 m/s.

- **Adhesion** (wetting): each wall acts as a half-space of virtual liquid behind its surface, weighted by Young's law $(1 + \cos \theta)/2$ — the cohesion pair force integrated over the half-space, a continuous function of the distance to the wall that attracts beyond $\sim h/2$ and repels closer. Near a container's lip it is limited by **Duez's inertial-capillary separation** (rounds 2–3, `params.DUEZ_SEPARATION`, `liquid.k_nonpressure`, `liquid.lip_zone`):
- Within 6 mm of the lip's rounded crest, adhesion cannot hold liquid sliding along the glass faster than $U_c = \sqrt{(We_c \sigma / (\rho e_o))}$, with $We_c = K (r_i + e_o/2) r_i / e_o^2 (1 + \cos \theta)$ (Duez et al. 2010 Eq. 2; $K = 0.44$ read from their Fig. 3b inset).
- r_i is the lip's own edge radius: the LEONOR finish's 1.5 mm round, half the 0.4–0.6 mm rim of the stemware, and 1.25–1.5 mm for the tumbler and carafe.
- e_o is the local liquid thickness normal to the glass: the particle's distance to the wall, plus how far its neighbours reach further out, plus half a particle, never less than one spacing.
- Beyond U_c the attraction falls as $(U_c/u)^2$. This roll-off is authored: the inertial load on the turning film grows as pu^2 .
- At 2 mm, U_c is about 0.04 m/s for a thick stream leaving the LEONOR's lip, 0.14 m/s for a one-particle trickle, and 0.04 m/s at a glass's thin rim.

Round 2's form set $r_i = e_o = dx$ on every glassware wall. So U_c moved with the spacing (0.156 / 0.174 / 0.195 m/s at 2.5 / 2.0 / 1.6 mm), and the closure also acted on flat walls, outside Duez's edge mechanism (qualitative audit, round 3). Without any closure, the kernel-range adhesion held a steady 25–45 mL/s stream round the bottle's lip and back along its neck, which Duez's measurements say a real lip does not. A slow trickle and a resting film keep their full adhesion, and so do flat walls, so the onset dribble and the final drops remain (section 10.7).

- **Contact-line friction** on the bar top and the floor: a particle touching the plane at the edge of the liquid ($s < 0.95$) has its sliding velocity relaxed at 100 s^{-1} , a closure for contact-angle hysteresis, which is what pins a real drop. Calibrated on the puddle law: a puddle stands $h = 2 l_c \sin(\theta/2)$ high ($l_c = \sqrt{\sigma/\rho g} = 2.2 \text{ mm}$), 2.0 mm at 55° , so 2 mL rests on about 9.9 cm² (the round-1 text said 1.9 mm and 10.5 cm²: an arithmetic slip the qualitative audit found); at 100 s^{-1} a 2 mL block released on the bar comes to rest on 14.3 cm² (round 2: 14.9; round 1: 14.6), 45 % more than the puddle law's 9.9 cm², the miss of the calibration (wetted radius 21.3 mm, fastest particle 0.003 mm/s after 4 s; V4); in calibration runs 30 s^{-1} gave 105 cm² and 10 s^{-1} was still spreading over 2400 cm² after 4 s. Without any of the three it spread to 1.16 m² and was still growing at 42 mm/s after 4 s. Not applied on glassware: liquid runs down a bottle.

V4 (evidence/v_surface/surface.json), each case with the forces off and on:

case	off	on
2 mL block on the bar, after 4 s	1.16 m ² , still spreading at 42 mm/s, 244 of 250 particles moving	14.3 cm ² , at rest (0 particles above 5 mm/s)
150 mL at rest in the PURE glass, 1 s	load = weight, level −0.63 mm, mean speed 3.9 mm/s	load = weight, level −0.80 mm, mean 4.0 mm/s
full bottle lifted 10 cm in 0.5 s, carried 20 cm in 0.6 s, stopped	0 particles out	0 particles out (the first surface model let six out at 3–7 m/s)

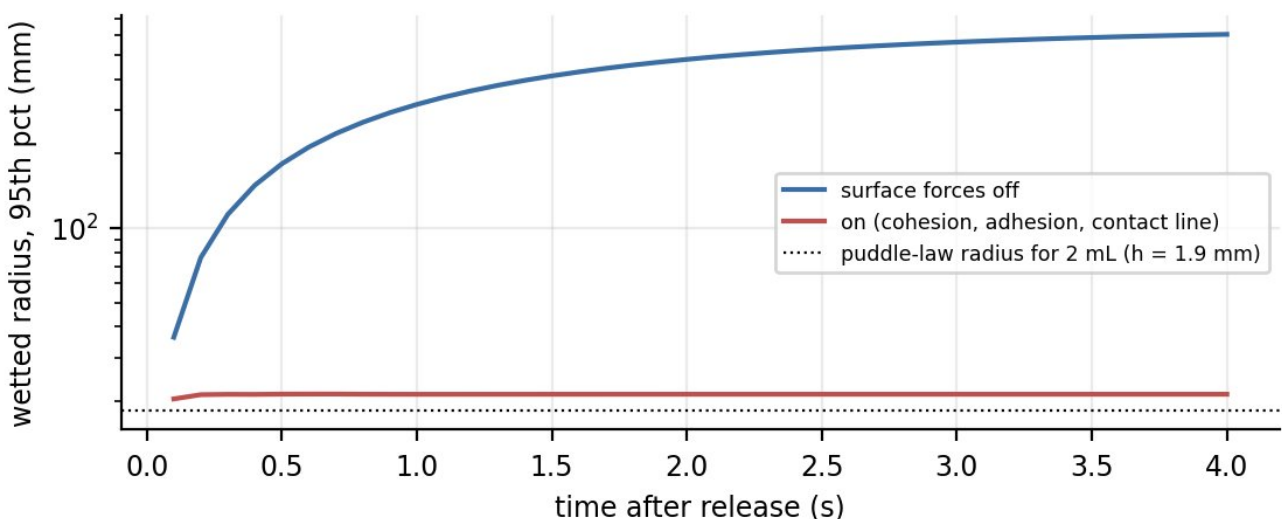


Figure 3b. V4: wetted radius of a 2 mL spill on the bar against time, with and without surface forces; dotted: the puddle-law radius.

What this buys and costs at 2 mm. A spill now stays where it lands, as drops and a patch near the glass; the stream holds together longer before it breaks into drops; wine wets the glass, so a slow stream follows the outside of the lip and the neck at a trickle (the teapot effect: E9, section 10.7). The cost: the tumbler's slosh frequency is 3 % above the cylinder mode instead of 1 % (section 10.3), and a puddle's 2.0 mm height is one particle, so a spill is one to two particles thick.

Numerical limiters and closures, each a measured fix (section 15). How often each acted is written into every evidence file (`Liquid.limiter_totals`: particle-substeps and ticks where the speed cap engaged, particles repaired out of a wall, squeezed particles, particles moved with a container's correction, impact-impulse ticks):

- XSPH 0.02 and the DFSPH factor floor (two particles pushed almost onto each other have vanishing kernel gradients; one was left at $3 \cdot 10^5$ m/s in a step).
- Squirt cap: 8 m/s per substep, applied before the guard (wine caught between a falling rim and the bar top was shot out at 25 m/s; free fall from 3 m is 7.7 m/s).
- **Squeeze closure** (rounds 2–3, `params.SQUEEZE_GAP_RATIO`, `liquid.k_guard`). A particle between two facing solids (normals within 60° of opposite) that are closer together than 1.25 dx cannot be represented in that gap.
 - It is placed on the gap's midline and carried with the gap's normal motion.
 - Along the gap, away from the moving body, it moves at the Reynolds squeeze-out speed $r V/(2h)$, with r its distance from the moving body's axis, V the closing speed and h the gap ($\geq \text{dx}/4$), bounded to 0.2–0.5 m/s.
 - It loads neither solid and takes no part in the pressure solve.

A real film under a glass set on a wet bar drains in milliseconds (Reynolds squeeze: 1 mm under a 42 mm foot at 3.5 N in ~ 7 ms). It is decelerated by its own lubrication pressure, $3\mu VR^2/h^3$, and leaves as a thin sheet that surface tension holds to the bar; a 2 mm particle cannot be that sheet, hence the bound.

- Without any closure (round 1), the guard pushed each trapped particle out of the bar at up to 2 m/s per substep, and the vessel took that momentum: a carafe set on 10 mL was thrown 197–420 mm and once the run blew up at 2.6 million m/s (evidence audit, round 2).
- With round 2's closure the vessel stayed put, but the momentum went into the liquid. Squeezed particles kept any faster tangential speed and stayed in the pressure solve, where the two walls' overlapping density maps made them look compressed. A glass set down at 5 cm/s on 5 mL threw the puddle at the 8 m/s cap, 0.94 m high and 4.9 m across the room (evidence audit, round 3; V7b, section 12.2b).
- **Position correction inside the clearance** (round 3, `liquid.guard_dv`). A particle found closer to a wall than the guard's clearance has its approach stopped and is moved out to the clearance; no velocity is added. Rounds 1–2 pushed it out at up to 2 m/s per substep. In the wedge where a glass's foot closes on the bar, the bar's push and the foot's alternately pumped particles to 3 m/s and threw them 0.35 m up.
- **Planes hidden behind glass** (round 2, `k_density`). A plane (bar top, floor, room wall) does not act on a particle whose nearest container wall is nearer and faces the same way: the particle is inside a glass whose wall lies between it and the plane. Before this, wine in a glass lying on the bar felt the bar top through the 1 mm wall as extra density.
- **The bar top is a box** (round 2, `wall_hit`): its front, back and end faces are walls too. As a plane limited to its footprint, liquid draping over the front edge just below the top surface counted as up to 30 mm inside the bar.
- **Room walls** (round 2, `params.ROOM`, `scene.liquid_planes`): the room's four walls bound the liquid. Under random legal actions some liquid had left the room through the back wall (evidence audit, round 2).
- Host residual read every 4th iteration (performance only).

4. Liquid in moving containers: coupling

A 0.5 kg bottle holding 0.75 kg of wine is the added-mass instability of partitioned fluid-structure coupling (Causin, Gerbeau & Nobile 2005): lifted by a phantom spring with explicit coupling it blew up within 25 ms. `coupling.py` lumps, every rigid step, the liquid inside each container (particle sums over its interior raster) into the container's MuJoCo body — mass at the liquid's centre of mass and the transverse part of its inertia tensor from second moments — and hands the rigid solver only the residual

$F_{\text{extra}} = F_{\text{liquid}} + m_l (a_{\text{lag}} - g)$ $\tau_{\text{extra}} = \tau_{\text{liquid}} + (c_l - c) \times m_l (a_{\text{lag}} - g) + I_l \alpha_{\text{lag}}$ (about the combined centre of mass c)

low-pass filtered over 8 ms (the per-step DFSPH wall force of a resting full bottle fluctuates ± 0.18 N about its 7.30 N mean, section 10.1; acting raw on the lumped body it once kicked the bottle 3 m/s). For liquid moving with its container the correction cancels F_{liquid} ; sloshing and outflow forces remain in F_{extra} . Two details found necessary: the compiled inertial frame (`body_iquat`) is kept and the tensor's diagonal in that frame is written (rewriting the frame produced 126 kN contacts in one step); and only the transverse inertia is lumped — an inviscid liquid does not spin with its container about the container's axis, and lumping that moment (0.8 of a bottle's own) with its lagged correction spun a phantom-held bottle to 5900 rad/s. The liquid's own motion is unaffected: it always feels the walls.

Impacts. A bottle landing on its coaster stops from 1 m/s in about 4 ms (25 g). The iterative pressure solve propagates such an impulse through a 25 cm column only a few particle layers per iteration, and the liquid's walls had moved over the impact tick at the velocity they had at its start, ending up to $v \Delta t = 1$ mm away from where their liquid was carried: the column compressed and rebounded and its top was thrown out of the mouth at the 8 m/s cap — 14–29 mL and a jet 2–3.5 m high from a 5 cm drop (evidence audit, round 1). For liquid with a level free surface the exact incompressible response to a vertical velocity change Δv of its container is that uniform change (the pressure impulse $P = \rho \Delta v (h - z)$ vanishes on the surface and meets every wall's normal velocity). The coupling therefore hands the vertical proper acceleration of each container beyond 2 g to its contained liquid directly, with the container's vertical position correction for that tick and the equal and opposite momentum on the container (`LumpedLiquid._impact_impulse`). With it the same drop lets nothing out (section 12.2). The impulse is exact only for a container whose free surface is level over a floor that stops it, so since round 2 it applies only to a container within 30° of upright (`params.IMPACT_MAX_TILT_DEG`). Round 1's report said handling stays below 2 g. That was wrong: the impulse acted 9–23 times in every robot pour (at the set-downs) and 525 times while random legal actions swung the bottle about (evidence audit, round 2). Those 525 had handed tilted liquid a uniform vertical kick.

Wall following (round 2, `LumpedLiquid._follow_walls`). The liquid steps against walls extrapolated over the tick with the velocities the bodies had at its start; the rigid step then moves them with the velocities after contact. At an impact the two differ by $\Delta v \Delta t$; for a glass landing on its bowl with its spin reversed from -9 to $+3.6$ rad/s, that is 2.5 mm. So liquid in a container, or within a kernel radius of its wall, follows the container's correction (the rigid map from its predicted pose to its actual pose) when the correction exceeds 0.2 mm at the container's extent. Ordinary handling never reaches that ($a \Delta t^2/2 \sim 10^{-5}$ m); impacts do, and the count is in the limiter totals.

5. The bottle's headspace

gas.py. Per rigid step one GPU pass counts the particles in each bottle and in 2 mm slabs along its neck. Walking inward from the mouth, the first slab holding $\geq 68\%$ of a full slab's particles seals the bottle if air (> 1 mL beyond the packing deficit) sits at the interior's highest point on the inner side of that slab. Sealed, the trapped air expands adiabatically (pressure changes over 8–21 ms, research M3, $\ll$ thermal diffusion ~ 5 s):

$$p = p_{\text{atm}} (n / V_g)^{1.4}, V_g = V_{\text{interior}} - V_{\text{liquid}} \text{ (8 ms-filtered count)}, p_{\text{atm}} - p \leq \rho g h_{\text{column}}$$

The pressure difference enters as the pressure drop it causes along the neck liquid: a uniform acceleration $(p - p_{\text{atm}}) / (\rho L)$ along the bottle axis on the liquid between the neck's bottom and the mouth (total $(p - p_{\text{atm}}) \times \text{neck area}$), reaction to the bottle. (Applied to the free-surface layer instead, it needed tension the free-surface condition forbids and surface particles tore off at 250 m/s.) Air re-enters only as bubbles through the mouth, not resolved: under suction the trapped air grows at the bubble-limited rate $Q_{\text{air}} = \beta A_{\text{mouth}} \sqrt{gD}$; a slab that is no longer full vents the bottle. Two constants were revised in the repair round against Kenton's measured bottle (section 10.6): the suction is bounded by the column it holds, $\rho g h$ with h the liquid's height above the mouth (air is drawn in through the mouth beyond that — the plug), instead of the authored $1.5 \rho g L_{\text{bottle}}$ that let the suction exceed the column and set the flux at 40° ; and $\beta = 0.80$, calibrated to Kenton's 80° rate (with Clanet & Searby's ideal-bottle 0.45 [S29] the 80° flux was exactly $0.45 A \sqrt{gD} = 51$ mL/s, 0.55 of the measurement). The 40° case is the check.

6. Robots, controllers, grasp and protective stops

6.1 Configurations (robots/catalog.py, entry `wine-pouring`)

alias	rig	instances
dual-fr3 (default)	robots/rigs/pour-fr3.json	two FR3 + Franka Hand on left_arm_mount / right_arm_mount at (± 0.25 , -0.18 , 1.00) m facing the guests; station camera on a ceiling post, D405-envelope wrist cameras
single-fr3	pour-fr3-single.json	the left arm only
dual-panda	pour-panda.json	two Panda arms + Franka Hand
single-panda	pour-panda-single.json	the left Panda only (added in round 2: the infrastructure audit found a single-Panda setup refused with no physical reason; same reach qualification as single-fr3, section 12.4)
g1	refused with its reason	see 6.4

All rigs use the joint-servo and hand profile qualified in barcode_scanning: PD position servos, hard joint stops, Franka Hand grasp at 50 mm/s until blocked then the commanded force, MuJoCo no-slip. The servos' references come from the arm's motion generator (section 6.2b). The force convention: Franka publishes "grasping force continuous 70 N, max 140 N" [S36] and libfranka's Gripper::grasp(width, speed, force) takes "grasping force in N" without saying per jaw or total; Franka's own simulated gripper (franka_ros franka_gazebo/src/franka_gripper_sim.cpp, HOLDING state) applies force_desired / 2.0 to each finger joint [S42], i.e. the commanded grasp force is the total of both jaws. This package does the same: 70 N commanded is 35 N per finger, the normal force at each pad. (The research's M5 arithmetic assumed 70 N per finger; its holding margins are therefore halved — the dry side grasp 145 mm up a full bottle still holds, section 11.) Deposit_cans and barcode_scanning use the same convention. probes/robots_smoke.py builds each alias, holds the ready pose 0.5 s (joint drift $\leq 2 \cdot 10^{-6}$ rad) and moves joint 1 by the commanded 0.1 rad (section 12.4).

6.2 Protective stops

The FR3 collision reflex compares an external-torque estimate with libfranka's setDefaultBehavior thresholds (10 N m nominal, 20 N m while accelerating) and stops the arm; it then refuses arm targets for 1 s. Two changes from the barcode profile were needed and measured here: the estimate is low-pass filtered (20 ms; libfranka compares its filtered tau_ext_hat_filtered), and the torques of the payload a hand holds are subtracted — a body touched by both fingers is the payload; its weight and inertial force $M(g - a)$ at its centre of mass map to the joints through the hand's Jacobian, as an application calls franka::Robot::setLoad for a bottle it picks up. Unfiltered and payload-blind, the reflex read 22–50 N m while the arm tilted a full bottle and stopped every pour. With the bottle hanging free, the payload model reproduced the measured contact torques on all seven joints. A velocity reflex stops an arm whose joint exceeds its datasheet speed.

6.2b Motion limits: what the Franka controller allows (round 3; motion.py, V8

evidence/v_motion/)

The qualitative audit (round 3) found that the simulated FR3 had no acceleration limit. Its servo reference slewed toward each target at the datasheet joint speed, so:

- one policy cartesian_delta (3 cm in 0.2 s) accelerated the joints at 49 rad/s^2 , and a named_joint_target at 27 rad/s^2 ;
- the scripted operator's 50 ms waypoints reached 71 rad/s^2 at the joints and 107 rad/s^2 at the hand;
- the FR3's control interface allows 10 rad/s^2 per joint, $5,000 \text{ rad/s}^3$ jerk, and 9 m/s^2 , 17 rad/s^2 at the TCP (FCI "Robot and interface specifications"; libfranka rate_limiting.h).

Those accelerations turned a held bottle 2.8° in the pinch in one staging move.

Now every arm command (policy action, keyboard chord, scripted waypoint) sets a goal and a duration. The goal is reached by a linear ramp from the previous goal over the command's duration. A time-optimal tracker moves the servo reference toward the ramp under the model's FCI limits:

- FR3: 10 rad/s^2 and $5,000 \text{ rad/s}^3$ on every joint;
- FER/Panda: $7.5\text{--}20 \text{ rad/s}^2$ and $3,750\text{--}10,000 \text{ rad/s}^3$ per joint;
- the datasheet joint speeds.

The tracker brakes on $v = \sqrt{2 a e}$, so it stops at a goal without overshoot ($1.6 \cdot 10^{-6}$ rad in a 0.35 rad move). A Cartesian check scales an arm's joint accelerations so its TCP stays within the FCI translation and rotation velocity and acceleration limits. A protective stop holds the arm where it is.

A command shorter than its motion needs is followed as far as the limits allow, and the motion continues into the next command, as on the real arm. A 0.35 rad rest-to-rest move takes 0.38 s, not E10's 0.15 s.

V8 re-runs the audit's own scripts on this code (probes/motion_checks/, adopted unchanged):

motion (evidence/v_motion/)	round 2 (audit, unlimited servo)	round 3
one policy cartesian_delta, 3 cm in 0.2 s, through NewtonWinePouringBackend.act (action_profile.json)	49 rad/s ² peak joint acceleration, 53 of 200 ticks over 10	5.0 rad/s ² , none over 10; 16.8 mm covered in the 0.2 s, the rest after it
one policy named_joint_target, joint 1 +0.2 rad over 200 ticks (action_profile.json)	27 rad/s ² , 147 ticks over 10	6.9 rad/s ² , none over 10
the pour's staging tilt, 45° in 1.0 s as 20 waypoints (one per 50 ms), bottle weighted with 658 mL (grip_stage.json, coarse)	71 rad/s ² joints, 107 rad/s ² hand; the bottle turned 2.82° in the pinch and ended 4.4° short of 45°	5.1 rad/s ² joints, 6.6 rad/s ² hand, 0.8 m/s ² hand; the bottle turned 0.007° and reached 45.15°
the same path as 1,000 waypoints (smooth)	4.2 rad/s ² , 0.01°	4.3 rad/s ² , 0.010°
the pour's upright carry, 30 / 1,500 waypoints (carry_coarse / carry_smooth)	—	2.7 / 7.8 rad/s ² joints, 2.2 / 8.6 rad/s ² hand; 0.018 / 0.056°
staging, then the 12°/s tilt-on to 86° (coarse_then_ramp)	+0.02° during the ramp	0.36° in the pinch throughout the ramp
upright to 86° at 12°/s, rigid world, empty / full bottle (grip_tilt_rigid.json)	≤ 0.31°	≤ 1.70° / 1.70°; grasp point slides ≤ 2.7 mm
the same with the environment's 658 mL of liquid, to 78° (grip_tilt_liquid.json)	≤ 0.28°	≤ 0.37°; coupling residual ≤ 0.021 N m; the lower finger at its 40.0 mm stop, the upper at 34.5 mm (independent fingers, section 6.3)

In the robot pour the bottle's turn relative to the hand peaks at 2.8° and ends at 0.85° (section 11.1; round 2: 7.2–7.4°, 3.8° permanent).

Not done: libfranka refuses or rate-limits a command *stream*; here a command is always accepted and followed as far as the limits allow. The TCP jerk limits (4,500 m/s³, 8,500 rad/s³) are not checked separately; the joint jerk limit bounds them in practice.

6.3 Grasping a bottle with two pads

Two parallel pads hold a bottle's tilt through a friction couple about the axis perpendicular to their contact line; a rotation about the contact line itself is resisted only by the pads' torsional friction, which is much weaker.

How much torsion the simulated pinch resists (V6, probes/v_pad_torsion.py, evidence/v_pad_torsion/). Newton's MuJoCo solver gives every geom condim 3 unless a shape sets otherwise, and condim 3 ignores the torsional coefficient (qualitative audit, round 2). Any torsion the pinch resisted came only from multiccd's separate contact points along each pad, and nobody had measured it. V6 holds each vessel in the scripted 70 N grasp, lifts it, and ramps a phantom torque at 0.25 N m/s about the pad line or about the approach axis. The slip torque is the torque at which the vessel has turned 2° in the hand:

grasp	about the pad line, condim 3 (rounds 0-1)	about the pad line, condim 4, $\mu_t = 3.5$ mm (round 3)	about the approach axis (friction couple), condim 4
bottle, 145 mm up the 75 mm body	0.040 N m	0.285 N m	> 2 N m (no slip by the end of the ramp)
1 L carafe by its 58 mm neck	0.055 N m	0.265 N m	> 2 N m
PURE glass by its 8 mm stem	0.025 N m	0.120 N m	0.335 N m

(Round 3 re-ran V6 under the motion limits. Round 2's condim-4 pinch, $\mu_t = 4.2$ mm, gave 0.335 / 0.405 / 0.140 N m.)

A flat pad pressed on a cylinder touches along a line of the pad's extent along the vessel axis. For the collision pad the scene builds (the rubber tip box, 17.5 × 15.2 × 18.5 mm), that line is $L = 17.5$ mm. Round 2 used the real pad's 21 mm height, 20 % more than the pad in the scene (qualitative audit, round 3).

Spinning about the contact normal with uniform pressure, the pad resists $\mu N L / 4$: $0.8 \times 35 \text{ N} \times 0.0175 \text{ m} / 4 = 0.12 \text{ N m}$ per pad, so 0.245 N m for the pinch. (A disc-shaped patch of radius a would resist $(2/3) \mu N a$.)

Condlim 3 ignored it. The pads carry condim 4 with torsional friction $\mu_t = \mu L / 4 = 3.5$ mm (params.PAD_TORSIONAL_M; an estimate from the pad geometry). The measured 0.285 N m on the bottle is that 0.245 N m plus the 0.040 N m that multiccd's contact spread gives.

The two fingers are independent force-capped drives (qualitative audit, round 3). The real Franka Hand drives both fingers from one motor through a rack; the URDF's mimic joint expresses that, and the importer drops it (barcode_scanning REPORT 3.3). Here each finger pushes with at most half the grasp force. With the pinch axis vertical (the bottle near horizontal), the bottle's weight presses on the lower finger, which yields to its 40 mm stop: the bottle shifts 2.4–2.6 mm along the pinch axis (one finger at 40.0 mm, the other at 35.1 mm, in the audit's rigid and liquid tilts). The grip still holds; on the real hand the coupling would keep the bottle centred.

The demonstrations therefore grasp from the bottle's side (approach along $\pm x$), close the pads front and back, and tilt about the approach axis toward the guests ($+y$), 145 mm up the body — above the bottle-plus-wine centre of mass (110–120 mm), so the bottle hangs stably.

The bottle's turn in the hand was slip, not compliance (qualitative audit, round 3). Rounds 0–2 reported that the bottle "lags the hand by about 7° at full tilt" and called it pad compliance:

- the operator's correction was 6.9–7.5° in every round-2 pour;
- the round-0 pours turned 8.2° and 6.7° in the fingers, with 3.5–4.5° left after the tilt-back;
- in a tasting flight a bottle turned 12° ahead of the hand between glasses, and half the second pour went onto the bar.

The audit showed it was slip, set off by the arm's unlimited accelerations (section 6.2b):

- in round 2's shipped pour the grasp point slid 8.3 mm and 3.8° of the turn remained after the bottle was upright again;
- it began in the staging move (1.0° → 4.7° while the hand turned at 28–41°/s);
- the same grasp, tilted smoothly, held within 0.3°.

Under the motion limits, the operator's correction in round 3's pour stays within -0.1 to 1.0° and the bottle turns at most 2.8° in the hand (section 11.1); V8 measures 0.007° for the staging move. A full carafe held by its neck still turns 30° and more under its own torque (section 11.5): that is the pinch's torsional capacity, not the arm's motion. The scripted operator keeps its correction from the bottle's measured tilt and mouth position (privileged pose) as a safeguard.

Two grasp layouts failed. Approaching along $+y$ and tilting toward the glass folded the FR3's wrist onto link 5 (1.1 kN self-contact) and the velocity reflex stopped the arm. Gripping at 105 mm let the bottle rotate in the fingers and the stream missed. Over a short glass (ISO 155 mm, tumbler 105 mm) a bottle tilted 45° with its mouth 6 cm above the rim put its heel into the bar top; the staging tilt and mouth height are chosen to keep the heel 3 cm clear and to stay 8° short of the geometric onset for the current fill (demos/robot_pour.staging).

6.4 Why the G1 humanoid is refused

Forward kinematics of the pinned G1 URDF (g1_29dof_with_hand_rev_1_0, pelvis at its 0.79 m standing height, as in barcode_scanning) puts the shoulder pitch joint at 1.08 m — the height of this bar top (1.09 m) — and the palm 0.42 m from the shoulder (upper arm 0.193 m, forearm and wrist 0.184 m, palm 0.042 m). Pouring into the Bordeaux glass at a delivery spot needs the bottle's mouth at 1.42 m, 0.34 m above the shoulder, and 0.6 m in front of a shoulder standing at the bar's robot-side edge: outside the arm's reach before any question of grasp. Beyond reach, the Dex3-1 hand has no qualified grasp of a 75 mm glass cylinder at 1.25 kg (filleting refuses the G1 hand at ~ 500 g) and the robot no qualified balance controller for a payload that loses 150 g per glass. A G1 configuration would need a different, lower work surface (a restaurant table at 0.75 m), i.e. a different scene, which was not built. The refusal is explicit in the catalog (gar inspect wine-pouring) and the backend (section 12.4).

Decision (round 3): no humanoid configuration for this environment. The infrastructure audits of rounds 1–3 asked either for a table-service preset with a G1 rig or for this decision to be recorded. The action group does include table-service pours where a standing server is the natural embodiment: rows 002 (coffee poured into a cup on a patron's table), 004 (ice water into a glass on the table), 010 and 011 (cups and glasses on a table). A G1 version would need:

- a new scene: a 0.75 m restaurant table, place settings, and the G1 standing at it;
- a grasp of the vessels it would pour from, qualified like the Franka Hand's (V6). The repository's only standing G1 rig, scan-g1 (barcode_scanning), holds a 214 g scanner by finger contact. The lightest vessel here, the 1 L carafe, weighs 1.34 kg full; the bottle weighs 1.25 kg. No repository evidence qualifies a Dex3-1 grasp above ~ 0.5 kg (filleting refuses it at that);
- a balance check for a standing robot whose payload changes by 150 g per glass.

That is a different environment, not a configuration of this one. The four Franka configurations cover the bar service the group's other rows describe, and the G1 stays refused with its reason.

7. Rendering and cameras

Isaac Sim RTX path tracing in a render-only worker (barcode_scanning lifecycle with a parent watch; the worker never runs physics). Since round 3 a live session keeps its worker for the session's life. Round 2 released it after 60 s idle, and a policy that waited longer paid a new worker inside its next command: 11 s, 1,096× wall/sim for a 10 ms command (infrastructure audit, round 3). `renderer_options={"idle_release_s": s}` restores a timed release. Each request carries the body poses and one liquid surface mesh per liquid.

- **Liquid surface** (`surface.py`): particle centres smoothed toward their neighbourhood mean (Yu & Turk 2013, $\lambda = 0.9$ over 2 dx), the Zhu & Bridson (2005) implicit surface $\phi(x) = |x - \bar{x}(x)| - 0.55 \text{ dx}$ (support 2 dx) on a 1 mm grid, GPU marching cubes, one grid per spatial cluster (a droplet thrown across the room must not coarsen the glass's grid), two Taubin passes and area-weighted vertex normals. Unsmoothed, a stream rendered as a string of 2 mm beads.
- **Optics** (`WinePouring.mdl`): glass and liquids are dielectrics with Beer-Lambert volume absorption. Red wine's decadic absorbance per cm ($R \leftarrow A620 \text{ } 0.8$, $G \leftarrow A520 \text{ } 3.6$, $B \leftarrow A420 \text{ } 3.3$; Glories values [S26, S28]) becomes the natural-log coefficient $A \cdot \ln 10 \cdot 100$ per metre, so a thin stream is light red and a full bowl near black-red; white and rosé likewise; bottle glass green (passes ~15/45/12 % through 3 mm, estimate) or flint; crystal stemware IOR 1.52.
- **Glassware** is drawn from the same polygon as its physics, with crease-aware normals (smoothing the rim corner bent its normals into the wall and drew a dark band round every glass rim).
- **Robots** use their visual meshes in their model's per-part colours (white shells, graphite joint rings and hand, status-light lens).

Cameras:

- policy: `station_rgb` (ceiling post above the robot, 75° HFOV, 1280×720), `left_wrist_rgb` and `right_wrist_rgb` (D405 envelope on link 7, 87°), and since round 3 `bar_side_rgb` (fixed, at bar level on the guest side, 0.15 m above the bar top, looking across the four delivery spots through the glasses, 62°);
- privileged: room, guest, side, closeup and probe cameras.

A session's policy cameras are `station_rgb` alone unless chosen at launch: `gar run wine-pouring --backend-options-json '{"cameras": ["station_rgb", "bar_side_rgb"]}'` (Python: `make(..., cameras=...)`). Resolution and samples per pixel are `renderer_options={"width": 1280, "height": 720, "samples": 24}`, and the shortest policy command scales with all three and with the number of cameras (section 13).

Which camera shows the fill level. From the ceiling, `station_rgb` sees the bottle and the hand over the glass during a pour but not the glass's level, and the left wrist camera sees the bottle's label and neck (`views/policy_cameras_pour.mp4`, the round-3 evidence audit). `bar_side_rgb` is the research's "side view through the bowls": it sees each glass's level from beside it, with the poured stream above. The scripted demonstrations stop on the privileged inventory; a policy would judge the level from `bar_side_rgb`.

All cameras are ideal pinhole RGB without a noise model, exposed as a photographer would for the bar (ISO 400, f/2.8, 1/60 s, filmic tone map) with a fixed white balance that corrects the 3500 K light half way (linear-sRGB gains 0.77 / 1.02 / 1.51 from the blackbody's chromaticity), as a camera's auto white balance does under warm light; uncorrected, every frame had a sepia cast.



Figure 3. `probes/surface_look.py`: the environment's own 150 mL fill plus a jittered stream, rendered from the closeup camera. The FR3's per-part colours and the crease-shaded glass rim are visible.

8. Parameter ledger

params.py holds every constant with its provenance label (datasheet, standard, literature-read, search, derived, estimate, authored, measured, calibrated); params.ledger() is written into evidence files.

quantity	value	provenance
red wine ρ , μ	992 kg/m ³ , 1.6 mPa·s	Hlaváč et al. 2016 [S22] (literature-read); AJEV 2019 Shiraz 1.49–1.70 mPa·s at 20 °C [S23] (search)
white wine μ	1.75 mPa·s (10 °C)	AJEV 2019 Chardonnay 1.45–1.53 mPa·s at 20 °C [S23], scaled (estimate)
water ρ , μ , σ	998.2 kg/m ³ , 1.002 mPa·s, 72.7 mN/m	IAPWS via the repository ledger
wine σ , contact angle on glass	47 mN/m, 20°	estimate (ethanol–water [S24, S25])
contact angle on the bar top / floor	55°	estimate (silica 20–30°, polyester binder ~70°)
cohesion: surface layer, scale	$s < 0.90$; $1.0 \times$ the σ -normalised γ	a numerical modelling choice that changes the physics, chosen against V2b (slosh frequency) and V4 (spill)
inertial–capillary separation (Duez), at lips	$U_c = \sqrt{(K(1 + \cos \theta) \sigma / \rho \cdot (r_i + e_0/2) r_i / e_0^3)}$, $K = 0.44$; within half the lip's wall + 6 mm of its crest; beyond U_c adhesion $\propto (U_c/u)^2$	K: Duez et al. 2010 [S43] Fig. 3b inset (read from the fitted line); r_i : each container's lip (geometry; LEONOR 1.5 mm round, stemware 0.2–0.3 mm, tumbler/carafe 1.25–1.5 mm); e_0 : local liquid thickness (computed); zone width and the $(U_c/u)^2$ roll-off: authored
contact-line friction (bar, floor)	100 s ^{−1}	calibrated on the puddle law (V4, section 3.4)
impact impulse threshold, tilt limit	2 g vertical proper acceleration; containers within 30° of upright	authored (an impact ~25 g; the impulse is exact only for a level surface over a floor; section 4)
wall following	correction > 0.2 mm at the container's extent	authored (ordinary handling ~10 ^{−5} m per tick; section 4)
squeeze closure	gap < 1.25 dx; escape 0.5 m/s	authored (a gap one particle cannot fill; Reynolds squeeze-out ~0.2 m/s at a 42 mm foot; section 3.4)
room walls (liquid bound)	$x \pm 3.4$ m, $y -1.7 / 3.6$ m	the dressing's room (section 2.2)
finger pad torsional friction (condim 4)	$\mu_t = 3.5$ mm	estimate from the collision pad's 17.5 mm extent along the vessel axis: $\mu L / 4$ (V6, section 6.3; round 2 used the real pad's 21 mm)
FR3 / FER motion limits	FR3: 10 rad/s ² , 5,000 rad/s ³ per joint; TCP 3.0 m/s, 9 m/s ² , 2.5 rad/s, 17 rad/s ² . FER: 7.5–20 rad/s ² , 3,750–10,000 rad/s ³ ; TCP 1.7 m/s, 13 m/s ² , 2.5 rad/s, 25 rad/s ²	FCI robot specifications [S45] (section 6.2b)
squeeze-out speed bounds	0.2–0.5 m/s around the Reynolds $r V/(2h)$	authored (section 3.4)
red wine absorbance per cm	A620 0.8, A520 3.6, A420 3.3	Glories values [S26]; OIV definition [S28]
white / rosé absorbance	(0.01, 0.03, 0.10) / (0.08, 0.6, 0.35) per cm	estimates
particle spacing dx	2 mm (8 μ L)	authored from R M3/M4 (bore 9 particles across); section 10.2
kernel support, CFL	2 dx, 0.4 dx per substep	authored
density / divergence tolerance	$1 \cdot 10^{-4}$ mean, $1 \cdot 10^{-2}$ max / $1 \cdot 10^{-3}$	SPLisHSPlasH defaults; at $1 \cdot 10^{-3}$ a resting glass's load swung 0.6–2.7 N
Jacobi relaxation, warm start	$\omega = 0.5$; 1.0 (density), 0.5 (divergence)	IISPH value; measured divergence at $\omega = 1$
XSPH	0.02	numerical (Monaghan 1989); section 10.3 measures its damping
squirt cap, DFSPH factor floor	8 m/s; $\frac{1}{4}$ of the lattice denominator	authored limiters (section 3.4)
air γ , p_{atm}	1.4, 101,325 Pa	literature
headspace suction bound, venting β	$\rho g h_{\text{column}}$ (the column held); 0.80	physical bound; β calibrated to Kenton's 80° rate [S33] (Clanet & Searby's ideal bottle: 0.4–0.5 [S29])
glass density, IOR	2500 kg/m ³ (scaled to datasheet masses); 1.52	literature; ISO 3591 crystal ≥ 1.520 [S6]
friction pad–glass / every glass contact	0.8 / 0.4 (effective MuJoCo values)	barcode_scanning pad; Engineering ToolBox table [S41] (section 2.1)
bar top height, depth	1.090 m, 0.66 m	[S9, S10]
delivery spots	$x = \pm 0.12, \pm 0.36$ m at $y = 0.42$ m, radius 55 mm	authored (Toni-style [S20])
FR3 payload, Franka Hand force/speed	3 kg; 30–70 N total of both jaws, 50 mm/s, 80 mm stroke	[S35], [S36], franka_gazebo [S42]
reflex thresholds, filter, recovery	10 / 20 N m; 20 ms; 1 s	libfranka setDefaultBehavior; estimate; authored

quantity	value	provenance
evaluation: spill, drip film, upright, settle	0.2 mL; 0.05 + 0.1 mL; 10°; robot still 2 s, target liquid mean speed $\leq$ 15 mm/s	authored
evaluation: spot, lined window, carafe return, bottle-glass contact	55 mm (inside the lit ring); 175–185 mL; 30 mm; none	ring radius; UK SI 1988/120 reg. 14 [S5]; authored; sommelier standard [S1]

9. Tasks and evaluation

Ten tasks in three presets (config.TASKS, layout.py); seeds 0-9999 vary the delivery spot, bottle fill (450-750 mL), the wine of a flight, leftovers and ± 15 mm / $\pm 20^\circ$ placement; seeds 0-999 split 80/10/10 train/validation/test by SHA-256 of (task, seed) (variants.py, benchmark.json).

preset	task	instruction (abridged)	goal checked
service	pour-glass	pour 150 mL of red wine into the Bordeaux glass at spot k	135-165 mL in the glass
service	taste-pour	the host's taste	30-45 mL
service	split-bottle	share an opaque bottle's 250-450 mL between two glasses	within 10 %, ≤ 15 mL left
service	water-tumbler	250 mL of water from the carafe	225-275 mL
service	carry-glass	carry a glass holding 150 mL to spot k	≥ 149.3 mL kept ($150 - 0.2$ spill tolerance $- 0.5$), glass on the spot
service	refill-glass	top a glass with 40-90 mL up to 150 mL	135-165 mL
tasting	tasting-flight	50 mL of the seeded wine into each of three ISO glasses	42-58 mL each
tasting	decant-carafe	empty the bottle into the carafe	≤ 30 mL left
tasting	dump-and-pour	dump the old sample into the bucket, pour 50 mL of the new wine	old ≤ 2 mL, new 42-58 mL
lined	lined-measure	fill the lined goblet at least to its etched 175 mL line, at most 10 mL over	175-185 mL (a short measure fails)

Common checks (task.py): spilled liquid (bar + floor + anywhere else) ≤ 0.2 mL; liquid within 3 mm of the outside of the source ≤ 0.15 mL (0.05 mL drip tolerance + 0.1 mL for particles resting against the wall: with wetting on this film is the dribble, section 10.7); the source upright on its coaster, or — the carafe, which has none — within 30 mm of where it stood; the source never touched a target glass or the dump bucket during the episode (sommelier standard [S1]); the environment samples the bottle-glass contact force every 10 ms, env.source_touch); targets upright ($\leq 10^\circ$) and on their spot (≤ 55 mm: inside the lit ring the guest sees; the research's 15 mm is tighter than any bar serves to); arms still for 2 s and released; liquid settled (in each target glass the mean particle speed ≤ 15 mm/s and the fastest ≤ 0.25 m/s, and nothing in flight — more than 2 cm above the bar, outside every container — faster than 0.25 m/s). The first version averaged every particle, so a deep bottle's 10-15 mm/s noise floor could decide the check (section 12.1); the second also failed whenever wine was spilt, because a spill keeps spreading over the bar and into the coaster dish at more than 0.25 m/s two seconds later (no surface tension), found by the evaluator tests (9.1). score is the fraction met; success needs all. Volumes are counted by liquid from the particles inside each container's exact interior polygon (inventory.py), never from a level estimate.

9.1 Adversarial evaluator tests (probes/evaluator_tests.py, evidence/evaluator/tests.json)

Each case starts from a task's reset state, edits it the way a particular failure would leave it (liquid moved by the environment's own fill and preparation; bodies moved by writing their pose, the liquid inside them with them), runs 2.2 s with the robot still and evaluates. These are evaluator tests (state writes), not physics evidence.

case	edit	checks failed	expected
positive_150mL	150 mL moved from the bottle into the glass	none (success)	none
untouched_start	nothing	volume	volume
over_pour_200mL	200 mL into the glass	volume	volume
spill_2mL_on_bar	150 mL into the glass and 2 mL onto the bar	no_spill	no_spill
glass_off_spot	150 mL in the glass, glass moved 9 cm	target_on_spot	target_on_spot
bottle_off_coaster	150 mL in the glass, bottle moved 25 cm off its coaster	source_on_coaster	source_on_coaster
lined_line_exact	the lined glass filled to its 175 mL line	none (success)	none
lined_short_10mL	165 mL in the lined glass	volume	volume
lined_short_3mL	172 mL: a short measure, illegal under SI 1988/120	volume	volume
lined_over_8mL	183 mL: inside the 10 mL over-measure allowance	none (success)	none
bottle_touched_glass	150 mL in the glass; the bottle stood against the glass for 0.3 s and back	never_touched_glass	never_touched_glass

case	edit	checks failed	expected
carafe_not_returned	water-tumbler: 250 mL in the tumbler, the carafe left 6 cm from where it stood	source_returned	source_returned

All twelve fail exactly what they remove, and the three positive cases pass. Getting there found faults in the harness and the evaluator: the harness first moved a glass without its wine (the wine stayed and spilled); the "settled" check averaged every particle and then counted spilt wine still spreading over the bar (section 9); a spill blob centred on the bar top put its lowest layer inside the bar; and the first bottle-glass contact case overlapped the two vessels' widest radii, which are at different heights, so they never touched.

10. Physics validation against theory and measurement

Each experiment states its prediction first. V-probes use a bench (probes/bench.py): one container held fixed or moved kinematically at the origin, filled and prepared exactly as the environment does, no robot — measured forces are the liquid's alone. Full results with per-sample series are in the JSON files named.

10.1 Level, volume and hydrostatic load at rest (V1; evidence/v_level/level.json)

Prediction: the free surface lies at the container's geometric level for the particle volume (exact interior integral, geometry.level_for) and the liquid presses on its container with exactly its weight. Measured after preparation plus 0.4 s free dynamics: surface = 98th percentile of particle centres in the inner 60 % of the radius + $dx/2$; load = vertical liquid force on the container, mean over the last 0.2 s.

container, liquid	particles	geometric level (mm)	simulated (mm)	error as volume (mL)	load / weight	per-ms load fluctuation (SD)
PURE, 150 mL red	18,750	151.9	151.1	-5.1	-0.11 %	0.30 N on 1.46 N
PURE, 37.5 mL red	4,687	132.4	131.2	-5.4	-0.04 %	0.009 N
ISO 3591, 50 mL white	6,250	80.1	79.3	-2.5	-0.09 %	0.11 N
lined goblet, 175 mL red	21,875	132.7	131.7	-4.7	+0.15 %	0.34 N
tumbler, 250 mL water	31,250	78.1	76.7	-5.7	+0.24 %	0.51 N
1 L carafe, 500 mL water	62,500	78.9	77.6	-8.9	+0.04 %	0.36 N
LEONOR, 750 mL red	93,750	247.0 (neck)	218.0	-9.6	+0.21 %	0.88 N

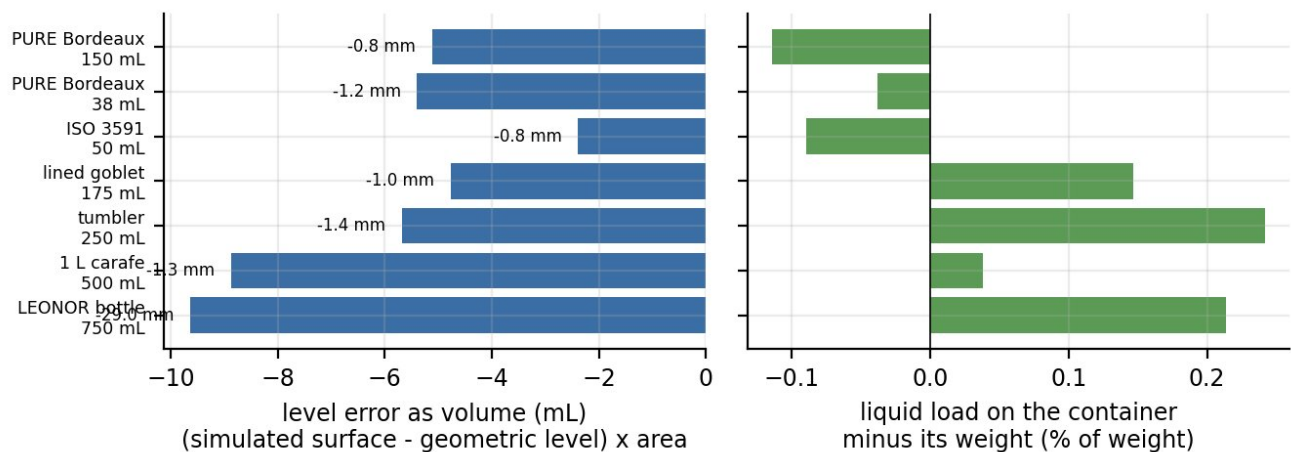


Figure 4. V1: level error expressed as volume, and the liquid's load on its container relative to its weight.

The hydrostatic load is exact to 0.3 %. The surface sits 0.8–1.4 mm below the geometric level in every open vessel (in the bottle the level is in the 18.5 mm neck, where 9.6 mL is 29 mm of height: the reason a full bottle's onset is late, section 10.5): the liquid occupies 1.6–5 % less volume than its particle count says, a deficit that scales with the wetted wall area. The particles settle slightly closer to the walls than the $dx/2$ resting distance the density map was scaled for, a little more with wall adhesion on (the first build, without surface forces: 0.6–1.2 mm). Consequence for the tasks: the evaluator counts particles (exact), so a policy that fills the lined goblet until the surface touches the 175 mL line delivers about 4.7 mL more than the line by count, which is why the lined window is 175–185 mL by count: a fill judged by eye at the line passes, a short measure fails. The per-millisecond load fluctuation is pressure noise of the iterative solver; its mean is exact and the coupling filters it over 8 ms (section 4).

10.2 Resolution (V1b; evidence/v_level/level_resolution.json)

The same fills at three spacings (each step down in spacing, $\times 1/1.25$, multiplies the particle count by 1.95 and the cost by about 2.4: more particles and proportionally shorter substeps):

case	dx 2.5 mm	dx 2.0 mm (default)	dx 1.6 mm
PURE 150 mL: level error, as volume	-0.94 mm, -6.2 mL	-0.77 mm, -5.1 mL	-0.64 mm, -4.2 mL
lined 175 mL: level error, as volume	-1.26 mm, -5.8 mL	-1.04 mm, -4.8 mL	-0.86 mm, -4.0 mL
load / weight (both)	≤ 0.02 %	≤ 0.15 %	≤ 0.03 %

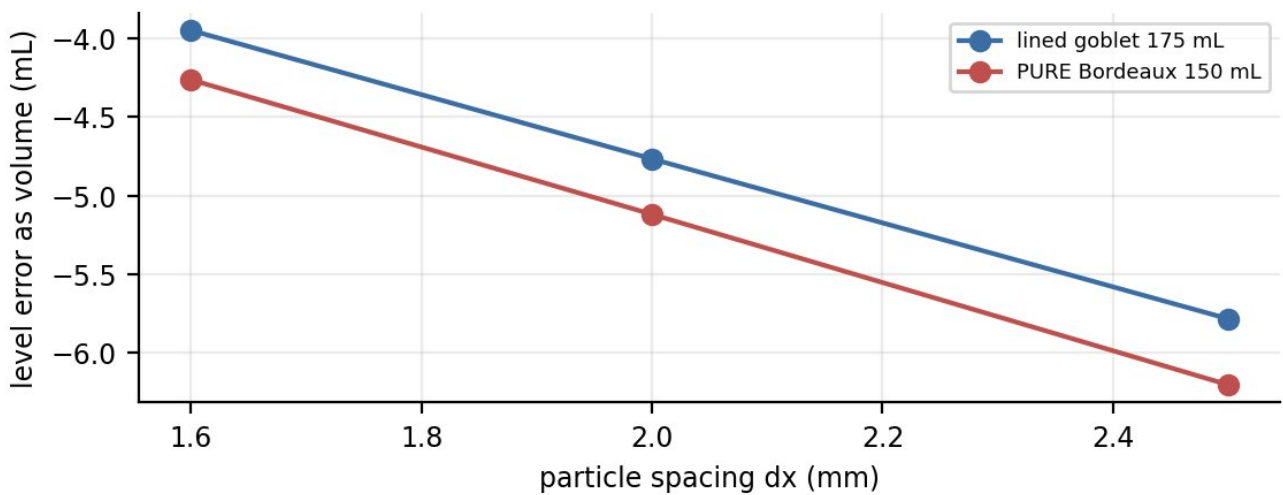


Figure 5. V1b: level error against spacing.

The level deficit shrinks with the spacing roughly in proportion (0.38–0.50 dx at 2.5 mm, 0.39–0.51 dx at 2.0, 0.40–0.54 dx at 1.6 mm): it is the near-wall packing at the particle scale. Refining to 1.6 mm improves the line reading by 0.15 mm at about 2.4 times the cost; 2 mm stays the default (research M3/M4: the bore is nine particles across). Section 10.3 shows the sloshing frequency is unchanged between 2.5 and 2.0 mm.

10.3 Sloshing frequency and damping (V2, V2b; evidence/v_slosh/slosh.json, ablation.json)

Prediction: after an 8 mm lateral step in 0.12 s, the liquid's centre of mass oscillates at the first antisymmetric mode of a circular cylinder of the surface radius R and depth h, $f = (1/2\pi)\sqrt{(gk + \sigma k^3/\rho) \tanh kh}$, $k = 1.8412/R$ (Ibrahim 2005; the capillary term adds 1.0 % to f for water in the 36 mm tumbler, Bond number 172). The tumbler is nearly cylindrical (33–40 mm radius); the PURE bowl tapers, so there the cylinder formula is only an estimate (research prediction: within 20 %).

case	f theory, gravity only (Hz)	error	damping ratio
tumbler 250 mL water, dx 2 mm	3.577	+2.9 % (+1.9 % against the capillary-corrected mode)	5.2 %
tumbler 120 mL water	3.642	+3.8 %	6.3 %
PURE 150 mL red wine	2.978 (cylinder estimate)	–8.5 %	4.4 %
tumbler 250 mL, CFL 0.2, ≥ 4 substeps	3.577	+3.3 %	5.7 %
tumbler 250 mL, dx 2.5 mm	3.577	+2.4 %	5.6 %
tumbler 250 mL, XSPH off	3.577	+1.0 %	4.8 %
tumbler 250 mL, wall shear off	3.577	+3.0 %	5.4 %
tumbler 250 mL, wall shear and XSPH off	3.577	+0.7 %	4.8 %
tumbler 250 mL, dx 1.6 mm	3.577	–1.6 %	2.8 %
1 L carafe , 400 mL water, straight 47 mm bore	3.088	+3.4 %	4.3 %

(Round 3's runs. Round 2's, r2_final/v_slosh/slosh.json, differ by ≤ 0.5 points in frequency and ≤ 0.2 points in damping.)

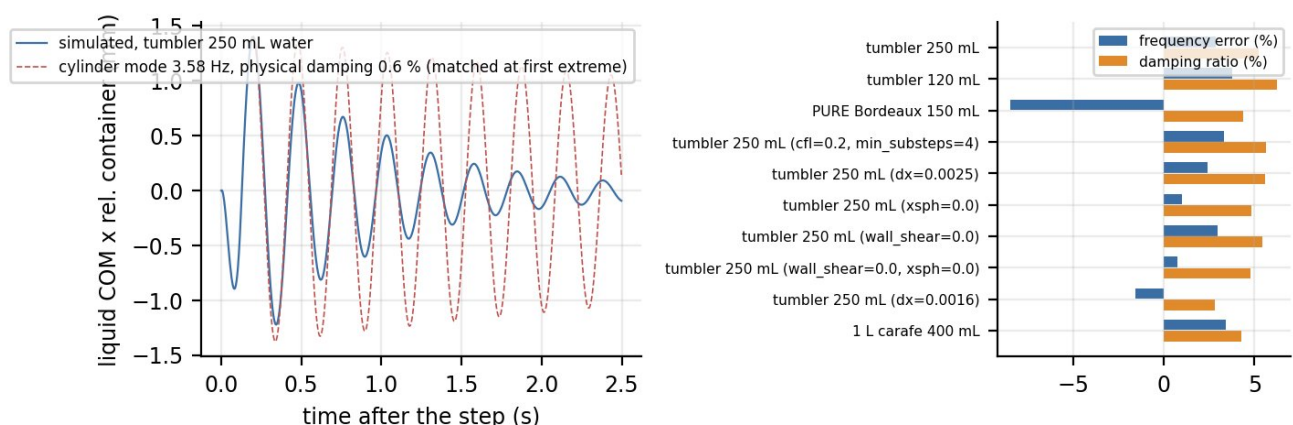


Figure 6. V2: centre-of-mass oscillation in the tumbler against the cylinder mode (left); frequency error and damping ratio of every case (right).

The sloshing frequency matches the cylinder mode to 2–3 % and hardly changes with the timestep or the spacing. The damping does not match: 5.2 % of critical at 2 mm, against ≈ 0.6 % for water in a clean 36 mm cylinder (Ibrahim ch. 4, $\zeta \approx 0.83 \sqrt{\nu / (g^{1/2} R^{3/2})}$) and 1–3 % typically measured in real glassware with contact-line losses.

V2b changes one numerical ingredient at a time (tumbler 250 mL):

ablation	f error	damping
default (surface forces on: cohesion in the surface layer, adhesion with the lip closure)	+2.9 %	5.2 %
surface forces off	+1.4 %	6.9 %
cohesion everywhere (round 1's first model)	+71 % (round 2: +7.9 %; the bulk lattice locks; see below)	10.7 %
cohesion where $s < 0.97$ instead of 0.90	+3.0 %	5.3 %
adhesion off (cohesion on)	+1.2 %	6.9 %
inviscid	+3.0 %	5.4 %
divergence tolerance $1 \cdot 10^{-5}$ (default $1 \cdot 10^{-3}$)	+3.0 %	5.5 %
no divergence warm start	+3.1 %	5.3 %
density tolerance $1 \cdot 10^{-5}$ (default $1 \cdot 10^{-4}$)	+3.1 %	5.4 %
guard clearance 0.1 dx (default 0.3 dx)	+4.5 %	5.8 %
Jacobi ω 0.3 (default 0.5)	+3.2 %	5.3 %

The rejected "cohesion everywhere" model now gives a slosh frequency 71 % high, against round 2's 8 %. Round 3 changed the guard (position correction inside the clearance) and the lip closure; which of them interacts with bulk cohesion this way was not investigated, because the package does not use that model. It is recorded so the ablation is complete.

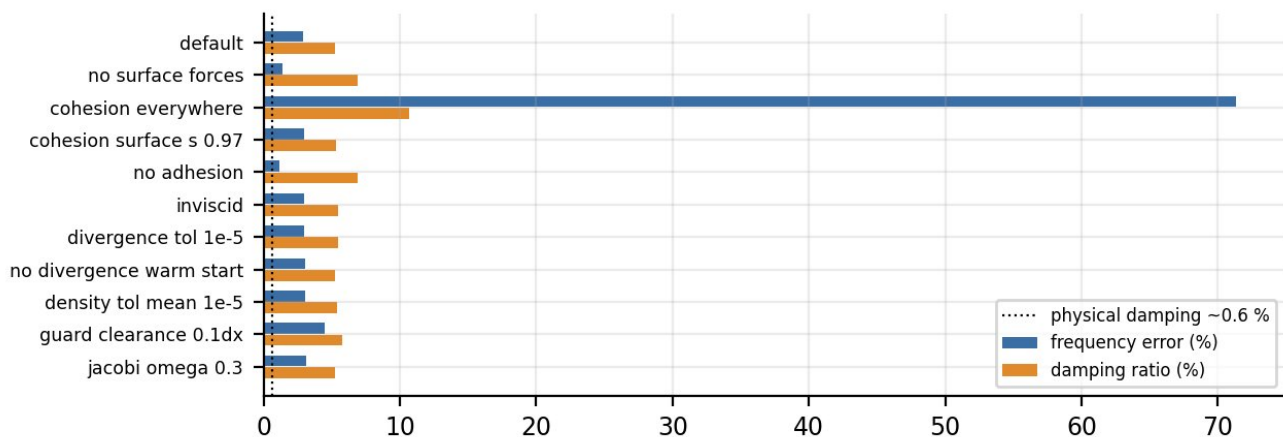


Figure 6b. V2b: frequency error and damping ratio of each ablation; dotted: physical damping.

What sets the frequency: the cohesion, through the lattice stiffness described in section 3.4 (+8 % applied in the bulk, +3 % in the surface layer), the wall adhesion (without it the frequency is the capillary-corrected cylinder mode, 3.612 Hz, to 0.03 %) and XSPH (off: +0.8 %). The ingredients do not add: each of adhesion and XSPH alone accounts for the 2 % above the capillary-corrected mode. What sets the damping is not the physical viscosity (inviscid: unchanged), the solver tolerances, the warm start or the Jacobi relaxation (all within ± 0.3 points); XSPH accounts for about 0.6 points, and adhesion lowers the damping by 1.5 points.

Where the remaining 4–5 % comes from (round 2). Round 1 asserted, without an experiment, that it comes from the discrete pressure projection and "not resolution" (the qualitative audit, round 2, called this a hypothesis). Two cases now separate the candidates:

- **Spacing.** Refining from 2.0 to 1.6 mm (61,035 particles instead of 31,250) halves the damping, 5.2 % $\rightarrow$ 2.8 %. At 2.5 mm it is 5.6 %. The excess is numerical dissipation that falls with resolution. It is not a fixed property of the solver at any spacing, as round 1's text said on the strength of the 2.5 mm case alone.
- **Container size.** The 1 L carafe's straight 47 mm bore damps at 4.3 %, against 5.2 % in the 36 mm tumbler. The ratio, 0.82, is the radius ratio, 0.77. A wall (Stokes-layer) loss and a free-surface discretisation loss with the wave number $k = 1.84/R$ both scale like $1/R$, so this case does not tell them apart. The spacing case does: viscosity is not the source (the inviscid case is unchanged), so the loss is numerical and comes from the particle discretisation of the moving surface and walls.

At 1.6 mm the decay time is $1/(\zeta \omega) = 1.6$ s instead of 0.83 s; extrapolating the two finest spacings, the physical 0.6 % would need about 1.2 mm (three times the particles of 1.6 mm, five times those of 2 mm). That is outside the runtime budget (section 13). Bound on its effect at 2 mm: a slosh decays with an e-folding time of 0.83 s in the tumbler instead of ~ 7 s, so a carried glass calms faster than a real one and carry-glass is

easier. In E6 a 0.36 m carry in 0.35 s at constant speed still throws wine over the rim (section 12.2), and a smooth 1.6 s carry does not; a real glass does the same at those accelerations. What the over-damping hides is the build-up of a slosh over successive accelerations. A real carry that brakes and accelerates at the slosh period (0.28 s) can pump the slosh up over several cycles; here the slosh amplitude falls by $e^{(-2\pi\zeta)} = 0.72$ per cycle (0.96 for the physical 0.6 %).

10.4 Tipping a glass (V3; evidence/v_tip/)

Prediction: a horizontal push at rim height h tips the glass about its foot rim (radius a) when $F h > M g a$: $F = M g a / h$ with $a = 42$ mm, $h = 267$ mm: 0.324 N empty ($M = 0.210$ kg), 0.553 N with 150 mL ($M = 0.359$ kg). The push (ASSISTED phantom force at rim height on the axis) ramps at 0.25 N/s; "tipped" is 2° of rotation.

glass	predicted (N)	simulated (N)	error	after
empty	0.324	0.362	+11.8 %	fell over
150 mL red wine	0.553	0.555	+0.2 %	fell over; 36.3 mL out of the glass by the end of the run (3.5 s)

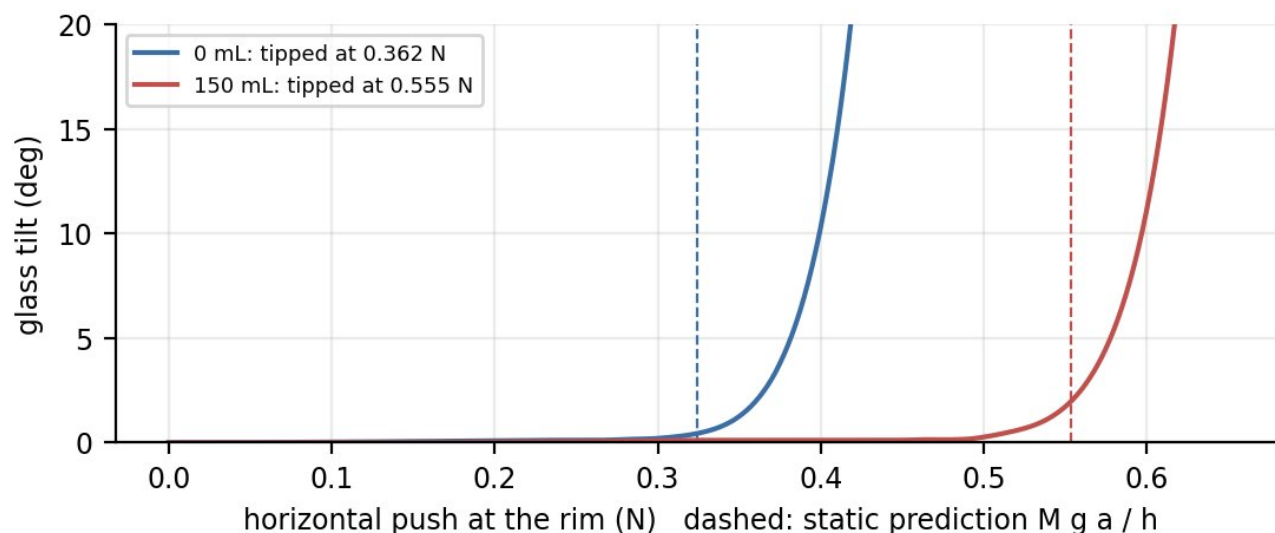


Figure 7. V3: glass tilt against push force; dashed lines are the static predictions. Videos v_tip/tip_pour-glass.mp4, v_tip/tip_carry-glass.mp4.

The empty glass reads 12 % high because the 2° criterion is reached ~0.04 N after the rotation starts at this ramp rate; the full glass's wine shifts toward the pivot as it starts to tip and offsets that. Both are inside the research prediction (± 15 %).

10.5 When a bottle starts to pour, and how fast (E2; evidence/e2_onset/)

Prediction (research M1, geometry only): a still liquid of volume V reaches the lip of the LEONOR when the bottle axis is at the elevation where the free surface passes the lip — 16.5° above horizontal at 750 mL, 7.75° at 600 mL, 4.25° at 450 mL (voxel integral of the interior, e2_onset.geometric_onset). Onset in the simulation is defined as 1 mL out; a phantom hand (ASSISTED) lifts the bottle 0.1–0.2 m, tilts it about its lip to 6° before the geometric onset, then on at 2°/s, and after onset holds 2, 5 and 10° further (3 s each, flow measured over the last 2 s).

fill	geometric onset (elevation)	simulated (1 mL out)	difference	steady flow at +2° / +5° / +10° (mL/s)	headspace at +10°	on the outside > 10 mm below the lip, during the holds
750 mL (e2_onset/fill750_rate2_gas.json, video)	16.5°	16.82°	0.3° early	3.2 / 6.4 / 28.4	+17 Pa (vented)	0.000 / 0.000 / 0.000 mL
600 mL (e9_dribble/e9_return10.json: E9's first run)	7.75°	7.95°	0.2° early	9.3 / 29.1 / 48.3	−0.2 Pa	0.000 / 0.000 / 0.000 mL
450 mL (e2_onset/fill450_rate2_gas.json)	4.25°	4.97°	0.7° early	7.1 / 25.4 / 38.8	+1.8 Pa	0.000 / 0.000 / 0.000 mL

These are round 3's runs, on the lip closure grounded in each lip's own edge (section 3.4). Two earlier versions of this table were wrong:

- Round 1's gave 0.0 mL on the outside while its own JSON said otherwise (qualitative audit, round 2).
- Round 2's REPORT quoted its pre-closure runs (onsets 17.1 / 8.07 / 4.97°, 1.2–6.9 mL on the outside, +18 Pa) under file names that by then held the closure runs (qualitative audit, round 3). Those runs are archived in r2_pre_closure/, and round 2's closure runs in r2_final/. Round 2's closure runs gave onsets of 16.82 / 8.11 / 4.89° and 0.000–0.072 mL on the outside.

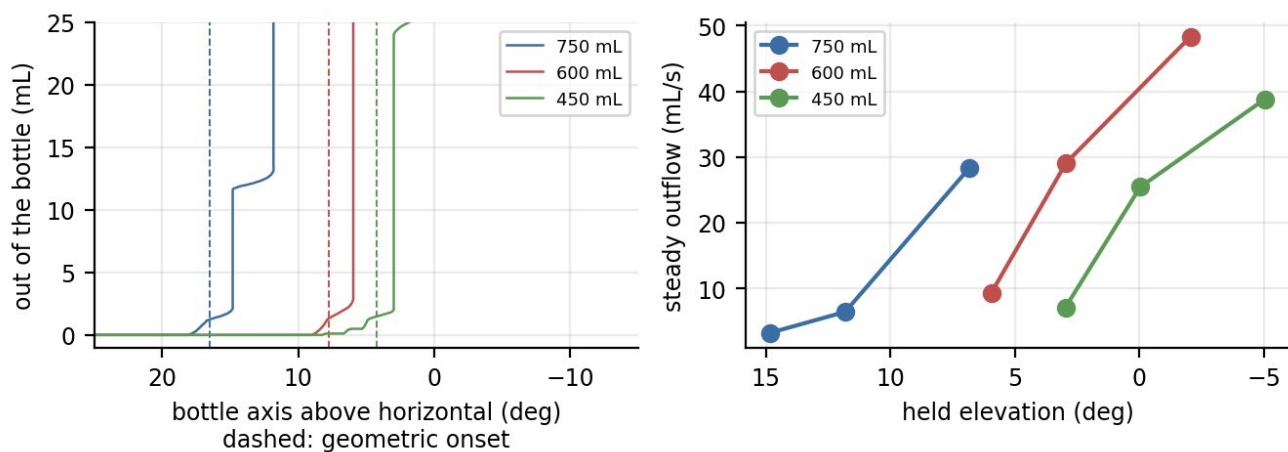


Figure 8. E2: liquid out of the bottle against axis elevation (dashed: geometric onset) and the steady outflow at each held tilt. Video e2_onset/fill750_rate2_gas.mp4.

The onset follows the geometric prediction across fills, 0.2–0.7° early (the first build, without surface forces: 0.7–2.4° late). The wetting meniscus at the lip stands slightly above the flat-surface level, so liquid reaches the lip a little before the still surface would. All three are inside the research's $\pm 1.5^\circ$ allowance.

The flow then grows with tilt from a trickle to 28–48 mL/s at 10° past onset, the research's weir estimate (a few mL/s at onset to 30–60 mL/s). The headspace stays vented throughout: in this window the stream does not seal the neck. Surface tension slows the first degrees after onset (750 mL: 3.2 and 6.4 mL/s at $+2^\circ$ and $+5^\circ$, against 4.7–11.9 and 13.7–14.9 mL/s without it).

Nothing runs down the outside of the neck during the holds at any fill. The lip holds a film of 0.09–0.16 mL while it pours, and drips from the lip fall with the stream. How this compares with a real bottle is section 10.7's question. The robot pours of section 11 found the onset the same way: flow started 1.3° (658 mL) past the geometric onset for the fill at a $12^\circ/\text{s}$ approach.

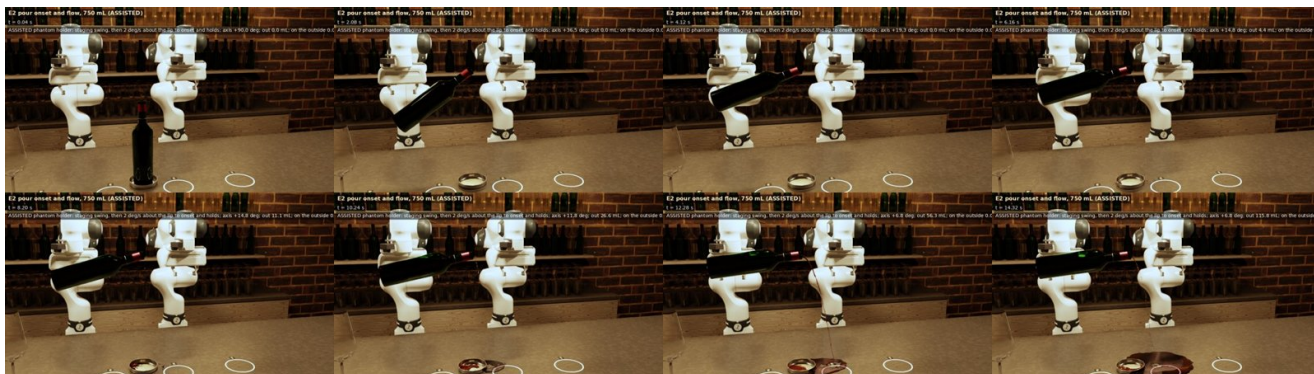


Figure 9. E2 750 mL, eight frames (guest-side probe camera; ASSISTED phantom holder).

10.6 Emptying a held bottle: Kenton's measurement (E3; evidence/e3_kenton/)

Kenton, Neufeld & Huppert (2012) [S33] emptied a 0.77 L wine bottle of water held at fixed angles: ≈ 95 g/s mean at 40° from horizontal and 85–95 g/s at 50 – 90° . Here a phantom hand lifts the LEONOR, filled to its 750 mL line with water, over the dump bucket in 1 s, swings it about its lip to the angle (axis below horizontal, mouth down) in 2 s and holds it 6 s; flux = the mean over the held phase. The venting coefficient β was calibrated on the 80° case (section 5); the 40° case is the check.

angle	gas model	simulated flux (held mean; from the first 5 g to 90 % out)	Kenton	ratio	round 1 / round 0
40°	on	88.4 g/s; 88.6 g/s	≈ 95 g/s	0.93	89.1 / 62 g/s
80°	on (calibration)	84.7 g/s; 86.7 g/s	85–95 g/s	0.89–1.0	85.1 / 51 g/s
40°	off	(83.2 g/s held mean, meaningless: the bottle is empty after 2 s); 334 g/s	≈ 95 g/s	3.5	334 / ≈ 300 g/s peak

(Round 3's physics; round 2's runs, r2_final/e3_kenton/, gave 89.2, 86.6 and 325 g/s. The no-gas run's held_flux_g_s is the mean over a held phase in which the bottle has already emptied, as the qualitative audit noted; rate_5g_to_90pct_g_s is the meaningful rate and is now written into every E3 file.)

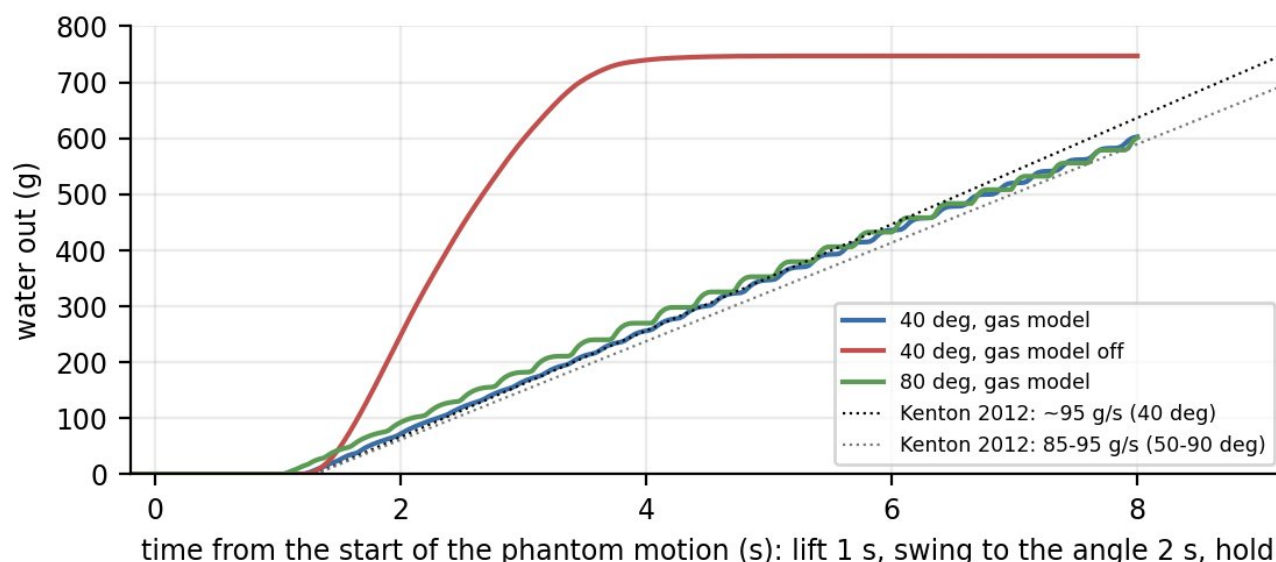


Figure 10. E3: water out of the held bottle against time with and without the headspace model; dotted: Kenton's mean rates. Video `e3_kenton/theta40_gas.mp4` (the 80° and no-gas runs were not filmed this round).

Without the headspace the bottle empties like a vented orifice. With it, the outflow is throttled to the measured rate and comes in pulses, the closure's seal/vent cycle standing in for the bubbles it does not resolve (no measured glug rate of this bottle was found to compare with). The two constants that set it were revised this round on physical grounds, not tuned to the 40° number: the suction bound is the column the bottle holds, $\rho g h$ (the first build's 1.5 $\rho g L$ let the suction exceed the column and itself set the 40° flux), and $\beta = 0.80$ reproduces the 80° measurement, the regime where the venting rate alone sets the flux. The bucket caught 511 g at 40° and 533 g at 80° (round 2: 497 and 535 g; round 1: 550 and 528 g); the rest splashed onto the bar beside it (visible at the end of the video). Service pours stay out of this regime as long as the bottle is not tipped far past onset: in E2's holds up to 10° past onset the neck stayed vented ($|\Delta p| \leq 16$ Pa, section 10.5).

10.7 The dribble: wetting, tilt-back speed and a hydrophobic lip (E9; evidence/e9_dribble/)

Prediction (Duez et al. 2010, research R\$7.5): a liquid stream leaving a wettable lip slowly follows the solid's outside instead of separating — the teapot effect — and the critical speed below which it does so grows with the wettability; on a hydrophobic lip the stream separates at any speed. A bottle tilted back slowly passes through a long low-flow phase and should therefore leave more wine running down its neck than one tilted back quickly. E9 runs E2's schedule at 600 mL (onset, holds at +2, +5, +10° past it), then tilts the bottle back upright at 10 or 90°/s and watches 2.5 s; a third run makes the glass hydrophobic ($\theta = 110^\circ$ instead of 20°) and repeats the slow tilt-back. Measured from the particles: liquid within 3 mm of the bottle's outer surface (film, includes the lip, where the stream itself passes), and liquid on the outside more than 10 mm below the lip (the dribble running down the neck).

Round 3's runs, on the closure grounded in each lip (section 3.4):

run	film at the lip / dribble during the holds at +2, +5, +10° (mL)	dribble, max during the tilt-back	at the end
θ 20°, tilt back at 10°/s	0.11 / 0.000, 0.16 / 0.000, 0.18 / 0.000	0.000 mL	0.0 mL below the lip; 0.0 mL on the lip
θ 20°, tilt back at 90°/s	0.11 / 0.000, 0.16 / 0.000, 0.18 / 0.000	0.000 mL	0.0 mL; 0.0 mL
θ 110° (hydrophobic), 10°/s	0.12 / 0.000, 0.16 / 0.000, 0.17 / 0.000	0.000 mL	0.0 mL; 0.0 mL

Rounds 1-2 kept wine on the outside of the neck:

- round 1, and `r2_pre_closure/e9_dribble/`, before any closure: 2.0–3.2 mL during the holds and 3.6–4.1 mL at the slow tilt-back;
- round 2's closure (`r2_final/e9_dribble/`): 0.02 mL during the holds and 0.10 mL at the slow tilt-back, against 0.02 mL at the fast one.

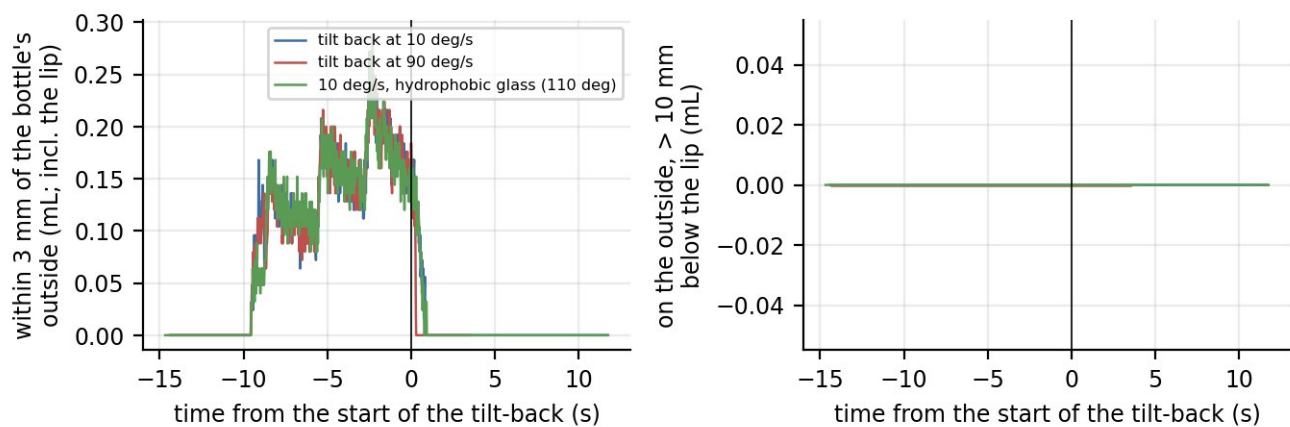


Figure 10b. E9: liquid on the outside of the bottle near the lip (left) and running down the neck (right), against time from the start of the tilt-back.

Is the dribble's size physical? (E9r, E9d, sensitivity; evidence/e9_lip/).

Round 2's closure was not converged. E9r runs the E2/E9 schedule at 600 mL with one 3 s hold 5° past onset, then a 10°/s tilt-back, at three spacings. With round 2's closure ($r_i = e_0 = dx$ everywhere), the tilt-back dribble grew with refinement, past the 0.2 mL spill tolerance at 1.6 mm (qualitative audit, round 3). Without any closure (round 2's first study) it grew from 0.16 to 3.9 mL, and round 1's claim that the 2 mm film was "an upper bound" was withdrawn. With the closure grounded in each lip's own edge radius and the local liquid thickness, applied only where the flow turns round the lip:

spacing	particles	onset (geometric 7.75°)	flow in the hold	outside > 10 mm below the lip (hold mean)	at the lip	max during the tilt-back	round 2's closure: hold / tilt-back
2.5 mm	38,400	6.83°	38.2 mL/s	0.000 mL	0.17 mL	0.000 mL	0.000 / 0.000 mL
2.0 mm (default)	75,000	8.07°	32.0 mL/s	0.000 mL	0.16 mL	0.000 mL	0.064 / 0.136 mL
1.6 mm	146,484	10.27°	20.2 mL/s	0.016 mL	0.15 mL	0.041 mL	0.166 / 0.594 mL

The tilt-back dribble is now bounded well below the 0.2 mL tolerance at all three spacings, and the film at the lip converges (0.17 / 0.16 / 0.15 mL).

What did not converge. The onset at 1.6 mm is 2.5° early, outside the research's $\pm 1.5^\circ$, and the flow just past it falls with refinement. Both are set by the static wetting at the lip, which the closure does not touch.

Against Duez. Duez et al. (PRL 104, 084503, 2010) find a threshold Weber number for dripping, $We_c = \rho U_c^2 e_0 / \sigma = k (R r_i / e_0^2) (1 + \cos \theta)$, with $R = r_i + e_0 / 2$ and $k \approx 0.44$ read from their Figure 3 inset. Here r_i is the edge's radius of curvature and e_0 the film thickness. For the LEONOR's lip ($r_i = 1.5$ mm, its geometry) and wine ($\theta 20^\circ$, $\sigma 47$ mN/m) this gives two regimes:

- **Steady pour** (25 mL/s, a stream about 4 mm thick): $U_c \approx 0.06$ m/s, while the stream leaves the lip at $U \approx 0.5$ m/s. Duez predicts clean separation. The model agrees: nothing on the outside at 7–49 mL/s.
- **Onset trickle and final drops** (1–3 mL/s, a real film $e_0 \approx 0.5$ –1 mm): $U_c \approx 0.35$ –0.9 m/s against $U \approx 0.25$ m/s. Duez predicts dribbling.

In the model, a trickle's film cannot be thinner than one particle, so its U_c is 0.14 m/s at 2 mm and it separates sooner than a real one would. The trickle clings to the lip (the 0.04–0.16 mL film) and drips from it rather than running down the neck. So the teapot dribble down the neck at a trickle is under-predicted at these spacings. E9 shows no difference between a slow and a fast tilt-back, or between a wettable and a hydrophobic lip, where Duez expects a small one. Round 2's closure showed that ordering, at 0.1 mL. For a pour this errs toward the clean side by at most the trickle dribble, which round 2 measured at 0.02–0.1 mL.

E9d: a pour slowed in steps (e9_lip/e9d_slowed.json, video; the research's acceptance test). The 600 mL bottle tilts at 20°/s to onset, is held 10° past it, and then is brought back in steps to 6, 3 and 1.5° past it (2.5 s each):

hold (past onset)	flow	on the outside > 10 mm below the lip	at the lip
+10°	89 mL/s (glugging)	0.000 mL	0.25 mL
+6°	27 mL/s	0.000 mL	0.14 mL
+3°	2.1 mL/s	0.000 mL	0.04 mL
+1.5°	0 (stopped)	0.000 mL	0.00 mL
tilt-back at 10°/s		max 0.000 mL	

The research's tests:

- (a) A slow stream wets the outside: only at the lip (a 0.04 mL film at 2 mL/s), not down the neck (above).

- (b) A fast pour does not: yes, nothing at 27–89 mL/s.
- (c) A hydrophobic lip does not: yes.
- (d) The transition is within a factor 2 of the Duez estimate: not shown, since the model's trickle does not dribble down the neck.

E9s: wetting sensitivity (e9_lip/e9s_*.json; the E9r schedule at 2 mm, one constant changed):

case	onset	flow in the hold	outside (hold mean)	at the lip	max during the tilt-back
σ 43 mN/m	8.03°	32.1 mL/s	0.000 mL	0.16 mL	0.000 mL
σ 47 mN/m, θ 20° (default)	8.07°	32.0 mL/s	0.000 mL	0.16 mL	0.000 mL
σ 51 mN/m	8.07°	32.1 mL/s	0.000 mL	0.16 mL	0.000 mL
θ 10°	8.11°	31.6 mL/s	0.048 mL	0.27 mL	0.136 mL
θ 30°	8.03°	32.2 mL/s	0.000 mL	0.18 mL	0.032 mL

Across the plausible range of wine's surface tension and contact angle, the onset moves by 0.1°, the flow by 2 %, and the dribble stays between 0 and 0.14 mL, the most at the most wettable lip (θ 10°). Single particles are 8 μ L, so the run-to-run spread of these small numbers is comparable to the differences.

Where the pours' spill comes from (probes/spill_origin.py, recordings only). It classifies every particle that ends on the bar by where it fell through the glass's rim plane. Before any closure, in the careful phantom pours of section 11.0, all of it fell on the bottle's side, 10–16 cm behind the glass axis: wine running back along the underside of the neck and dripping off. In round 3's robot pour-glass, 0.064 of the 0.136 mL fell through the rim plane 30–59 mm beyond the glass axis, just past the far rim

(e4_robot_pour/service_pour-glass_s0_dual-fr3_spill_origin.json), and none on the bottle's side.

10.8 The robot is a liquid boundary (V5; evidence/v_robot_boundary/)

A diagnostic state edit, not a policy action. In service/carry-glass at the ready pose, three blobs of 2.7 mL each (343 particles; round 1 used 1 mL, too small to see in the video) are released at rest, each 6 cm above the highest robot surface beneath it:

- A over the middle of the right hand: that is the wrist (link 6), which stands above the hand there;
- C 8 cm from link 7 along the hand, over its end: the qualitative audit's placement;
- B over the forearm (link 5).

Prediction: each lands on the link beneath it, is held back while the link carries it, runs off and falls on. (Round 2's first placement of A, 6 cm above link 7's top, was inside link 6. The guard pushed the blob out of link 6 downward, and it fell through the hand at free-fall speed: a placement error in the test, found because the blob showed no contact at all.)

blob	free fall would reach	at 0.2 s	first on the bar
A, over the wrist	1.536 m	1.626 m, falling at 0.92 m/s (free fall 1.96 m/s): resting on link 6, then running off	0.40 s
C, over the hand's end	1.378 m	1.424 m: grazes the rounded end and is shed at once	0.32 s
B, over the forearm	1.492 m	1.643 m: the forearm carries it (0.032 N at 0.1 s against the blob's 0.027 N weight, the landing included)	0.42 s; it slides off the robot's side to the floor behind the bar

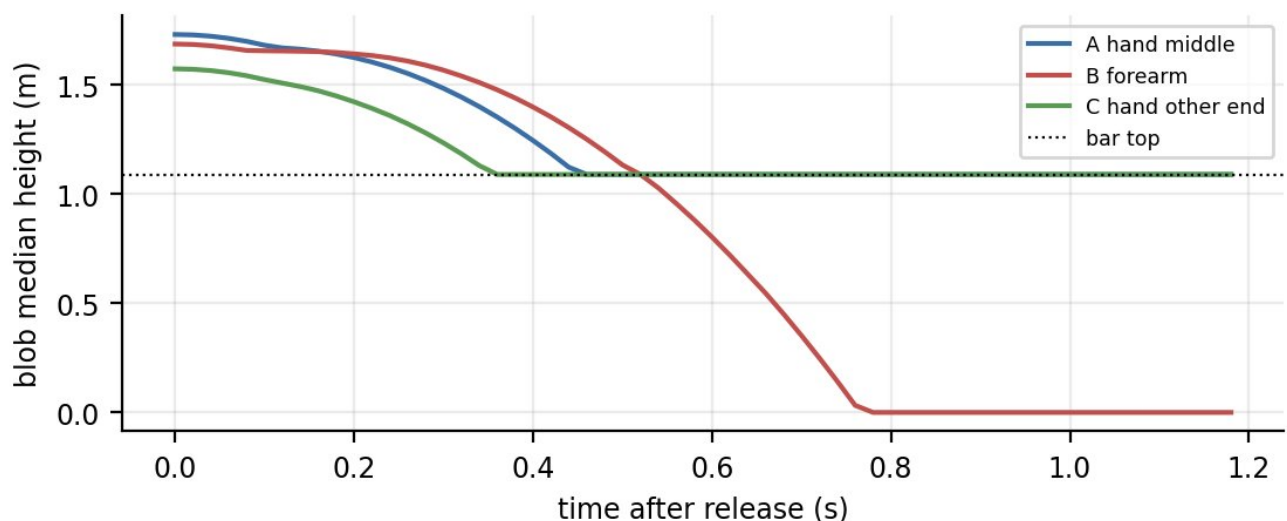


Figure 10c. V5: median height of each blob against time; dotted: bar top.

Sixteen polytopes (72–79 planes each) stand for the two arms' links and hands. The speed cap engaged in 18 of 1,200 ticks (18 particle-substeps; round 2: 17). Which particles were capped was not traced: nothing in this scene moves faster than the falling blobs' 2–4 m/s, so a few particles were briefly squeezed at an impact. There were no wall repairs and no fault. The video `v_robot_boundary/robot_boundary.mp4` (round 3) draws isolated particles as dots: the surface render alone lost the blobs once they broke up on the links (evidence audit, round 3).

11. Assisted demonstrations

Each is one continuous rollout from the task's reset state, rendered from the recording, evaluated by the unmodified task evaluator at the end. ASSISTED: a scripted operator streams IK joint targets to the arm's physical position servos and closes the Franka Hand by its force mode (70 N total); it reads the privileged inventory to decide when to stop (the eyes a policy would have to replace with its cameras). Nothing is attached: bottles and glasses are held by finger contact and friction; the liquid feels the walls, the bar, the fingers and the robot's links.

The pour primitive (`demos/scripted.py`, `demos/robot_pour.py`; operator version 6c, the final code; 11.1 lists the earlier versions):

1. side grasp 145 mm up the body (bottle) or by the neck (carafe); lift 10 cm; carry upright to beside the glass;
2. tilt to the staging angle 3 cm above the staged pose, then lower into it. The staged pose has the mouth 3.5 cm above the rim and 0.7 of the rim radius toward the bottle side of the glass centre. The staging angle and mouth height keep the bottle's heel 3 cm clear of the bar and stay 8° short of the onset for the current fill (section 6.3);
3. tilt on at 12°/s about the lip until wine arrives, and 3° more at that rate (a decisive start keeps the trickle phase short); then at 3°/s while the stream is below 45 mL/s. The commanded hand pose is corrected every 50 ms from the bottle's measured tilt and mouth position (integral gain 0.25);
4. stop when the glass plus the stream in the air plus a lead reaches the goal; the lead is what flows during the tilt-back, $\max(4 \text{ mL}, 0.3 \text{ s} \times \text{the current flow})$;
5. tilt back at 60°/s about the lip while raising the mouth 4 cm (so the neck is not pivoted onto the rim), and hold 2 s over the glass;
6. return the bottle to its coaster and release.

Version 5 (round 1's final) had the mouth 5 cm above the rim and a third of the rim radius from the centre, no flow limit, no staging lift and a 0.6 s hold; version 4 held the stream at $\leq 35 \text{ mL/s}$ and did not raise the mouth.

With wetting on, the first operator of this round failed in two instructive ways, kept in `e4_robot_pour/attempts/`: slowing the stream before the stop (a "taper", as a cautious pourer might) let the trickle follow the outside of the lip and the neck and drip onto the bar — 8 mL in 3 s while the glass gained nothing (`r1_taper_pour-glass.log`, the teapot effect, section 10.7); and with the mouth 3.5 cm over the rim and 1 cm from centre, a 90°/s tilt-back with the mouth drawn back 3 cm threw 9.4 mL of the 67 mL/s stream's tail outside the near rim and brought the neck onto the rim (28.8 N for 0.73 s) — 137.5 mL in the glass, 10.1 mL spilt (`r1_v3_pour-glass.json`).

11.0 Is a clean pour possible? Careful phantom pours (E4p; `probes/e4p_phantom_pour.py`, `evidence/e4p_phantom_pour/`)

Both round-2 audits required this. The evidence audit asked for one assisted pour that meets the 0.2 mL tolerance (phantom allowed), or a tolerance set from the model. The qualitative audit asked for proof that a clean pour is physically possible in this world. E4p takes grip lag, IK and reflexes out of the loop. A phantom hand (ASSISTED fixture) carries the bottle from its coaster to the staged pose, tilting on the way: 8° short of onset, the lip 2.5 cm over the rim and 0.7 rim radii toward the bottle side of the glass axis, the body behind the glass. It then:

- tilts at 15°/s to 3° past onset (a decisive start);
- holds the flow at 45 mL/s by the tilt rate;
- stops when the glass, the stream in the air and 0.35 s of flow reach the goal, with a 20° tilt-back at 80°/s and a 90° twist about the bottle's axis (the sommelier's twist);
- holds 2.5 s with the lip over the glass;
- goes back along the staging path and sets the bottle on its coaster.

The task's own evaluator scores the result.

run	physics	in the glass	spilt	evaluator
design A: lip 3 cm over the rim, −0.25 radii, 25 mL/s, no twist (<code>r2_pre_closure/e4p_phantom_pour/service_pour-glass_s0.json</code>)	round 2 before the Duez closure	141.8 mL	39.1 mL: 15.6 during the pour, 3.8 at the stop, 16.4 on the way back (an operator bug, fixed: the stop left the bottle 4° below the staged tilt and the return path snapped it back up toward the pour)	9/10
design B (<code>r2_pre_closure/.../service_pour-glass_s0_B.json</code>)	before the closure	140.3 mL	5.9 mL: 4.2 during the pour, 1.4 at the stop	9/10
design B (<code>r2_final/e4p_phantom_pour/service_pour-glass_s0.json</code>)	round 2's closure	143.5 mL	0.016 mL	success, 10/10
design B (<code>e4p_phantom_pour/service_pour-glass_s0.json</code> , video)	round 3 (shipped)	143.3 mL	0.008 mL (one particle, during the pour)	success, 10/10

The same design on the other glassware, on round 3's physics:

task (e4p_phantom_pour/...)	in the target	spilt	evaluator
lined-measure, UK 175 mL lined goblet (lined_lined-measure_s0.json, video)	175.2 mL (window 175–185)	0.008 mL	success, 10/10
tasting-flight, first ISO glass, 44 mm rim (tasting_tasting-flight_s1_t0.json, video)	42.9 mL (goal 50 ± 8)	0.088 mL (at the stop)	the flight's volume check fails because the other two glasses are empty; this glass is inside its window, and spill, touch and every other check pass
water-tumbler, 1 L carafe into the tumbler, onset at 15°/s (service_water-tumbler_s0.json, video)	243.1 mL (window 225–275)	0.248 mL: 0.208 at the onset, 0.040 at the stop	9/10: fails the spill check
water-tumbler, onset at 40°/s (service_water-tumbler_s0_onset40.json, video)	243.0 mL	0.024 mL	success, 10/10
decant-carafe, the whole bottle into the carafe, lip 0.7 radii toward the bottle, 2.5 cm over the 52 mm mouth, 45 mL/s (attempts/r3_decant-carafe_offset07_killed.json)	727 mL, bottle empty	22.8 mL, all of it as the bottle passed horizontal with 250–40 mL left and glugged	not scored: killed at t = 32 s (its stop waited for a carafe volume the spill made unreachable; fixed)
decant-carafe, lip 0.45 radii, 2.2 cm, 35 mL/s, onset 40°/s (tasting_decant-carafe_s0.json, video)	677 mL, 15 mL left in the bottle	57.2 mL: 56.4 during the pour, at 17–27 s	8/9: fails the spill check

Water from the carafe. Round 2's carafe pour spilt 6.3 mL. On round 3's physics, with the closure grounded in the carafe's own 1.5 mm lip, it spills 0.25 mL, almost all of it as the onset trickle runs down the carafe's side. Tipped decisively (40°/s to onset, as a waiter does), it meets the tolerance.

Decanting a whole bottle into the narrow carafe was not reached. Where the second attempt's 57.2 mL went (tasting_decant-carafe_s0_spill_origin.json):

- 17.8 mL fell through the rim plane just beyond the far rim (median 28 mm from the axis against a 26 mm mouth): the stream overshot;
- 6.0 mL fell on the bottle's side;
- 33 mL never crossed the rim plane near the mouth: it struck the carafe's lip and outside and ran down.

All of it left between 17 and 27 s, while the bottle, past horizontal, glugged. Its outflow pulses between 15 and 56 mL/s and the stream's reach changes with it, so a 52 mm mouth cannot catch all of it from 2 cm above.

The first attempt, with the lip nearer the bottle's side of the mouth, lost less (22.8 mL), in the same phase. A sommelier decants with the bottle's lip inside the decanter's neck, pouring down its wall; the evaluator's never-touch-the-glass rule forbids that here. Whether decant-carafe can meet the 0.2 mL tolerance is therefore not shown. Its tolerance is unchanged, and the task is listed as unshown (section 16).

So with a careful hand a clean pour is reachable in this world, and the 0.2 mL tolerance stands: it is not relaxed to what an operator achieves. The phantom is not a policy. It has no grip lag, and it reads the privileged inventory to know when to stop, where a policy would have to see the level through its cameras.

11.1 A 150 mL glass of red wine (pour-glass, seed 0;

e4_robot_pour/service_pour-glass_s0_dual-fr3.*, attempts in attempts/r1_*)

The left FR3 takes the LEONOR (658 mL, 82,255 particles) from its coaster, carries it to the PURE glass at spot 2, stages at 45°, tilts on; wine reaches the lip at a commanded 81.0° (bottle 81.5°) against a geometric onset of 80.5° for 658 mL. Every attempt of this round, each a full rollout scored by the task evaluator (10 checks; success needs all):

operator	mouth over the rim, offset; stream; stop	glass	spilt	where the spill came from	bottle touched the glass	score
v1 (log only, stopped)	6 cm, 2 cm; slowed before the stop	131 mL when stopped	13.5 mL and rising	onset trickle beside the near rim (6 mL); the slowed stream ran down the neck (8 mL in 3 s)	—	—
v3	3.5 cm, 1 cm; full; 90°/s, mouth drawn back 3 cm	137.5 mL	10.1 mL	onset 0.6 mL; tail at the tilt-back 9.4 mL	28.8 N, 0.73 s (neck on the rim)	8/10
v4	5 cm, ½ rim radius; held ≤ 35 mL/s; 60°/s	144.8 mL ✓	6.2 mL	onset 1.0 mL; along the neck while pouring 1.1 mL; tail 3 mL	20.0 N, 0.87 s (neck on the rim during the tilt-back)	8/10

operator	mouth over the rim, offset; stream; stop	glass	spilt	where the spill came from	bottle touched the glass	score
v5 (round 1's final code)	5 cm, $\frac{1}{3}$ rim radius; full; 60°/s rising 4 cm	130.2 mL	32.1 mL	the full stream from 5 cm overshot the far rim (8 mL); tail thrown at the rising tilt-back (20 mL)	none ✓	8/10
v5, round-2 physics before the Duez closure (r2_pre_closure/e4_robot_pour/)	the same	130.8 mL	31.9 mL	6.6 mL behind the glass (run back along the neck), 9.6 mL beyond it, 15.7 mL at the stop	none ✓	8/10
v5 with the closure (attempts/r2_v5_closure_pour-glass.json)	the same	87.2 mL	294 mL	the stream now leaves the lip cleanly and, from 5 cm up, landed beyond the far rim; the stop never came	none	8/10
v6 / v6b (attempts/r2_v6*_pour-glass.json)	the careful phantom's geometry: 2.5 cm, 0.7 rim radii toward the bottle; held ≤ 45 mL/s; 2 s over the glass after the stop (v6b: tilted 3 cm higher, then lowered)	148.1 / 147.4 mL ✓	0.048 / 0.048 mL ✓	—	12.3 N for 90 ms / 11.7 N for 30 ms: the neck brushed the near rim as the tilt began (probes/touch_times.py: 9.96–10.36 s / 10.28–10.96 s)	9/10
v6c, round 2 (r2_final/e4_robot_pour/service_pour-glass_s0_dual-fr3.json)	as v6b with the mouth 3.5 cm over the rim	145.4 mL ✓	0.008 mL ✓	—	none ✓	success, 10/10 (with the bottle slipping 7° in the pinch)
v6c, round 3 (motion limits; attempts/r3_v6c_pour-glass*.json)	the same	151.0 mL ✓	0.0 mL ✓	—	5.3 N for 60 ms: the bottle no longer slipped, so lowered into the staged pose at 45° its neck, 26 mm below the lip, reached the near rim (0.35 mm; attempts/r3_v6c_pour-glass_touch.json)	8/10 (also read before the arm had been still for 2 s)
v6d, first run (attempts/r3_v6d_pour-glassUntitled.json)	the mouth 4.5 cm over the rim	149.8 mL ✓	0.072 mL ✓	—	none ✓	9/10: read 1.9 s after the arm stopped (its last move runs on past the command under the motion limits)
v6d (final) (service_pour-glass_s0_dual-fr3.json)	as v6c with the mouth 4.5 cm over the rim; waits until the evaluator reads the scene settled	150.1 mL ✓	0.136 mL ✓	0.064 mL just beyond the far rim, one splash particle (..._spill_origin.json)	none ✓	success, 10/10

On round 3's physics and under the FR3 motion limits, the scripted robot pours a 150 mL glass that the task's evaluator accepts:

- 150.1 mL in the glass and 0.136 mL on the bar;
- the bottle never touches the glass and goes back on its coaster;
- no limiter acted: no speed cap, no impact impulse, no wall repair, no squeeze, no wall following.

The bottle now holds in the pinch (v_motion/grip_axis_pour-glass.txt, the qualitative audit's grip_axis_r3.py on this recording):

- its turn relative to the hand peaks at 2.8° and ends at 0.85°;
- the grasp point slides at most 2.8 mm;
- the operator's tilt correction stays between −0.1 and 1.0°.

In round 2 the bottle turned 7.2–7.4°, the grasp point slid 8.3 mm, and 3.8° remained after the tilt-back. That slip had lowered v6c's neck away from the rim; without it, v6c's staged pose put the neck on the rim, hence v6d's 4.5 cm.

Round 0's best pour without surface forces (152.2 mL, 0.89 mL spilt, attempts/pour-glass_*) is not comparable: there was no dribble to make. The trade-offs that remain are geometric. A stream thrown from higher

overshoots the far rim. A mouth low and close to the near rim lets the neck touch it at the 45° staging tilt. The neck must not be pivoted onto the rim at the stop. In the failed versions every other check passed: glass upright on its spot, bottle back upright on its coaster, released, settled, no protective stop, no drip film left on the bottle.

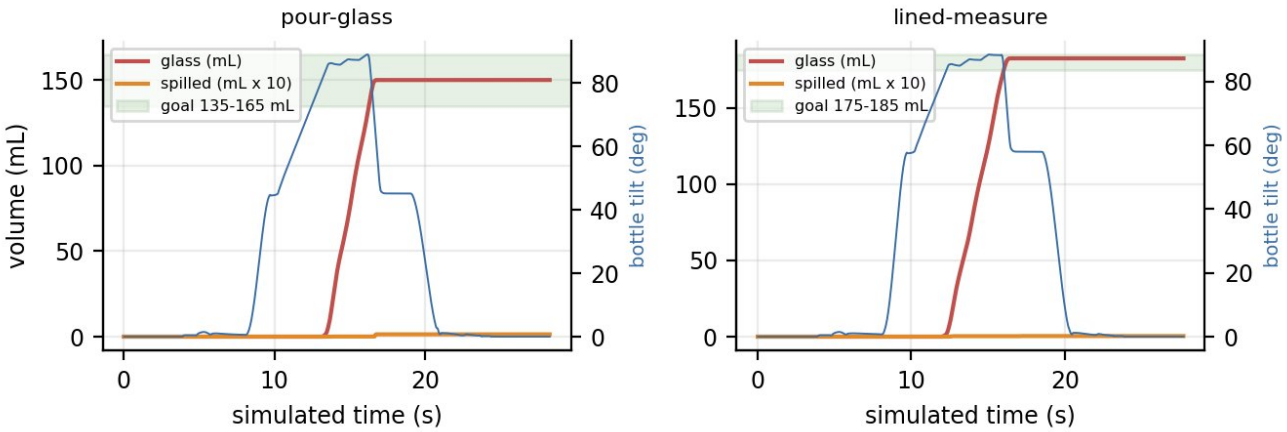


Figure 11. E4: glass contents, spill (× 10) and bottle tilt against time for pour-glass and lined-measure (operator v6d, round 3).



Figure 12. pour-glass, v6d: eight frames of the video (closeup | station_rgb).

11.2 A legal 175 mL measure (lined-measure, seed 0;

e4_robot_pour/lined_lined-measure_s0_dual-fr3.*)

The UK lined goblet (677 mL in the bottle). The heel clearance over the 183 mm glass sets the staging tilt: 62.5°.

operator	in the glass	spilt	evaluator
v5 (round 1)	159.7 mL (short of the 175 mL line)	13.0 mL: the onset trickle 1.1 mL, the rising tilt-back threw 11.6 mL of the tail over the near rim	8/10
v5, round-2 physics before the closure (r2_pre_closure/e4_robot_pour/)	160.5 mL	12.6 mL	8/10
v6b, round 2 (mouth 2.5 cm over the rim, 0.7 radii toward the bottle, ≤ 45 mL/s, 3 cm staging lift, 2 s hold; r2_final/e4_robot_pour/)	175.1 mL (window 175–185)	0.072 mL	success, 10/10
v6d, round 3, aimed at the line (attempts/r3_v6d_lined-measure_short.json)	172.4 mL	0.040 mL	9/10: short
v6d, aimed at the line + 5 mL (attempts/r3_v6d_lined-measure_aim180.json)	174.75 mL	0.056 mL	9/10: short by 0.25 mL
v6d, aimed at the line + 9 mL (lined_lined-measure_s0_dual-fr3.json)	182.5 mL	0.040 mL	success, 10/10

On this goblet the operator's stop lands 2.6–5.3 mL below its aim. The legal window is the line to 10 mL above it, and aimed at the line itself the robot poured a short measure, which the evaluator fails as a UK inspector would. So v6d aims 9 mL above the line. The final run delivered 182.5 mL, with 0.040 mL on the bar (..._spill_origin.json: 0.032 mL of splash out of the glass and one 8 µL drop 17 cm behind it, on the bottle's side). The bottle never touched the goblet, and no limiter acted.



Figure 13. lined-measure, v6d: eight frames of the video (closeup | station_rgb).

11.3 A tasting flight (tasting-flight)

The robot flight was not re-run on the round-2 or round-3 physics: its 50 s rollout is the costliest, and round 3's campaign went to the motion limits and the single-glass operator. Round 0's run is kept in attempts/r0_tasting_tasting-flight_s1_dual-fr3*.json, on the old physics only, as the evidence that one bottle serves three glasses in one rollout: 46.0, 43.0 and 41.5 mL into three ISO glasses, 16.1 mL spilt, four protective stops at full reach. A pour into the first ISO glass of the flight is demonstrated by the careful phantom (section 11.0): 42.9 mL (inside the 50 ± 8 mL window) with 0.088 mL spilt into a 44 mm rim.

11.4 A hand holding a glass (carry-glass, seed 0; e4_robot_pour/service_carry-glass_s0_right_*) demos/robot_glass.py carry. The PURE glass holding 150 mL stands at (0.11, 0.14) m; its bowl (94 mm) is wider than the hand's 80 mm opening, so the right FR3 grips the 8 mm stem from the side, just under the bowl (100 mm above the foot), at 70 N; lifts 5 cm; holds 1 s; carries it 0.28 m to spot 3 in 2.5 s; sets it down upright and lets go (ASSISTED; the glass is held by the pads' friction only).

run (round 3: motion limits, $\mu_t = 3.5$ mm)	turn / slide of the glass in the hand (max)	kept	spilled	evaluator
70 N, 100 mm up the stem	0.56° / 0.92 mm (round 2: 0.51° / 0.95 mm; round 1, condim 3: 2.8° / 4.6 mm)	150.0 mL	0.0 mL	success, 6/6
30 N, 90 mm up the stem	0.53° / 0.77 mm (round 2: 0.96° / 1.56 mm)	150.0 mL	0.0 mL	success, 6/6 (round 1, condim 3: turned 50° in the pinch, 106.8 mL spilt when it fell on release, 1/6)

(Turn and slide relative to the hand, from the recordings: the round-2 evidence audit's held_and_rest_motion.py, e10_legal_stress/rest_motion.json.) The carry is now driven under the FR3 motion limits. With the slosh still over-damped about ninefold (section 10.3), carrying a full glass is easier here than it would be in reality: the qualitative audit's point, which the motion limits address on the robot's side only.

Round 1's 30 N failure was the pads' missing torsional friction. With condim 3 the stem pinch resisted 0.030 N m about the pad line; with the pads' torsional friction from their geometry it resists 0.120 N m at 70 N (measured, $\mu_t = 3.5$ mm) and about 0.05 N m at 30 N (scaled with the grip force) (V6, section 6.3). Held 90 mm up the stem, just under the bowl, the glass-plus-wine centre of mass (~120 mm up) sits almost over the pads, so its weight hardly twists the pinch. The 30 N grip now holds through the lift, the carry and the set-down. So round 1's statement that 30 N is the limit, and its explanation, came from the condim-3 pads and no longer hold. Where a soft or low pinch fails now has not been measured: the carry at lower forces or further down the stem is not in this round's evidence.

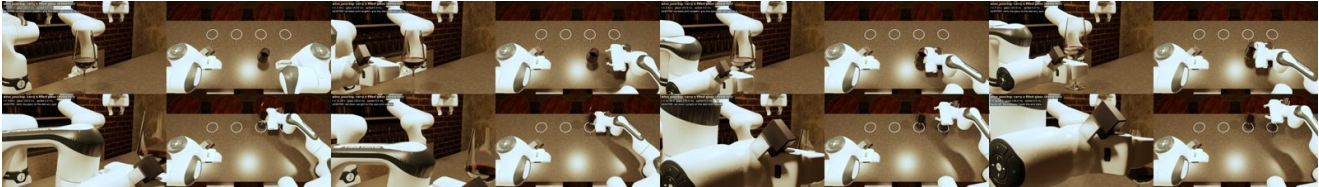


Figure 15b. carry-glass at 70 N: eight frames (guest | station_rgb).

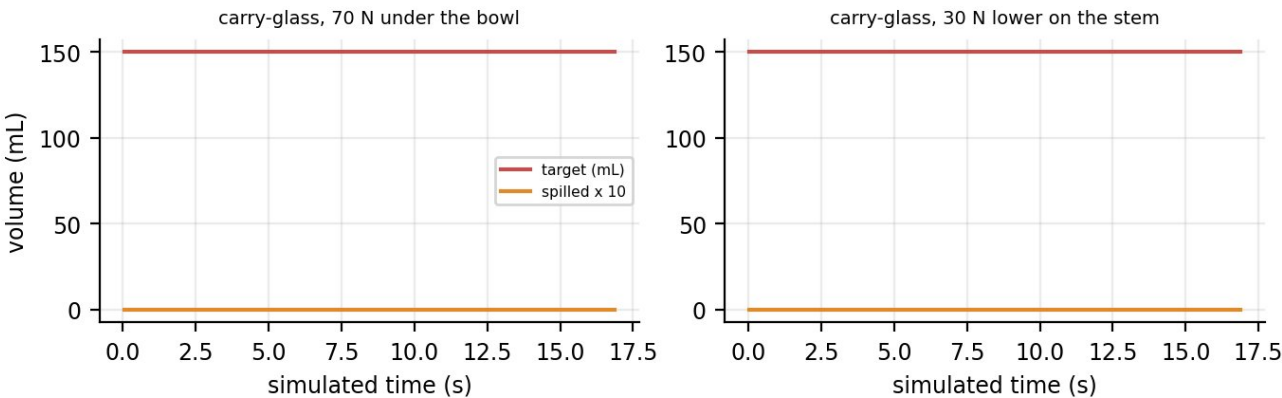


Figure 15c. E4c: the hand-held glass at 70 N and at 30 N: glass contents and spill ($\times 10$) against time.

11.5 A carafe of water by its neck (water-tumbler, seed 0): the hand cannot hold it

The 1 L carafe (584 g, 759 mL of water in seed 0: 1.34 kg) is wider in the body (100 mm) than the hand opens, so the only one-hand grasp is the 58 mm neck, 205 mm up, with the carafe-plus-water centre of mass 77 mm up: 13 cm below the pinch. Tilting it for the pour turns it in the fingers about the pad contact line: by the staging tilt the carafe lagged the hand by 30° (commanded 67°, carafe 37°); with the operator's correction saturated (20°, 50 mm) water started to flow with the carafe 22° short of the commanded tilt, and 0.8 mL had reached the bar and 0.1 mL the tumbler when the run was stopped (e4_robot_pour/attempts/r1_water-tumbler-neck-grasp.log). The arithmetic agrees. At 37° the weight's torque

about the pinch is $1.34 \text{ kg} \times g \times 0.128 \text{ m} \times \sin 37^\circ \approx 1.0 \text{ N m}$. A real pad pair resists rotation about the pad line with about $\mu \text{ N L} / 4$ per pad: 0.25 N m at 70 N for the 17.5 mm pad (section 6.3). The simulated pinch resists 0.265 N m on the carafe's neck (V6, condim 4, round 3; round 2 with the 21 mm pad length: 0.405 N m). That measurement replaces the round-1 estimate of " $\sim 0.5 \text{ N m}$ at most", which the round-2 qualitative audit showed rested on an unmeasured capacity. All are well under the 1.0 N m the full carafe exerts, so a real Franka Hand would fail the same way. The attempt above was made with condim 3, whose capacity was 0.22 N m; with the corrected pads the conclusion stands. The water-tumbler task therefore needs what one parallel gripper does not give: a second hand under the carafe's base or at its body, or a regrasp; it is kept as a task for that reason (two-hand coordination is one of the group's robot capabilities) and is not demonstrated by the robot.

The liquid side of the task, with a phantom hand (E4p,

e4p_phantom_pour/service_water-tumbler_s0_onset40.json, video). The careful phantom pour (section 11.0; 14° staging margin, because the wide-mouthed carafe starts to flow earlier than its geometric onset; tipped to onset at 40°/s) takes the carafe from its place, pours and puts it back where it stood:

- 243.0 mL in the tumbler, inside the 225–275 mL window, with the carafe returned within 3 cm and upright;
- 0.024 mL spilt (0.008 during the pour, 0.016 at the stop): **success, 10/10**;
- the carafe set back down on the bar beside the tumbler with no hop, no cap and no fault.

So water-tumbler is reachable in this world: the robot's limit is its one-hand grasp of the carafe, not the liquid. Tipped at 15°/s instead, the same pour spilt 0.248 mL, 0.208 of it as the onset trickle ran down the carafe's side (service_water-tumbler_s0.json; round 2: 6.3 mL).

Before the Duez closure, the same design spilt 218 mL: half the flow ran back along the carafe's outside and dripped 11–21 cm behind the tumbler (r2_pre_closure/e4p_phantom_pour/service_water-tumbler_s0*.json). In the first build this motion ran only in the rigid world, with an empty carafe, where the lag was $\sim 10^\circ$.

11.6 Dumping a tasting sample (dump-and-pour)

- **By the robot** (demos/robot_glass.py dump; round 2's run archived in r2_final/e4_robot_pour/tasting_dump-and-pour_s0_*; round 1's log attempts/r1_dump-and-pour.log). The arm on the glass's side grips the ISO glass holding the old sample (34.7 mL of red) by its bowl. **It never emptied the glass into the bucket.** The round-3 qualitative audit tracked the bucket in round 2's recording (bucket_track.py), and round 2's account here, "the 34 mL went onto the bar beyond the bucket", was wrong:
 - on its way to the glass, the left arm's forearm pushed the 300 g bucket at about 2.0 s and 3.2–3.8 s, in step with the two collision reflexes the run recorded (2.01 s, 21.2 N m; 3.77 s, 13.3 N m);
 - it dragged the bucket off the robot-side edge of the bar at 6.0–7.4 s, and the bucket lay on the floor by 8.0 s;
 - the glass was tipped from 8.7 s over the place where the bucket had stood, and the 33 mL puddle is centred there.

The physics of a light bucket knocked off a bar is plausible; the account was not, and nobody had looked at the video. In round 3 a rigid-world dry run of the same script, under the motion limits, still moved the bucket 1.17 m (from 2.1 s). With a raised approach over the glass it moved 95 mm (from 4.05 s): the left arm's forearm passes over the bucket in the glass's grasp pose itself. The robot dump was therefore not re-run (another 3 h rollout would show the same), and the task's dump half is shown by the phantom below. The left arm's fresh pour of white wine in round 2's run put 41.4 mL into the ISO glass (goal 50) with 3.3 mL on the bar, on round-2 physics. Moving the bucket or re-planning the dump arm's path is the open work.

- **By a phantom hand** (probes/e7_dump.py, evidence/e7_dump/). A phantom hand lifts the same glass, carries it over the bucket's rim with its foot 2 cm above it, and turns it 150° about its rim in 2 s while lowering it, so that the foot keeps 2 cm above the rim's height and the rim it turns about ends 3 cm over the opening, 3 cm on the far side of the bucket's axis. It holds, turns back and sets the glass down. **All 34.70 mL reached the bucket:** 0.0 mL stayed in the glass, 0.0 mL fell on the bar, the bucket did not move (0.0001 mm), the glass stood upright again (0.001°), and no limiter acted. Two earlier versions failed:
 - round 1 carried the glass with its foot 9.5 cm below the bucket's rim and pushed the 300 g bucket 58 mm across the bar (round-2 evidence audit);
 - this round's first version turned the glass 17 cm above the rim in 1.2 s, and the wine, leaving the rim sideways, landed 15 cm beyond the bucket's centre: all 34.6 mL on the bar (e7_dump/attempts/r2_v1_dump.json).

On round 3's physics the phantom dump again puts all 34.70 mL into the bucket, which moves 0.0002 mm. The physics of the dump works; where the inverted glass is held over the bucket decides it.

- **Pushing a glass over** (the negative control of every task): V3 (section 10.4) and E6 tip (section 12.2).

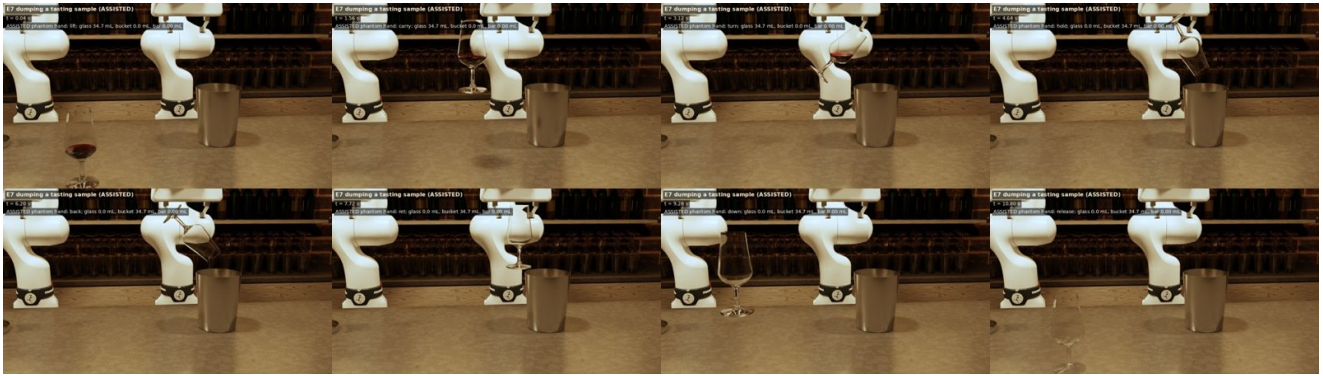


Figure 16. E7: the phantom-held ISO glass emptied into the dump bucket.

12. Stability at rest and under legal abuse

12.1 Rest (E0; evidence/e0_rest/)

30 s, bottle and glass (service/refill-glass seed 0: a bottle with 595 mL on its coaster, a glass with 57 mL at spot 1; 81,560 particles; nothing commanded; round-3 physics):

quantity	after 30 s
liquid centre-of-mass drift	bottle 0.20 mm, glass 0.15 mm
largest body displacement / tilt	bottle 0.016 mm / 0.0012°, coaster 0.007 mm, glass 0.0005 mm / 0.0007°
contents	bottle 595.19 → 595.19 mL, glass 57.29 → 57.29 mL; nothing spilled
particle speed	mean 11–22 mm/s throughout (17.6 mm/s at 30 s), fastest 62–169 mm/s; no trend
limiters, faults	none (no speed cap, no wall repair)

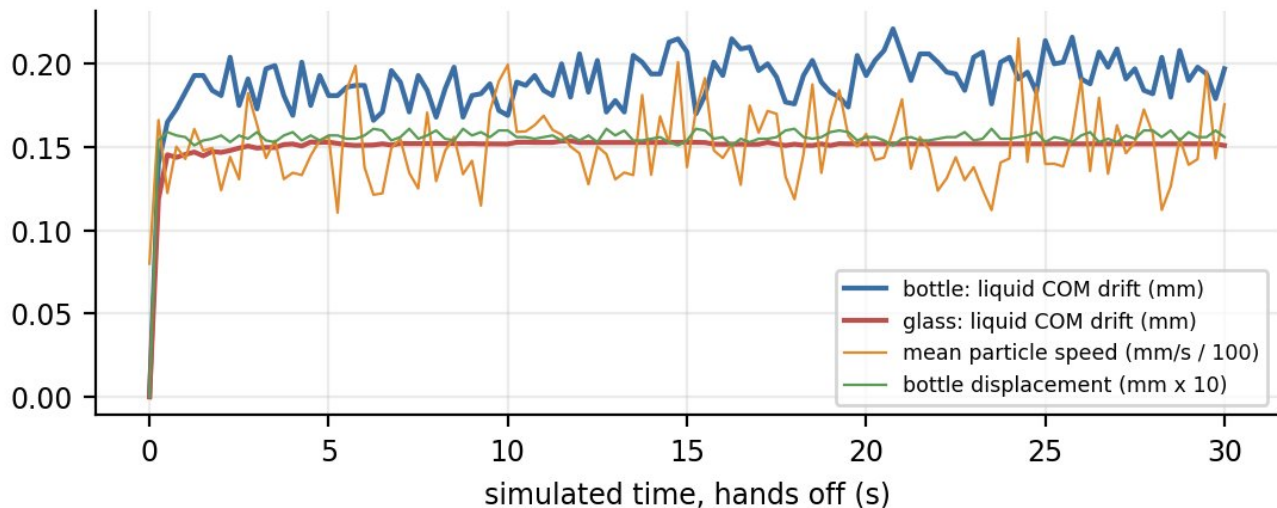


Figure 17. E0 soak: liquid centre-of-mass drift and mean particle speed over 30 s hands off (e0_rest/service_refill-glass_s0.json, video service_refill-glass_s0.mp4).

Nothing moves at the scale of the task (0.2 mm of liquid drift, 16 μ m of bottle motion in 30 s) and nothing accumulates. The particles of a deep column do not come to rest: they jitter at a 11–22 mm/s mean with single particles up to 0.17 m/s, a noise floor of the iterative pressure solve that neither grows nor drifts; it is why the evaluator's "settled" check reads the target glass (0.19 mm/s here) rather than every particle. The reset itself runs preparation (40 packing steps, hydrostatic pressure, 150 ms damped dynamics) and returns at the task's $t = 0$; the first seconds after reset are the free relaxation the series shows.

A 150 mL glass at rest for 10 s (service/carry-glass seed 0, nothing commanded): the liquid's centre of mass moved 0.14 mm, the glass 0.0012 mm and 0.0009°, 150.0 mL stayed in the glass, nothing spilled, mean particle speed 3.3–5.0 mm/s with single particles up to 0.10 m/s (the resting noise floor), no limiter acted.

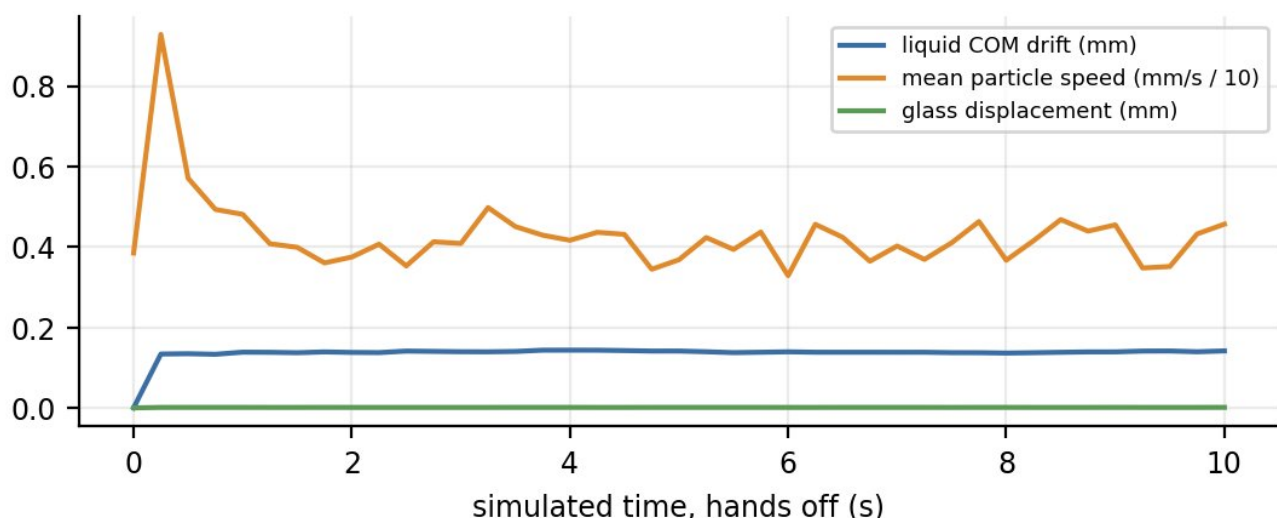


Figure 18. E0: liquid centre-of-mass drift, mean particle speed and glass displacement, hands off.

12.2 Assisted handling and abuse cases (E6; evidence/e6_glass/)

Each case is a phantom fixture on one body (labelled in its video), run on the repaired physics; "limiters" are the counts every evidence file now records.

case	what the fixture does	result (round 3)	limiters
carry_smooth	150 mL glass lifted 4 cm, carried 0.36 m in 1.6 s (smoothstep), set down	0.0 mL spilled	none
carry_jerk	the same 0.36 m in 0.35 s at constant speed (abrupt start and stop)	8.3 mL thrown over the rim (round 2: 10.4 mL)	speed cap in 4 of 4,500 ticks (4 particle-substeps); wall following in 4 ticks; squeeze in 14; no in-wall repair
tip	3 N push at the bowl for 0.15 s	the glass falls; 126.0 mL leaves it, 96.5 mL of that reaches the floor; 24.1 mL stays in the bowl (details below)	speed cap in 253 ticks (3,826 particle-substeps; round 2: 443 ticks, 9,447); 9 particles repaired in one tick; squeeze in 21 ticks; wall following in 12
shake_bottle	full bottle lifted 6 cm and shaken ± 30 mm at 3 Hz in a phantom hand	0.11 mL out of the mouth (round 2: 0.35 mL)	none
drop_bottle	full bottle (658 mL) lifted 5 cm and released onto its coaster	lands upright on the coaster; 0.0 mL leaves the bottle	impact impulse in 10 ticks; wall following in 3; no cap, no repair

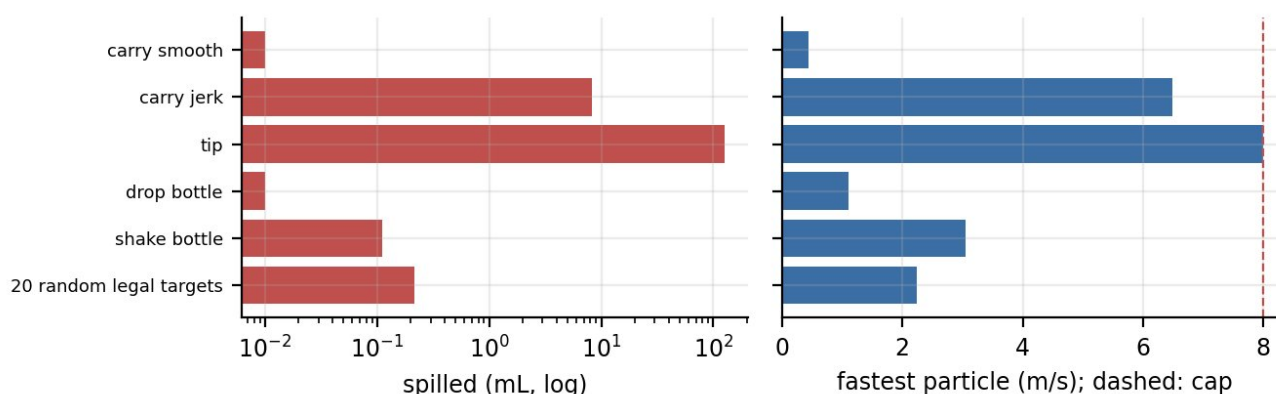


Figure 19. E6 and E10b: liquid that left its container in each assisted handling and abuse case (log scale), and the fastest particle (dashed: the 8 m/s cap).

The tipped glass throws its wine far (far_liquid_tip.json, tip_particles.mp4; the round-3 evidence audit found that this section's round-2 numbers came from a broken analysis: the adopted far_flung.py never computed a speed, and "nothing above 1.35 m" was the height at the first frame past 1 m, not the run's maximum of 1.69 m). probes/far_liquid.py replaces it. On round 3's run:

- the glass's rim centre reaches 2.86 m/s at 0.48 s;
- the wine leaves the rim as the glass strikes the bar, at 4.1–4.7 m/s (frame-to-frame), about 1.5 times the rim's speed, at 10–14 cm above the bar;
- it crosses the bar's edge and falls the 1.09 m to the floor; that fall takes it to 6–7.9 m/s;
- the frames with particles faster than 5 m/s (22 of 112) are all of liquid below the bar top, falling;
- the highest particle is 0.52 m above the bar (1.61 m above the floor);
- up to 76 mL is more than 1 m from where the glass stood, and 18 mL more than 2 m; the farthest particle is 4.3 m away; 96.5 mL ends on the floor.

Wine leaving at ~ 4.5 m/s from 1.1 m lands about 2 m away, which accounts for the liquid beyond 2 m. Faster than the rim it leaves from, it is a jet off the striking bowl. Whether a real 150 mL glass knocked over on a bar throws its wine 2 m is not established here; it is a limitation (section 16). Most of the cap's 253 ticks are in the fall to the floor: the cap then clips free fall, a small distortion. On the floor the liquid breaks into single particles and slides (the qualitative audit's floor atomisation). The surface render cannot show single particles: tip_particles.mp4 draws them as dots.

The dropped bottle is the case the evidence audit found broken (up to 112 particles inside the 9.6 mm base at once, 14–29 mL and a 2–3.5 m jet out of the mouth); with the swept guard and the impact impulse (sections 3.1, 4) the wine stays in the bottle, as an incompressible column stopped by its base does. The abrupt carry is the opposite regime: a 1 m/s velocity step applied to the glass by a stiff phantom spring. In round 1 the guard's in-wall repair acted there, moving out the particles pressed into the glass wall faster than the guard removes them per substep. Wall following (section 4) now carries the liquid with the glass's correction in those ticks, and no repair was needed; the wine thrown out went over the rim.

12.2b Setting a vessel down on spilt liquid (V7, V7b, V7c; evidence/v_wet_setdown*/)

V7: an arrival at almost zero speed (probes/v_wet_setdown.py). This is the round-2 evidence audit's procedure, adopted unchanged:

1. A phantom hand lifts the vessel 3 cm.
2. A one-layer disc of water at rest (a diagnostic state edit) is placed where the vessel stood: on the bar top, or on the coaster's floor for the bottle.
3. The hand lowers the vessel over 1 s to 0.5 mm above its place, holds 0.3 s and releases it.
4. Hands off for 1.5 s.

vessel	puddle	before the squeeze closure (audit, round 2)	round 3
carafe, 758.6 mL water, bar	10 mL	thrown 197–420 mm up, turned over (131–180°); once the run faulted at $2.6 \cdot 10^6$ m/s with both arms' velocity reflexes firing	hops 0.27 mm, tilts 0.001° , at rest; no cap, no repair, 3 impact ticks
PURE glass, 150 mL red, bar	5 mL	thrown up at 1.04 m/s, 94 mm, falls over	0.33 mm, 0.001° , at rest; no cap, no repair
empty PURE glass (210 g), bar	5 mL	leaves the bar at 4.28 m/s, 1.07 m up	0.49 mm, 0.001° , at rest; no cap, no repair
bottle, 658 mL, into its coaster dish	5 mL	stays; 3,927 in-wall repairs	0.26 mm, 0.001° , at rest; no repair
carafe, dry bar (control)	none	stays (0.27 mm)	0.27 mm, 0.001° ; no squeeze (nothing to squeeze), 3 impact ticks

V7b: an arrival at a finite speed (round 3; probes/v_wet_setdown_speed.py, the round-3 evidence audit's wet_setdown_hard.py adopted unchanged in its procedure). V7 and the scripted place() both hover just above the bar and drop. A policy arrives at decimetres per second. The audit kept V7's hand and puddle and changed only the approach: a constant descent at v_down to 1 mm into the bar, so the hand presses the vessel onto it. On round 2's physics, at 5 cm/s the 5 mL puddle left as a ring at about 7 m/s:

- the 8 m/s speed cap fired on 22,118 particle-substeps;
- liquid rose 0.94 m above the bar, and 1.36 mL travelled more than 1 m, the farthest 4.9 m;
- faster arrivals threw it 0.2–1.4 m high.

Round 3 changed two things (section 3.4). Squeezed particles leave the pressure solve and escape along the gap at a bounded Reynolds speed. A particle already inside a wall's clearance is moved out by position rather than pushed at up to 2 m/s. The audit's runs on this code:

run	vessel	puddle	arrival	speed cap	highest rise of the puddle	farthest	puddle within 15 cm at the end	vessel after release
audit, round-2 physics	PURE glass, 150 mL	5 mL, 1 layer	0.05 m/s	22,118	0.94 m	4.9 m	3.17 of 5 mL	at rest
glass150_5mL_v005	same	same	0.050 m/s	0	30 mm	106 mm	5.0 mL	at rest, 0.001°
glass150_5mL_v010	same	same	0.102 m/s	0	27 mm	138 mm	5.0 mL	at rest
glass150_5mL_v020	same	same	0.202 m/s	0	39 mm	147 mm	5.0 mL	at rest
glass150_5mL_2L_v015	same	5 mL, 2 layers	0.152 m/s	0	40 mm	139 mm	5.0 mL	at rest
glass_empty_5mL_v025	empty PURE glass	5 mL	0.253 m/s	0	42 mm	127 mm	5.0 mL	at rest
carafe_10mL_2L_v010	carafe, 759 mL water	10 mL, 2 layers	0.101 m/s	0	40 mm	120 mm	10.0 mL	at rest
carafe_10mL_v020	same	10 mL	0.201 m/s	0	40 mm	100 mm	10.0 mL	at rest
bottle_coaster_5mL_v010	bottle, 658 mL, into its coaster dish	5 mL	0.099 m/s	0	8 mm	51 mm	5.0 mL	on the coaster
glass150_10mL_slide40	PURE glass, 150 mL, then slid 40 mm at 0.1 m/s	10 mL	0.050 m/s	0	40 mm	136 mm	10.0 mL	at rest, 0.15°

In every run the fastest puddle particle stays under 1.6 m/s, and the cap never acts. The squeeze closure acts on 8,900–272,000 particle-substeps while the base comes down. 200–625 particles are found inside the glass's foot (counted as wall repairs; the hand presses the foot 0.1 mm into the bar, so no film remains there) and are moved out. What a real glass set briskly on a wet bar makes, a splash ring of millimetres to centimetres, is what these runs show.

V7c: the robot, through the policy's action path (probes/v_wet_setdown_robot.py, the audit's robot_wet_setdown.py adopted unchanged in its procedure). The scripted operator (ASSISTED) carries the 150 mL glass by its stem at 70 N to 5 cm over its delivery spot. A 5 mL puddle is placed under the foot. Then one policy-style IK joint target puts the foot 3 mm below the bar top, as a GAR cartesian_pose does: since round 3 that target is ramped over the command's duration under the FR3 limits.

command	foot at contact	speed cap	highest rise	farthest	puddle within 15 cm
audit, round-2 physics (servo slew)	0.23 m/s	42,967	1.22 m	4.7 m	— (1.70 mL beyond 1 m)
600 ticks	0.149 m/s	0	39 mm	142 mm	5.0 mL
200 ticks	0.202 m/s (0.33 m/s peak before contact)	0	32 mm	118 mm	5.0 mL
200 ticks, dry bar (control)	0.204 m/s	0 (no squeeze)	—	—	—

The glass keeps its 150 mL and stands upright on its spot in all three. The dry control's carry-glass evaluation passes 5 of 6 checks. settled fails only because the probe reads the scene 1.5 s after the hand opens: the evaluator wants the arm still for 2 s, and under the motion limits the arm's last move runs on past its 200-tick command.

12.3 Random legal actions (E10b; evidence/e10_legal_stress/)

After the scripted pick (side grasp, 70 N; ASSISTED), 20 random joint-target actions go through the GAR action interface exactly as a policy's would (action_codec.action_from_dict → NewtonWinePouringBackend.act: the same validation, motion limits and protective stops). Every joint of the holding arm moves by a uniform ± 0.35 rad within its limits, 0.15 s per action. Then a GAR gripper action opens the fingers wherever the bottle is, and empty holds follow until the scene is at rest or 8 s have passed.

Since round 3 these legal actions are executed the way the real arm would execute them (section 6.2b). A rest-to-rest 0.35 rad move needs at least 62 rad/s² in 0.15 s and the FR3 allows 10. So the arm follows each target as far as its limits allow, and the next action redirects it. The round-3 qualitative audit's point was that round 2's E10 used motions the real arm cannot make. Results (stress.json):

- all 20 actions were accepted, none rejected;
- the collision reflex fired twice (20.02 and 20.01 N m against 20 N m while moving);
- no fault and no non-finite value;
- 0.22 mL of 658 mL left the bottle (round 2: 118 mL), and the bottle's mouth reached 0.69 m/s (round 2: 2.56 m/s);
- limiters: no speed cap (round 2: 240 ticks), the impact impulse in 5 ticks, wall following in 2.

Where the liquid went (far_liquid.json): all of it within 0.43 m of where the bottle stood, at most 0.41 m above the bar, and no frame-to-frame speed above 1.8 m/s.

Round 2's E10 under the unlimited servo is archived (r2_final/e10_legal_stress/), and the round-3 evidence audit recomputed it from the recording. Its report numbers were wrong:

- frame-to-frame speeds reached 7.7 m/s in 31 frames, not "no particle faster than 5 m/s";
- liquid reached 2.85 m high, not 2.5 m;
- 63 mL went more than 1 m and 32 mL more than 2 m, the farthest 4.4 m.

The audit's energy argument made that throw plausible for the motion commanded: the bottle reversed from 1.3 m/s mouth-up to 2.5 m/s mouth-down in 0.16 s. A real FR3 would not have made that motion.

The released bottle comes to rest. It leans at 54.9° in the open fingers with its base 3 cm above the bar. From $t = 10$ s to the end: height range 0.10 mm (SD 0.028 mm, unchanged between the two halves), tilt 54.84–54.98°, drift 0.32 mm (rest_motion.json; the round-2 evidence audit's held_and_rest_motion.py, re-run on this recording). In round 1 the same analysis found a sustained 2.5 Hz oscillation of ± 0.8 mm with 12 mm of drift.

12.4 Robot configurations (evidence/robots/smoke.json)

probes/robots_smoke.py builds the service scene with each catalog alias, holds the ready pose for 0.5 s and moves joint 1 of the first arm by 0.1 rad in a 1 s command:

alias	arms	joints	hold drift	joint 1 moved	reflexes
dual-fr3	left, right	18	$6.3 \cdot 10^{-7}$ rad	0.0914 rad	none
single-fr3	left	9	$1.4 \cdot 10^{-6}$ rad	0.0914 rad	none
dual-panda	left, right	18	$1.9 \cdot 10^{-6}$ rad	0.0917 rad	none
single-panda	left	9	$1.4 \cdot 10^{-6}$ rad	0.0917 rad	none

(Round 3: the 0.1 rad command is now ramped over its 1 s, so at its end the joint is still moving. The 0.0085 rad not yet covered is the PD servo's velocity lag, $k_d v / k_p = 45 \times 0.1 / 600 = 0.0075$ rad, plus the tracker's 0.001 rad. Round 2's servo stepped the reference and read 0.0997–0.1000 rad.)

g1	—	—	refused: "the G1's shoulder is at the 1.09 m bar-top height ..." (section 6.4)		
----	---	---	--	--	--

probes/robots_reach.py (evidence/robots/reach.json) solves, with each configuration's own IK and joint limits, the hand poses the scripted pour uses — the side grasp of the bottle on its coaster, and the bottle's mouth over the glass at every delivery spot tilted 70, 85 and 100° about the lip (the range a 450–750 mL bottle needs, section 10.5): all 16 poses per configuration solve within 2 mm / 2° (in fact within 0.001 mm), for all four configurations, including the single FR3 and the single Panda serving spot 4 (0.88 m from the shoulder). The infrastructure audit ran all three configurations through gar run with liquid (named, keyboard and Cartesian actions at 51–60×). The pours themselves were run with dual-fr3 only; the single arm at spot 4 works near full extension, where the flight's protective stops occurred (section 11.3), so a single-arm pour there is qualified by reach, not by a pour.

13. Runtime

probes/perf.py drives the public WineEnv.act path (what GAR's backend calls) with 0.2 s joint-target actions, returning the policy camera image(s) every action, after a 2-action warm-up, on the default robot, on round 3's physics with the motion limits. They ran after the evidence campaign, one at a time (load average 4.9–5.9). Another creator's jobs shared the GPU at times, and the round-3 lined-measure rerun overlapped the first two.

condition	wall/sim per action (mean / worst)	physics part	per-observation render	startup
pour-glass (82,255 particles), liquid at rest, station_rgb	52.0× / 52.6×	47.0–48.2×	frame 0.57–0.66 s + surface 0.25–0.27 s	36.9 s build + 0.4 s reset
the same, all three policy cameras	56.7× / 57.6×	46.4–49.0×	1.43–1.48 s for the three frames + surface 0.25 s	37.1 s
the same, liquid flowing (a phantom tilts the bottle: 658 → 375 mL during 33 actions), station_rgb	41.5× / 52.6×; the 17 cycles with wine flowing: 33–40×	26.1–48.2×	frame 0.57–0.70 s + surface 0.19–0.32 s; up to 7 liquid substeps per tick	35.0 s
the heaviest scene: tasting/dump-and-pour seed 131 (100,246 particles: a full bottle, two glasses' samples, the bucket), at rest, station_rgb	64.6× / 65.3×	58.1–60.6×	frame 0.60–0.62 s + surface 0.32–0.33 s	35.5 s
through gar run over HTTP (gar/http_transaction.json), pour-glass	see section 14			

The motion generator (section 6.2b) adds forward kinematics to each moving arm's tick; the rest run's physics ratio, 47–48×, is that of round 2 (48.5–50.1×).

Round 2's "up to 86× while pouring" was measured while the campaign's last render worker shared the GPU and before any wine flowed (infrastructure audit, round 3). On an idle GPU the audit measured pouring cheaper per simulated second than rest, and round 3's flow run agrees: 33–40× while the wine flowed. Why the solver is cheaper then was not investigated.

After an idle period. Round 2 released the render worker after 60 s idle. A policy that waited longer paid for a new worker and its first frame inside its next command: the infrastructure audit measured 11.0 s (1,096×) for a 10 ms command, and 13.3 s (66×) for a 0.2 s command. Since round 3 the worker is kept for the life of the session, so the floor holds whatever the policy's pace (section 7). `renderer_options={"idle_release_s": s}` restores a timed release for a host that needs the GPU memory back.

(Round 1 called pour-glass the heaviest scene; the infrastructure audit, round 2, found dump-and-pour seed 131 at 100,246 particles. Its row is now measured.)

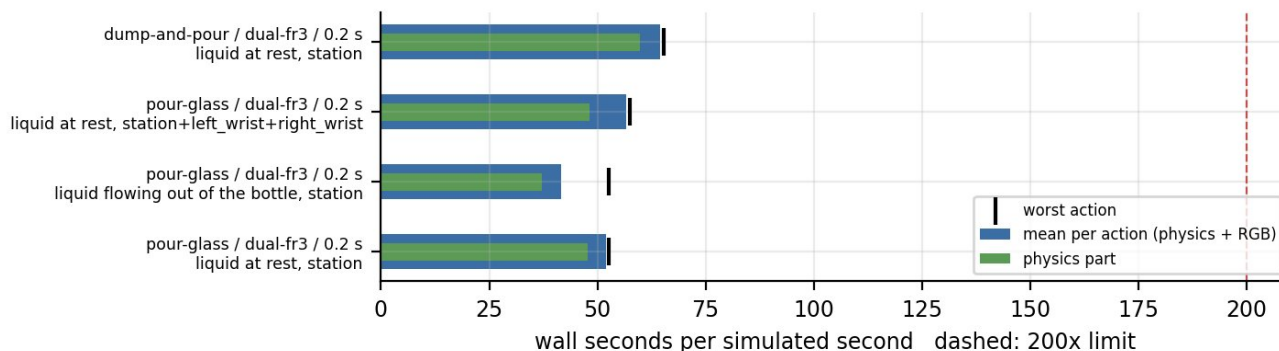


Figure 20. E16: wall seconds per simulated second through the public API, with the 200× limit.

All are under the 200× limit. What the budget depends on:

- **Render cost per observation** is 0.57–0.62 s per 1280×720 RTX frame at 24 samples, plus 0.17–0.38 s to reconstruct the liquid surface (more while a stream is in the air); three cameras cost 1.41–1.45 s together. A 0.2 s action therefore spends 4–8× of its budget on its observation. The shortest policy command scales with the cameras in the observation (`min_command_s` in the backend: the command T must satisfy $70T + 0.6\text{ s} + 0.66\text{ s per camera} \leq 200T$, using the worst measured physics ratio and frame and surface costs). That gives 10 ms with one camera, 15 ms with two and 20 ms with three. The floor is published in the backend's capabilities and checked over HTTP (`min_command_published`, `short_command_refused`). Round 1 used one 10 ms floor for any number of cameras; with three cameras the infrastructure audit (round 2) measured 195–214× for such a command. The round-3 infrastructure audit measured every admitted command at the floor on an idle GPU at $\leq 152.9\times$ in Python and $\leq 166\times$ over HTTP, in the heaviest scene, at rest and while pouring, with 1, 2 or 3 cameras. Owner commands are not floored (section 14).
- **Physics** is 47–61× for a full bottle at rest. It needs 2 substeps per 1 ms at rest and up to 7 while pouring at 50–90 mL/s (24 at the most violent tilts).

- **Concurrency budget.** One full-bottle scene with its policy camera leaves room for about two more GPU-heavy jobs before the 200× limit is reached: at the floor even one other render job pushes a command past 200× (infrastructure audit, round 3). Round 3's evidence campaign ran eight to ten simulations and render workers at once, besides another creator's jobs, and individual runs then went at 180–400× (the robot pours: 6,000–7,300 s wall for 28 s). That is the cost of batch evidence, not of a policy session.
- Rigid-only build and settle take 8–13 s; the liquid preparation (density maps cached per container kind and spacing) 20–30 s; a variation change rebuilds the world in the same session in 30–45 s.

14. Interfaces

GAR backend `gym_anything_robotics/backends/newton_wine_pouring.py` registers wine-pouring through the unified contract (`docs/unified-contract.md`): one live state and clock shared by the Python API, the CLI and HTTP.

- `gar inspect wine-pouring / gar tasks wine-pouring`: presets, tasks and the qualified robots `dual-fr3`, `single-fr3`, `dual-panda`, `single-panda`; task selectors `preset`, `task`, `seed` (the robot is chosen with `gar run --robot`). `--task works` (`gar run wine-pouring --task lined-measure; benchmark.json cli_task_selector`): task names are unique across presets, so a task alone selects its preset. The G1 is refused with its reason when it is selected (`gar run wine-pouring --robot g1; catalog unsupported`).
- Robot setups: an alias, or a `gar.robot-setup.v1` object equal in content to one of the four qualified rigs (`wine_pouring/robot_setup.py`, the catalog's validator); any other setup is refused with what it would need (same arm model from `fr3@1/panda@1`, mounts on the bar pedestal, a Franka Hand per flange, the `barcode_scanning` servo profile). The scene is written for the Franka family (pedestal mounts, ready pose, `libfranka` reflexes, finger pads); `robot_configuration` reports the live handles, setup id and fingerprint.
- `gar run wine-pouring --select preset=service --select task=pour-glass --select seed=0 --json`: a session with the HTTP API and the shared dashboard (`/dashboard`). `POST /v1/variations/select` (or `reset` with options `{preset, task}` and a seed) rebuilds the world for another catalogue variation inside the same session (30–45 s).
- Actions: `named_joint_target` (joint names `left/fr3_joint1 ...`, SI), gripper (half width 0–40 mm per finger, or inward force up to 70 N total), `cartesian_pose` and `cartesian_delta` (per arm, IK on the robot's chain), keyboard (chords W/S A/D R/F ... held ≤ 10 s; they command the robot only); every action states its simulated duration, and arm motion is ramped over it under the FCI limits (section 6.2b); arm targets during a protective stop are refused with the reason in `info`. With cameras, a policy command shorter than the session's floor is refused, and the floor is published (`capabilities.notes.min_command_s` and the policy's `hardware.min_command_s`). It scales with the observation load: $\text{floor} = (0.6 \text{ s} + n_{\text{cameras}} \times 0.66 \text{ s} \times \text{pixels}/(1280 \times 720) \times \text{samples}/24) / (200 - 70)$, at least 10 ms. That gives 10 ms with one camera, 15 ms with two and 20 ms with three. The single 10 ms floor of round 1 admitted 10 ms commands with three cameras that measured 195–214× on an idle GPU (infrastructure audit, round 2). Owner commands are not floored: single-tick stepping is a diagnostic. The shared dashboard's default action is such a 1-tick joint target. It returns in about 1 s, i.e. about 1,000× wall/sim (the round-3 infrastructure audit measured 1,044× for one tick and 255× for five): fine for a diagnostic click, and not a policy's rate.
- Keyboard teleoperation page: `python -m wine_pouring.manual --session <id>` attaches to a live session and serves a local page (the GAR key stays in the proxy's memory; the browser holds a random token) that turns held keys into the same keyboard actions a policy sends, showing the task instruction, the station camera and the arms' stop state. (Until round 2 it read the instruction from the wrong field of `GET /v1/task`, so the line stayed empty.)
- Observations (policy): the session's policy cameras (`station_rgb` by default; `left_wrist_rgb`, `right_wrist_rgb`, `bar_side_rgb` on request, section 7; PNG), joint positions and velocities (the finger joints give the gripper widths; no grip force is exposed), each arm's remaining protective-stop time, the count of refused targets, and the simulated time. The liquid's state, inventories and evaluator are privileged (`/v1/state`, `/v1/evaluation`, owner key).
- Geometry (owner): `/v1/geometry` lists every glassware item (body id, pose, mass, role), the bar-top and floor planes, the delivery spots, the goal and the robot bodies; `/v1/geometry/<item>` returns the item's exact (r, z) wall and liquid-space polygons, rim, fill lines and mass properties; `/v1/geometry/assets/profile/<kind>` the polygons alone.
- Per-command physics traces in GAR's `PhysicsTraceStore` format (`wine_pouring/trace.py`): one sample per 1 ms tick (body poses and velocities, joint state and targets, grip commands, the liquid step's substeps, iterations, guard and limiter counts, the liquid wrench on each glassware item) and events (protective stops, speed cap engaged, wall repairs); read by `GET /v1/history/{id}/trace`, `/trace/events` and `gar export --include-traces`.
- Checkpoints save and restore the whole state, including every particle and the headspace gas. They are kept in memory for the session only and are lost when it stops (the restore answer says "storage": `"in_memory_session"`). A checkpoint restores only into the variation (preset, task, seed, robot) it was saved in. `GET /v1/checkpoints` lists each one with its variation and whether it is restorable; restoring one from another variation is refused (400) with the reason. Since round 3 that refusal, and a reset with an unknown preset or task, are pre-execution rejections: the session stays ready. Until then the shared HTTP service recorded them as runtime faults, and every policy action and evaluation returned 502 until the owner called `POST /v1/recover`; the creator's own dashboard screenshot showed "Owner · not ready" (infrastructure audit, round 3). The fix is in the backend (`ActionRejectedError`) and in the shared restore path, which now exempts pre-execution rejections as reset already did (`test_http_server.py`, `test_wine_pouring.py`). In round 1, a checkpoint saved before a variation change restored into the rebuilt world and left a hybrid: decant-carafe's carafe stood at pour-glass's glass position holding pour-glass's wine (infrastructure audit, round 2).

- `gar doctor wine-pouring` checks `scipy`, `PIL`, `Warp` and a `CUDA` device, the `Isaac` render interpreter, a writable density-map cache (`WINE_POURING_CACHE`), disk for traces, fonts and the pinned dressing assets.
- Tests: `gym_anything_robotics/tests/test_wine_pouring.py`. On the CPU: discovery, robot setups, catalogue, configuration, geometry, the lined window, the task selector and the camera-scaled command floor. With `GAR_GPU_TESTS=1`, on a live scene: the backend's robot configuration, geometry routes, trace contract, action rejections, policy view, variation selection, checkpoint scoping, the published floor, and that a refused restore or reset leaves the HTTP service ready.

`probes/http_transaction.py` runs a real `gar run` session under both principals and records each call's status and timing, screenshots the dashboard as the owner and as the policy, and exports the session (`evidence/gar/`).

Result (`evidence/gar/http_transaction.json`, round 3; all 25 checks pass):

- dashboard page, script, `/v1/session` and `/v1/help`: 200;
- policy reads (capabilities, task, observation, the `station_rgb` PNG): 200;
- policy actions: 200; a 1-tick policy command refused (400, nothing executed); a keyboard chord accepted;
- policy checkpoint save and `/v1/state`: 403;
- owner checkpoint save, an action, then restore: 200, with the tick returned to the saved value and all particles restored;
- `/v1/evaluation`: 200;
- the last action's physics trace (complete) and its events through `/v1/history/{id}/trace` and `/trace/events`;
- robot configuration with setup `pour-fr3` and its fingerprint; the three geometry routes; `/v1/variations`, a variation selection and a variation start;
- a stale checkpoint refused after the variation change, and a reset with an unknown preset refused. Health then still reads ready, and the next policy action is admitted (200), the round-3 fix;
- `gar tasks resolve --task`;
- `gar pause` then `gar export --include-traces` (manifest, history and traces with their SHA-256; the export itself is ignored); the session stopped cleanly.

Wall/sim per policy action: 60× (0.1 s joint target), 50× (0.5 s joint target, gripper and Cartesian step).

The screenshot displays the GAR session dashboard for the policy principal. At the top, it says "GYM ANYTHING / ROBOTICS" and "Session console". A status indicator shows "Restricted policy connected". Below this is a section for API key management, including a "Connect" and "Disconnect" button. The main area is divided into two panels. The left panel, titled "Camera observation", shows a top-down view of a robot arm on a table with a bottle and four numbered spots (1-4). A "Refresh observation" button is present. The right panel, titled "Task & session", contains the task instruction: "Pour the host a tasting measure: 30-45 mL of red wine into the glass at spot 4 (far right); no spills or drips; bottle back on its coaster." Below this is a "Send robot command" section with a text area for "Action JSON" and a "Send action" button. At the bottom, there is a link to "Available actions & capabilities".

Figure 21. The GAR session dashboard (policy principal): task instruction and the policy camera.

15. Failures found and corrected during the build

Each row is a failure observed in a run of this build, its cause as diagnosed, and the change; the lessons file docs/lessons/wine-pouring-v1.md records the reusable ones.

#	observed	cause	change
1	a resting 120 mm column circulated up the wall at ~1 cm/s and never settled	Akinci boundary samples on curved thin walls carried $\pm 50\%$ volume weights	exact-geometry density maps (section 3.2)
2	+10 % liquid volume in a glass	boundary samples on the wall surface instead of dx/2 inside it	map scaled to the dx/2 resting distance
3	750 mL did not fit the 769 mL bottle	a cubic lattice cut by a curved wall leaves a dx/2 gap	wall-conforming ring fill
4	deep columns never converged; a full bottle blew up	no hydrostatic initial pressure; warm start masked to compressed particles; plain Jacobi on signed increments	$\kappa = g \cdot \text{depth}$, unmasked warm start, $\omega = 0.5$ projected Jacobi
5	the first free step of a full bottle pressed 1.6 N on 6.3 N and rang ± 5 N	quasi-static preparation (velocities zeroed) let the zeroing carry the weight	damped dynamic preparation
6	a phantom-lifted full bottle blew up within 25 ms	added-mass instability (0.75 kg liquid in a 0.5 kg bottle)	lumped liquid inertia + filtered residual (section 4)
7	126 kN contacts in one step	rewriting MuJoCo body_iquat at runtime	keep the compiled inertial frame
8	a bottle spun to 5900 rad/s about its axis	axial liquid inertia lumped and cancelled with a lag	lump transverse inertia only
9	surface particles tore off at 250 m/s when the neck sealed	headspace suction applied to the free-surface layer needs tension	pressure drop along the neck liquid
10	an upright full bottle read as sealed with 10 mL of gas	packing deficit near the walls counted as air	seal requires the gas at the interior's highest point
11	particles left the mouth at 3–7 m/s during a gentle lift	Akinci's curvature term (dimensionally inconsistent with the σ -normalised cohesion) and a discontinuous adhesion force on meniscus particles	surface forces off in the first build; rebuilt and on since round 1 (rows 29–30, section 3.4)
12	liquid on the floor under the bar launched upward at 2 m/s per substep	the one-sided bar-top plane treated the floor liquid as 1 m inside the bar	plane limited to its 30 mm slab
13	a splash droplet fell forever past a ± 4 m floor and faulted at 20 m/s	finite floor footprint	unbounded floor
14	one particle at $3 \cdot 10^5$ m/s in one step	DFSPH factor $\alpha \sim 1/r^2$ for two nearly coincident particles	α floor at $\frac{1}{4}$ of the lattice value
15	a toppled glass faulted the episode at 25 m/s	a gap closing between the rim and the bar top squirts	8 m/s squirt cap
16	the gripped bottle did not lift	the robot links' loose hull boxes in the liquid guard reached through the glass into the wine	guard the fingers only (replaced in round 1 by tight convex polytopes, row 28)
17	every tilt of a full bottle tripped the collision reflex (22–50 N m)	the reflex counted the payload and unfiltered grip micro-slip	filtered, payload-aware reflex (section 6.2)
18	the bottle ended at 93° in the hand; joint-speed reflex	the scripted IK switched arm branch mid-move	reject IK branch jumps; side grasp geometry (section 6.3)
19	1.1 kN link 7 – link 5 contact while carrying the bottle	approach along +y folded the FR3 wrist	approach from the bottle's side, tilt toward +y
20	the bottle rotated in the fingers and the stream missed the glass	gripped below its centre of mass	grip 145 mm up the body
21	a phantom-held bottle spun up at 1400 rad/s	phantom damping from the largest moment of inertia was unstable about the slender axis	per-axis inertia-scaled phantom gains
22	a stream rendered as a string of beads; a dark band round every rim	unsmoothed particle centres; rim normals smoothed into the wall	Yu-Turk smoothing, Taubin, crease-aware glass normals
23	the arm drawn as white blobs	one material for every robot mesh	the model's per-part colours
24	the delivery spots a task names were invisible	not drawn	numbered illuminated spot rings (section 2.2)
25	two thirds of every liquid step spent waiting	one host sync per Jacobi iteration on the shared GPU	residual read every 4th iteration
26	137 MB of intermediate video frames committed	a failed render left its frames in the evidence folder	frames ignored and removed (history kept, section 17)
27	a dropped full bottle drove ~120 particles through its 9.6 mm base; a 2–3.5 m jet out of the mouth (round-1 evidence audit)	guard tested walls only within $4 \text{ m/s} \times \Delta t$, pushed $\leq 2 \text{ m/s}$ per substep, cap after integration; the iterative pressure solve cannot stop a 25 cm column in one impact tick; walls moved at the tick's start velocity	swept, exact guard with in-wall repair, cap before the guard; vertical impact impulse handed to the contained liquid with the walls' position correction (sections 3.1, 4)

#	observed	cause	change
2 8	liquid fell straight through the arm and hand (round-1 qualitative audit)	only the finger pads were liquid boundaries	convex polytopes circumscribing every link and the hand (section 3.2)
2 9	a 2 mL spill spread over 1.2 m ² and ran off the bar; the drip check could not fail by a real dribble	no surface tension, wetting or contact line	cohesion, half-space adhesion, contact-line friction on (section 3.4)
3 0	surface forces raised the tumbler's slosh frequency 9 % and its damping to 9.3 %	bulk cohesion pair forces at the rest spacing = the kernel's attraction/repulsion crossover gave the lattice a shear stiffness	cohesion only in the surface layer ($s < 0.9$): +3.1 %, 5.2 % (section 10.3)
3 1	the first robot pour with wetting on spilt 13.5 mL and still running (log kept: e4_r obot_pour/attempts/r1_taper_pour -glass.log)	the operator slowed the stream before the stop: a trickle follows the outside of the lip and runs off the bottle (the teapot effect); the onset trickle, poured 2 cm toward the bottle side from 6 cm up, fell outside the near rim	decisive start (+3° at full ramp after the flow starts), no taper, quick 90°/s stop, mouth 3.5 cm over the rim 1 cm from centre (section 11)
3 2	re-rendered videos lost their probe camera	scene.py binds task_layout at import; a re-render in a process that had built a scene used the original	patch every binding (video.rerender)
3 3	the E6 drop video showed the arms, not the bottle	probe camera on the robot side	probe cameras on the guest side
3 4	E9's "on the outside of the neck" read 7.3 mL before anything was poured	the outer-radius lookup took the largest polygon vertex within 3 mm in height, found none along the straight body and counted liquid on the bottle's axis	edge-crossing outer radius (e2_onset.outer_radius); the two runs made before the fix recounted from their recordings (probes/e9_recount.py; it first picked a robot link as the bottle, caught by a zero reading before the pour)
3 5	the 1 L carafe turned 30° in the fingers before it poured	its weight hangs 13 cm below a neck pinch; two pads resist 0.4 N m about their contact line (V6) against ~1 N m	none in the operator: a one-hand neck pour of this carafe is not possible (section 11.5)
3 6	a carafe set down on 10 mL of spilt water was thrown 197–420 mm and turned over, once faulting at $2.6 \cdot 10^6$ m/s with both arms' velocity reflexes firing; an empty glass on 5 mL left the bar at 4.3 m/s (evidence audit, round 2)	particles trapped in a gap thinner than a particle: the guard pushed each out of the bar at up to 2 m/s per substep, and the descending vessel took that momentum (a pump from the static bar into the vessel)	squeeze closure (section 3.4); V7 (section 12.2)
3 7	wine in a glass lying on the bar felt the bar top through its 1 mm wall	a plane's density map has no notion of an intervening glass wall	planes hidden behind the nearest container wall when it faces the same way
3 8	liquid draping over the bar's front edge just below the top counted as up to 30 mm inside the bar	the bar top was a plane limited to its footprint	the bar top is a box
3 9	the vertical impact impulse acted 525 times while random actions swung the tilted bottle about	the uniform impulse is exact only for a level free surface over a floor	upright containers only (within 30°)
4 0	a glass landing on its bowl reversed its spin in one tick and its walls jumped 2.5 mm into its wine	the liquid steps against walls extrapolated with the tick's starting velocities	wall following (section 4)
4 1	the pinch resisted 0.045 N m about the pad line, a seventh of a real pad's line contact (V6)	Newton's MuJoCo solver gives condim 3, which ignores torsional friction	finger pads condim 4, $\mu_t = \mu L/4 = 4.2$ mm from the pad geometry (section 6.3)
4 2	E7's phantom carried the ISO glass into the bucket and pushed the 300 g bucket 58 mm (evidence audit, round 2)	the carry ran at 8 cm above the bar, with the glass's foot 9.5 cm below the bucket's rim	carried over the rim, foot 2 cm above it, and lowered while turning so the wine leaves 3 cm over the opening (the first fix poured from 17 cm and missed the bucket; section 11.6)
4 3	a checkpoint saved before a variation change restored into the rebuilt world as a hybrid (infrastructure audit, round 2)	checkpoints were not tied to their world	stored with their variation; a mismatched restore is refused (section 14)
4 4	with three cameras, the 10 ms policy commands admitted ran at 195–214x (infrastructure audit, round 2)	one floor for any camera load	floor from the session's camera load, published (section 14)
4 5	the teleop page never showed the instruction	it read task.instruction from GET /v1/task, which returns instruction	fixed (section 14)
4 6	every careful pour lost wine behind the glass (1.4–3.5 mL/s from the bottle, 11 mL/s from the carafe), and the dribble grew from 0.03 to 3.6 mL as the spacing was refined (qualitative audit, round 2: the lip physics unvalidated)	the half-space adhesion acts over a kernel radius and held a steady stream round the lip and back along the neck, where Duez measures clean separation	Duez inertial-capillary separation limit on glassware adhesion (section 3.4, 10.7); probes/spill_origin.py locates where each spilt millilitre fell

#	observed	cause	change
4 7	with the closure the round-1 operator (v5) poured 294 mL onto the bar	its mouth height and offset had been set against a stream that clung to the lip; a clean stream from 5 cm landed beyond the far rim	operator v6c: the careful phantom's geometry, mouth 3.5 cm over the rim (section 11.1)
4 8	the robot's neck brushed the near rim with the mouth 2.5 cm over it (v6, v6b; 11.7–12.3 N for 30–90 ms)	the bottle slipped 7° in the pinch (round 3 found the cause: the unlimited arm accelerations, row 53) and the hand correction lowered the hand 11–15 mm	staging lift and 3.5 cm (v6c); probes/touch_times.py finds when a source touches its target
4 9	the first careful phantom pour spilt 16 mL on the way back	the stop left the bottle 4° below the staged tilt and the return path snapped it up toward the pour	the return starts from wherever the stop left the bottle
5 0	V5's blob over the hand fell through it at free-fall speed in the first round-2 placement	placed 6 cm above link 7's top, which over the hand's middle is inside link 6; the guard pushed it out of link 6 downward	blobs 6 cm above the highest robot surface under them (section 10.8)
5 1	E7's first round-2 version put all 34.6 mL on the bar beyond the bucket	turned 17 cm above the rim in 1.2 s; the wine left the rim sideways	lowered while turning, the rim 3 cm over the opening on the far side of the axis (section 11.6)
5 2	the robot pours' videos were not produced (and the HTTP probe's export failed) in the first campaign of this round	renders cut off by the 9,000 s per-run timeout; gar export refuses an existing directory	record-only robot runs rendered afterwards from their recordings at 960×540; the probe removes its own previous export
5 3	one policy command accelerated the FR3's joints at 27–49 rad/s², the operator's 50 ms waypoints at 71 rad/s², and a held bottle slipped 2.8–7.4° in the pinch (qualitative audit, round 3)	the servo reference slewed at full joint speed with no acceleration or jerk limit	the FCI motion generator (section 6.2b, motion.py); V8
5 4	a glass set down on a puddle at 5 cm/s threw it at the 8 m/s cap, 0.94 m high and 4.9 m across the room; the robot's policy-style set-down did the same (evidence audit, round 3)	squeezed particles kept fast tangential speeds and stayed in the pressure solve, where the two walls' density maps overlap; the guard pushed particles inside its clearance out at up to 2 m/s, pumping them in the foot-bar wedge	squeezed particles leave the pressure solve and escape at a bounded Reynolds speed; position correction inside the clearance (section 3.4); V7b, V7c
5 5	the Duez closure's separation speed moved with the spacing, and the tilt-back dribble grew past the spill tolerance at 1.6 mm (qualitative audit, round 3)	edge radius = film thickness = dx, on every glassware wall	each lip's own edge radius, the local liquid thickness, only near lips (section 3.4, 10.7)
5 6	a refused checkpoint restore, and a reset with an unknown preset, left the session "not ready": every policy action and evaluation returned 502 until the owner ran recovery (infrastructure audit, round 3)	the backend raised plain ValueError; the shared HTTP restore path recorded any ValueError as a runtime fault	ActionRejectedError in the backend; the shared restore path exempts pre-execution rejections as reset did; tests at both levels
5 7	a 10 ms policy command after 60 s idle took 11 s (1,096×) (infrastructure audit, round 3)	the render worker was released after 60 s idle and restarted inside the next command	the worker stays for the session's life (renderer_options.idle_release_s to release)
5 8	REPORT 12.2–12.3 said "no particle faster than 5 m/s" (E10) and "nothing above 1.35 m" (E6 tip) (evidence audit, round 3)	the adopted far_flung.py never computed a speed (it compared len((x, t)) with the particle count); the heights quoted were at the first frame past 1 m	probes/far_liquid.py; the numbers re-stated
5 9	REPORT 11.6 said the robot's dump "missed the bucket"; the bucket had been knocked onto the floor before the dump (qualitative audit, round 3)	nobody watched the video; the account came from the end-state puddle	the account corrected from the recording; the robot dump not re-run (its path still meets the bucket, section 11.6)
6 0	REPORT 10.5 quoted round 2's pre-closure E2 numbers under file names that held the closure runs (qualitative audit, round 3)	the table was not regenerated after the closure was adopted	regenerated from round 3's JSON; archives named
6 1	v6c's neck reached the near rim at the 45° staging tilt once the bottle stopped slipping (5.3 N, 60 ms)	v6c's 3.5 cm had been set with the bottle drooping 7°	v6d: 4.5 cm (section 11.1)
6 2	the lined measure stopped 2.6–5.3 mL below the operator's aim (172.4, 174.75 mL: short measures)	the stop's lead and the goblet's flow at the stop	aim 9 mL above the line (182.5 mL, section 11.2)
6 3	the first decant phantom never stopped (the bottle empty, the carafe 23 mL short of its target)	its stop waited for the carafe to reach the bottle's fill minus 15 mL while 22.8 mL had been spilt	a decant stops on what is left in the bottle

16. Limitations

Each is unfinished work or a modelling choice, stated with its measured consequence.

- **The lip physics rests on a closure** (sections 3.4, 10.7).
 - Without it, the surface-force model's dribble grew with refinement (0.03 / 1.5 / 3.6 mL at 2.5 / 2.0 / 1.6 mm), and a steady stream ran back along the neck, contrary to Duez's measurements.
 - Round 3's closure is grounded in each lip's own edge radius and the local liquid thickness. The steady stream separates cleanly, and the tilt-back dribble is bounded at 0.000 / 0.000 / 0.041 mL across the three spacings.
 - A trickle's film cannot be thinner than one particle, so it separates sooner than a real one and drips from the lip instead of running down the neck. The real teapot dribble at the onset and the last drops is therefore under-predicted, by at most the 0.02–0.1 mL round 2 measured.
 - The onset at 1.6 mm is 2.5° early: static wetting, not converged.
 - The closure's 6 mm zone and its $(U_c/u)^2$ roll-off are authored.
- **The scripted robot pours succeed, but the operator is not robust.** pour-glass with v6d (150.1 mL, 0.136 mL) and lined-measure with v6d (182.5 mL, 0.040 mL) are 10/10 (sections 11.1–11.2). This round took two more versions: the neck on the rim at 3.5 cm once the bottle stopped slipping, and two short lined measures. A policy faces the same margins without the privileged inventory.
- **decant-carafe is not shown to meet its tolerance** (section 11.0). Two phantom decants spilt 22.8 and 57.2 mL as the emptying bottle glugged past horizontal and its pulsing stream overshot or struck the carafe's 52 mm mouth.
- **The carafe cannot be poured with one Franka Hand** (section 11.5). The pinch resists 0.27 N m about the pad line (V6) against ~1 N m, so water-tumbler needs a second hand or a regrasp and is not demonstrated by the robot. Its liquid side is, by a phantom hand: 243 mL in the tumbler, 0.024 mL spilt, 10/10.
- **The robot's dump in dump-and-pour is not shown.** The dump arm's forearm meets the bucket on its way to the glass (section 11.6); the phantom dump puts all 34.7 mL in.
- **The robot tasting flight** has round-0 physics only.
- **Sloshing is over-damped:** 5.2 % of critical at 2 mm and 2.8 % at 1.6 mm, against ≈ 0.6 % (section 10.3). The excess is numerical and falls with resolution. A carried glass calms in 0.8 s instead of ~7 s, so carry-glass is easier than it is, even with the arm's accelerations now limited.
- **A knocked-over glass throws its wine far** (section 12.2). The wine leaves the striking bowl at ~4.5 m/s, 1.5 times the rim's speed, and lands up to 2 m from the bar (18 mL beyond 2 m; the farthest particle 4.3 m away after sliding along the floor). Whether a real glass knocked over on a bar throws its wine that far is not established. The cap's 253 ticks are mostly the fall to the floor. Liquid that hits the floor at several m/s breaks into single particles instead of forming a puddle (the qualitative audit's floor atomisation), a resolution limit of impacts. It is out of every policy camera's view and does not change any score, since every tip-over fails the spill check.
- **The two fingers are independent force-capped drives** (section 6.3). A near-horizontal bottle shifts 2.4–2.6 mm toward the lower finger; the real hand's single motor and rack would keep it centred.
- **Surface forces cost accuracy elsewhere:** the slosh frequency is 3 % above the cylinder mode (1 % without them), and the level deficit is 0.8–1.4 mm (from 0.6–1.2 mm). A puddle's 2.0 mm height is one particle, so a spill is one to two particles thick, and its calibrated contact-line rate leaves a 2 mL puddle on 14.3 cm², 45 % more than the puddle law's 9.9 cm².
- **The headspace closure is calibrated on one of Kenton's angles** ($\beta = 0.80$ at 80°) and checked on the other (section 10.6). It is the right magnitude, not a resolved glug.
- **Level bias of 0.8–1.4 mm low** in every open container (section 10.1): judged by the visible surface, a "to the line" fill delivers 2–5 % more than counted; the lined window 175–185 mL is set by count.
- **Numerical limiters and closures act**, and every evidence file records how often (section 3.4).
- In the careful phantom pours and the robot pours: none in the robot pours of round 3; wall following and the impact impulse in the phantom pours.
- In set-downs on a puddle: the squeeze closure, and 200–625 particles found inside a glass's foot pressed into the bar, moved out (V7b).
- In abuse: the speed cap (tip-over: mostly liquid falling to the floor).
- **Wetting constants are estimates:** $\sigma = 47$ mN/m and $\theta = 20^\circ$ on glass for wine, $\theta = 55^\circ$ on the bar. The sensitivity runs (σ 43/51 mN/m, θ 10/30°; $e9_lip/e9s_*$) measure how the dribble depends on them.
- **Not built:** the lever dispenser (rows 001, 005, 009), condiments/cream, granular side work (rows 006, 018–020, 025), hot liquids as a separate preset, a G1 configuration (section 6.4: decided out of scope in round 3; a table-service preset with a G1 rig would be a new environment, and no humanoid grasp above ~0.5 kg is qualified in this repository).

- **Motion limits, partly:** commands are always accepted and followed as far as the FCI limits allow; libfranka would refuse some command streams outright. TCP jerk is not checked separately.
- **Protective stops** fire when a nearly full bottle is moved fast at full reach (random actions: 20.01–20.02 N m against 20 N m); a stopped arm holding a tilted bottle keeps pouring.
- **Runtime under sharing:** one RTX 5090 runs one full-bottle scene within the 200× limit with its policy cameras (section 13); this round's evidence ran with six to eight simulations and render workers at once, at 180–400×.
- **Cameras** are ideal pinhole RGB (no noise, exposure or depth model).

17. Reproduction and evidence custody

Every evidence file has its producing command in evidence/README.md; evidence/regenerate.sh runs them all in eight parallel lanes and then the interface lane. All commands run from the repository root after source /home/daytona/robotics/environment.sh (Newton 1.5.0, Warp 1.16, MuJoCo 3.11; Isaac Sim 5.1 for RTX). Physics results are deterministic for a given GPU state only up to floating-point reduction order (atomic sums), so a rerun reproduces the numbers to their stated precision, not bit for bit.

```
python -m wine_pouring.probes.v_surface --out wine_pouring/evidence/v_surface/surface.json # V4
python -m wine_pouring.probes.v_level --out wine_pouring/evidence/v_level/level.json # V1
python -m wine_pouring.probes.v_slosh --out wine_pouring/evidence/v_slosh/slosh.json # V2
python -m wine_pouring.probes.v_slosh_ablation --out wine_pouring/evidence/v_slosh/ablation.json # V2b
python -m wine_pouring.probes.v_tip --video --out wine_pouring/evidence/v_tip # V3
python -m wine_pouring.probes.v_robot_boundary --out wine_pouring/evidence/v_robot_boundary/robot_boundary.json
python -m wine_pouring.probes.v_wet_setdown_speed --task carry-glass --item glass --puddle-ml 5 --v-down 0.05 --end-mm -1 --re
python -m wine_pouring.probes.v_wet_setdown_robot --ticks 200 --out wine_pouring/evidence/v_wet_setdown_robot/wet_5mL_200ticks
python wine_pouring/probes/motion_checks/grip_stage_rigid.py wine_pouring/evidence/v_motion/grip_stage.json coarse smooth # V
python -m wine_pouring.probes.e2_onset --fill 750 --rate 2 --video --out wine_pouring/evidence/e2_onset
python -m wine_pouring.probes.e2_onset --fill 600 --rate 2 --return-deg-s 10 --tag e9_return10 --video --out wine_pouring/evid
python -m wine_pouring.probes.e3_kenton --theta 40 --seconds 6 --video --out wine_pouring/evidence/e3_kenton
python -m wine_pouring.probes.e3_kenton --preset service --task pour-glass --seed 0 --out wine_pouring/evidence/e4_robot_pour
python -m wine_pouring.demos.robot_glass carry --force 70 --z-grip 0.10 --out wine_pouring/evidence/e4_robot_pour
python -m wine_pouring.probes.e6_glass --mode drop_bottle --video --out wine_pouring/evidence/e6_glass
python -m wine_pouring.probes.e0_rest --preset service --task refill-glass --seed 0 --seconds 30 --video --out wine_pouring/ev
python -m wine_pouring.probes.evaluator_tests --out wine_pouring/evidence/evaluator/tests.json
python -m wine_pouring.probes.http_transaction --out wine_pouring/evidence/gar/http_transaction.json
python -m wine_pouring.probes.perf --out wine_pouring/evidence/perf/pour-glass_dual-fr3_0.2s_rest.json
python -m wine_pouring.report_pdf && python -m wine_pouring.manifest
```

Custody. Every JSON written in this round carries a provenance stamp made at the moment it was written (wine_pouring/provenance.py): the commit, the sources that differed from it, a digest over the modules that decide what the simulation does (liquid, coupling, gas, world, scene, params, geometry, layout, env, task, inventory, fixtures, kinematics, config, motion, and the shared robot inputs: the backend, robots/catalog.py, the four rigs, the FR3 and Panda models and the Newton robot builder) and the SHA-256 of every source. Since round 3 the digest hashes each Python module's syntax tree without docstrings (physics_code_sha256 per file), so a documentation fix does not change it. robots/catalog.py is shared with other environments, which edited their own entries during round 3's campaign. The manifest therefore accepts a file whose only changed physics source is the catalog, if the catalog's wine-pouring entry at the stamp's commit equals the current one, and says so in the file's note. This is why current evidence carries three digests (b343ff76..., 71fd8483..., eb2c470c...). evidence/MANIFEST.json records every evidence file's SHA-256 and size, and for each JSON its stamp and whether its physics digest equals the code's at manifest time; report/dossier_manifest.json the PDF's and REPORT.md's hashes. Raw recordings of every video (body poses and particle positions at 25 fps, positions quantised to 0.1 mm) are in evidence/**/raw/ (gitignored; kept by the workspace backup to the durable store) and re-render without physics: python -m wine_pouring.video --rec <file> --out <mp4> --cameras closeup,station_rgb (--particles draws isolated particles as dots). The robot demos also write a per-frame series (*_series.json). One past checkpoint (3119c5d95) committed 137 MB of intermediate video frames, removed in the next commit; history was not rewritten because the branch is shared.

18. Sources

Numbering follows research.md §16 ([S1]–[S37]); added: [S40] EN 12464-1 recommended illuminances (search: bar 300 lx, food preparation 500 lx); [S41] Engineering ToolBox, *Friction and friction coefficients for various materials* (engineeringtoolbox.com/friction-coefficients-d_778.html; glass-glass, glass-metal, rubber rows), fetched 2026-10-03; [S42] Franka Robotics, franka_ros franka_gazebo/src/franka_gripper_sim.cpp (HOLDING: $f_d = \text{force_desired} / 2.0$ per finger joint), fetched 2026-10-03; [S43] Duez C., Ybert C., Clanet C., Bocquet L., *Wetting controls separation of inertial flows from solid surfaces*, Phys. Rev. Lett. 104, 084503 (2010), arXiv:0910.3306 (threshold Weber number for dripping, Eq. 2 and Fig. 3b; read 2026-10-04); [S44] Reynolds squeeze film: the time for a disc of radius R under load F to squeeze a film of viscosity μ to thickness h , $t = 3\pi\mu R^4/(2Fh^2)$ (standard lubrication result; any lubrication text). [S45] Franka Robotics, FCI documentation, *Control Interface Specification and Robot Limits* (frankarobotics.github.io/docs/robot_specifications.html: FR3 and FER joint and Cartesian velocity, acceleration and jerk limits), fetched 2026-10-04; libfranka include/franka/rate_limiting.h.

- Bender J., Koschier D., *Divergence-free SPH for incompressible and viscous fluids*, IEEE TVCG 23(3) 2017.
- Koschier D., Bender J., *Density maps for improved SPH boundary handling*, SCA 2017.
- Akinci N., Akinci G., Teschner M., *Versatile surface tension and adhesion for SPH fluids*, ACM TOG 32(6) 2013.
- Monaghan J. J., *On the problem of penetration in particle methods*, JCP 82 (1989) (XSPH).
- Ibrahim R. A., *Liquid Sloshing Dynamics: Theory and Applications*, Cambridge University Press 2005 (cylinder modes, ch. 2; damping, ch. 4).
- Causin P., Gerbeau J.-F., Nobile F., *Added-mass effect in the design of partitioned algorithms for fluid–structure problems*, CMAME 194 (2005).
- Clanet C., Searby G., *On the glug-glug of ideal bottles*, J. Fluid Mech. 510 (2004) [S29].
- Kenton Z. A., Neufeld J. A., Huppert H. E., *Emptying bottles: a study of glugging*, Cambridge CMS report 2012 [S33].
- Antkowiak A., Bremond N., Le Dizès S., Villiermaux E., *Short-term dynamics of a density interface following an impact*, J. Fluid Mech. 577 (2007).
- Zhu Y., Bridson R., *Animating sand as a fluid*, SIGGRAPH 2005; Yu J., Turk G., *Reconstructing surfaces of particle-based fluids using anisotropic kernels*, ACM TOG 32(1) 2013; Taubin G., *A signal processing approach to fair surface design*, SIGGRAPH 1995.
- Franka Robotics libfranka examples/examples_common.cpp (setDefaultBehavior), fetched 2026-10-02.
- The research document's full list: Court of Master Sommeliers service standards [S1, S2], Nicolet College [S3], edu.wine [S4], UK SI 1988/2039 [S5], ISO 3591 [S6–S8], bar dimensions [S9, S10], Saverglass sheets [S11–S14], EN 12726 via Diam [S15], Schott Zwiesel PURE [S16, S17], Riedel Vinum [S18], Makr Shagr Toni [S20], Richtech ADAM [S21], wine properties [S22–S28], bottle emptying [S29–S33], Duez et al. teapot effect [S34], Franka datasheets [S35, S36].