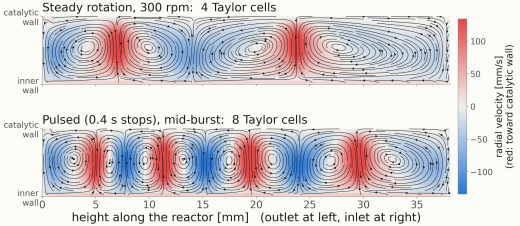
Brief stops, more vortices

Deep reinforcement learning discovers transferable control strategies for catalytic-wall Taylor–Couette reactors



Isaiah N. Bascom[1], Min Wang[1], Yuhe Wang[2]

[1]Department of Mathematics, University of Houston, Houston, Texas
[2]Institute for Scientific Computation, Texas A&M University, College Station, Texas

The Taylor–Couette reactor

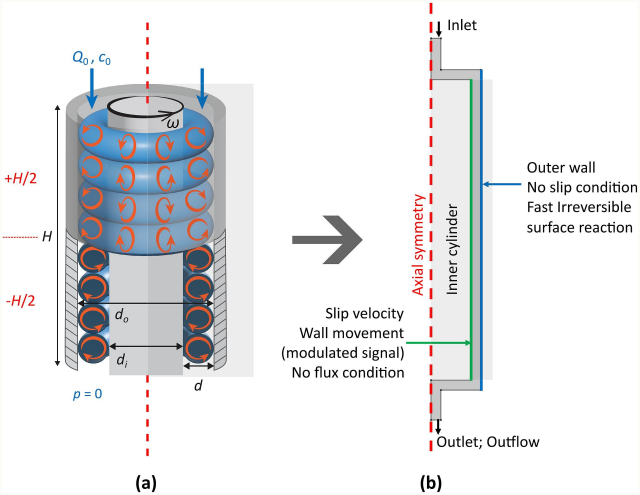


Figure: López-Guajardo et al., *Chem. Eng. J.* 489 (2024) 151174, Fig. 2.

Gap d	6.35 mm
d_i / d_o	50.8 / 63.5 mm
Height H	190.5 mm
$\Gamma = H/d$	30
Feed Q_0	100 mL/min

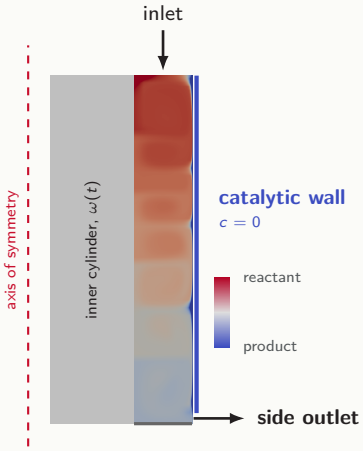
Inner cylinder spins; catalyst on the outer wall.

The reaction happens only at the wall

Taylor vortices carry reactant to the catalytic wall.

Problem

Choose $\omega(t)$ for the most conversion per unit power.



Non-turbulent flow and scalar transport

Equations

Laminar incompressible flow with constant viscosity and no turbulence model

$$\nabla \cdot \mathbf{u} = 0$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u}$$

Scalar-transport equation

$$\frac{\partial c}{\partial t} + \mathbf{u} \cdot \nabla c = D \nabla^2 c$$

Numerics

Finite volume (OpenFOAM), 2D axisymmetric; 3330 cells graded to 25 μm at the catalytic wall

Inner wall: $\omega(t)$
Catalytic wall: $c = 0$
Inlet: $c = c_0$

$\mathrm{Re} \approx 470, \quad \mathrm{Sc} = 1075$

Objective: max $X - P/P_{\text{max}}$ s.t. $\overline{\omega} = 300$ rpm

$X = 1 - \bar{c}_{\text{out}}/c_0$ (conversion), $P =$ motor power

Three ways to search an expensive simulator

Discover: reinforcement learning

TD3 actor–critic. State: conversion and time. Action: burst and stop lengths. Reward: conversion. No physics given.

Check: Bayesian optimization

Gaussian-process surrogate (Matérn-5/2) with expected improvement; 220 CFD runs over fixed waveforms. Replicate noise $\sigma \approx 0.002$.

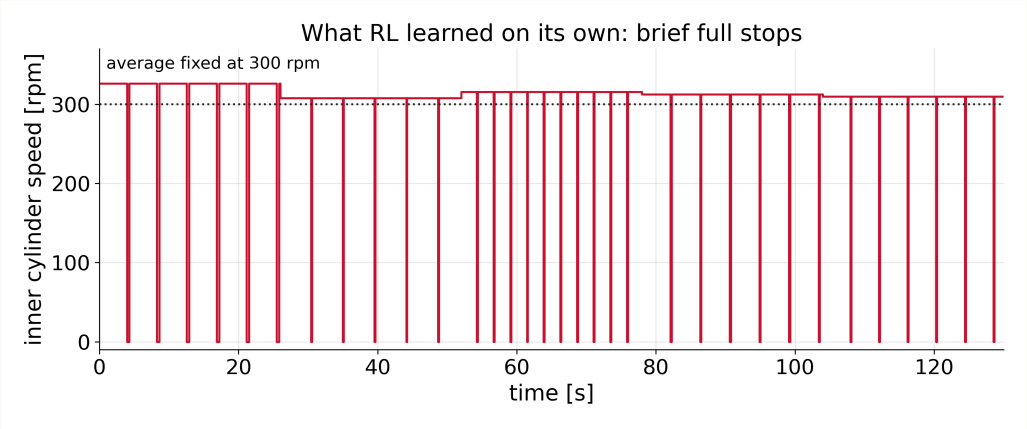
Explain: designed experiments

Pre-registered hypotheses, 62 targeted runs, flow diagnostics logged every 0.02 s.

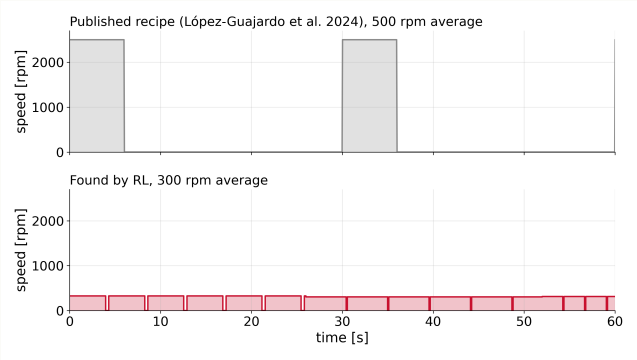
Each 130 s simulation costs about 3–4 CPU-hours, so every method has to be sample-efficient.

Reinforcement learning found brief full stops

All 5 runs: brief full stops (0.1–0.4 s), never long idles.



The opposite of the published recipe



Bayesian optimization

Stop 0.5 s every 5 s

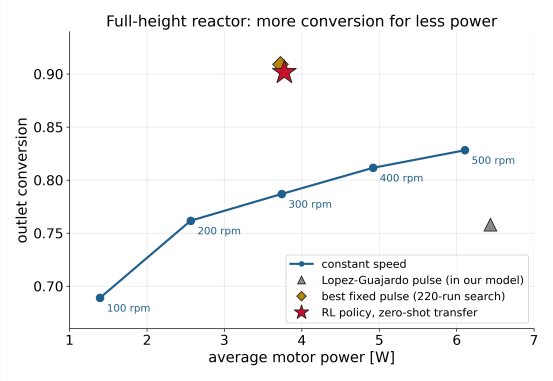
RL

Within 0.01–0.02 of it

Published recipe

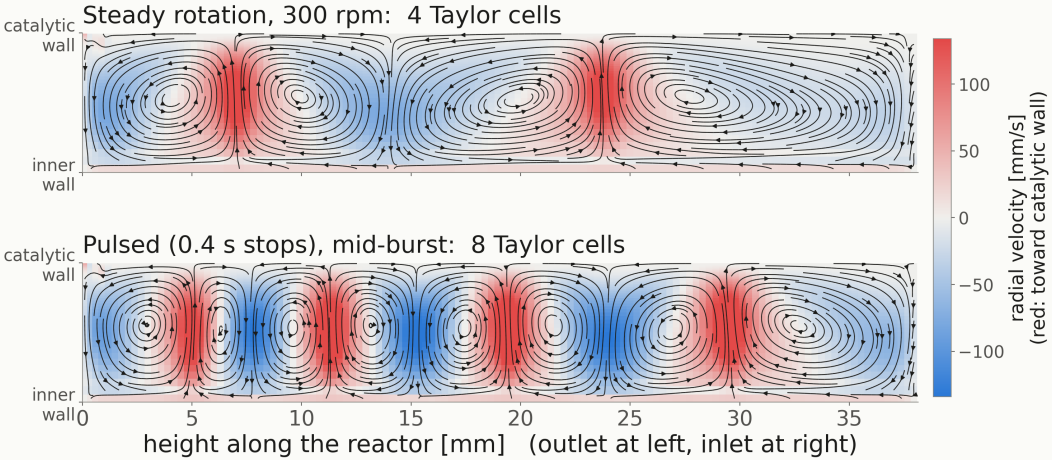
Best in their model, not ours

It transfers to a reactor five times taller



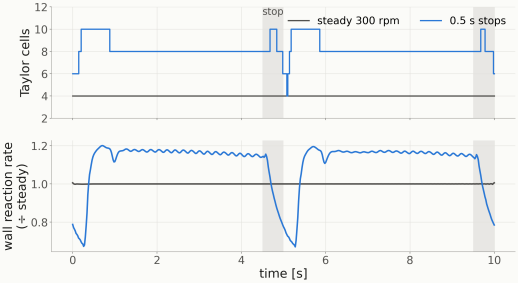
Tall reactor	Conversion
RL, no retraining	0.90
Best fixed waveform	0.91
Steady, same power	0.79
Published, 1.7× power	0.76

Why it works: each restart builds more vortex cells

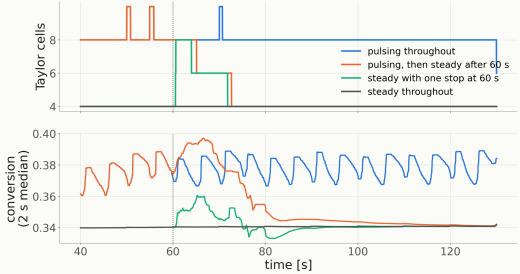


Restarts make the extra cells; steady rotation erases them

Each cycle: 8 cells, +15% wall reaction

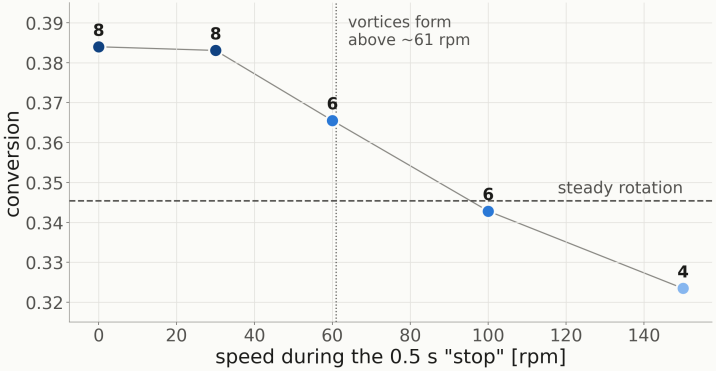


Stop pulsing: back to 4 cells in 10–15 s



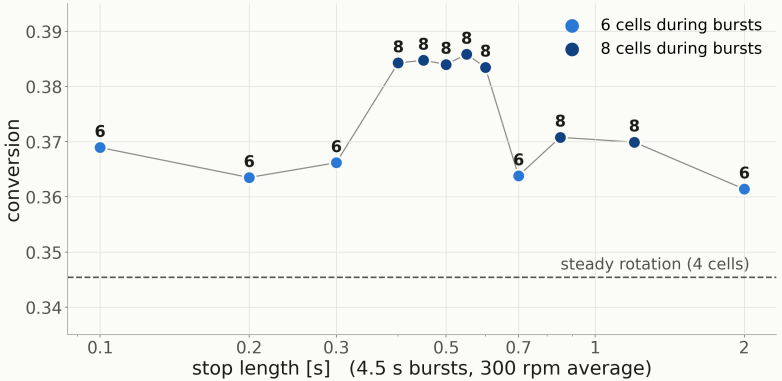
Stops must repeat.

It has to be a full stop



Vortices must die first.

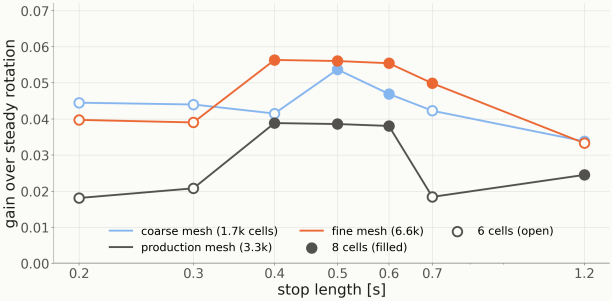
Conversion follows the cell count



+0.04
conversion in the 8-cell window (0.4–0.6 s stops)

4 → 8
cells, steady vs best pulsing

Numerical verification



3 meshes: same cells, gain $+0.04$ to $+0.06$

Checks

- Half time step: outlet artifact gone
- Mass balance: 0.3%
- Replicate noise: 0.002

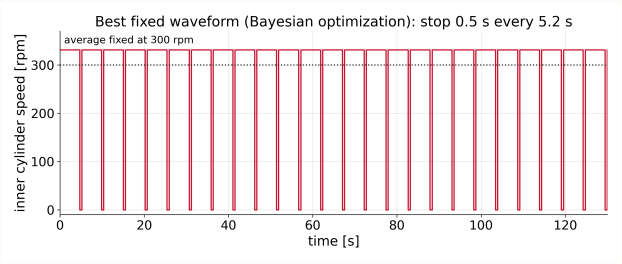
Three takeaways

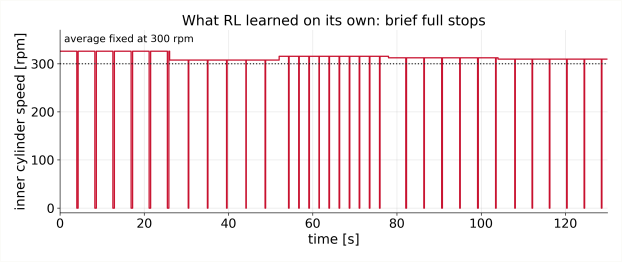
1. Brief full stops beat steady rotation and the published recipe.
2. RL learns near-optimal control: matches 220-run optimization, transfers $5\times$ taller.
3. Mechanism: restarts select more Taylor cells (8 vs 4); stops must repeat.

Next: messy conditions (feed, fluid drift) $\rightarrow$ feedback control; 3D check.

Limits: 2D axisymmetric; $\mathrm{Sc} = 1075$.

Backup: best fixed waveform vs the RL policy





Backup: why 0.5 s? Two time scales, two predictions

Swirl decay

$$\tau = \frac{d^2}{\pi^2 \nu} \approx 0.4 \text{ s}$$

Slowest decay mode of momentum diffusion across the gap. Predicts best stop $\propto \nu^{-1}$.

Wall layer

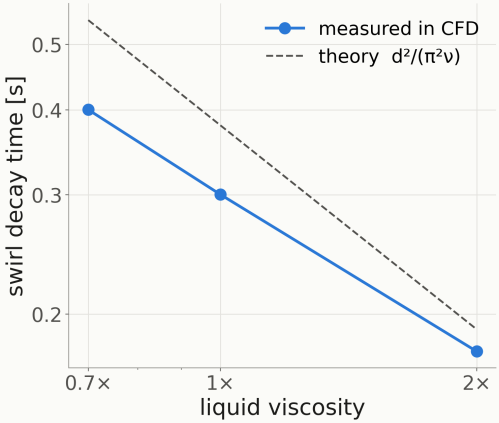
$$\tau = \frac{\delta^2}{D} \approx 0.8 \text{ s}$$

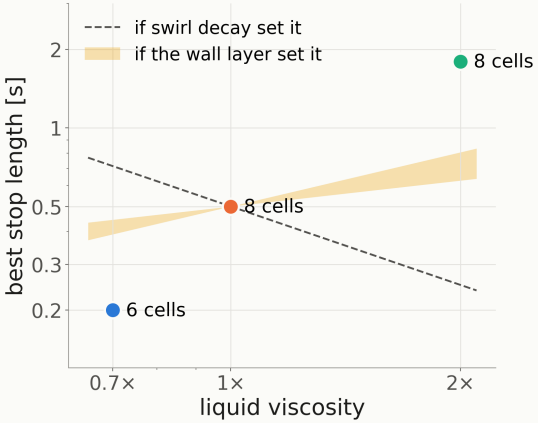
Regrowth of the depleted layer at the wall. At fixed D, predicts best stop $\propto \nu^{0.3}$ to $\nu^{0.7}$.

Test: vary ν at fixed D and fit best stop $\propto \nu^a$. The sign of a decides.

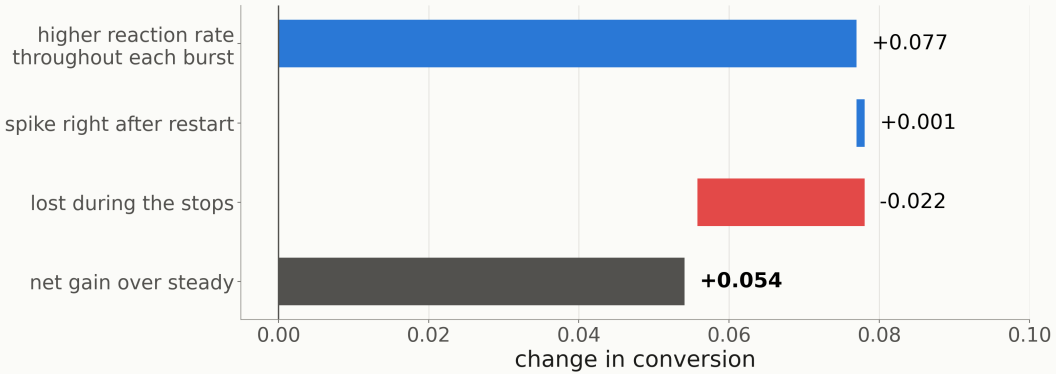
Backup: neither guess sets the best stop

The swirl decays exactly as theory says (left), but the best stop follows neither prediction (right).



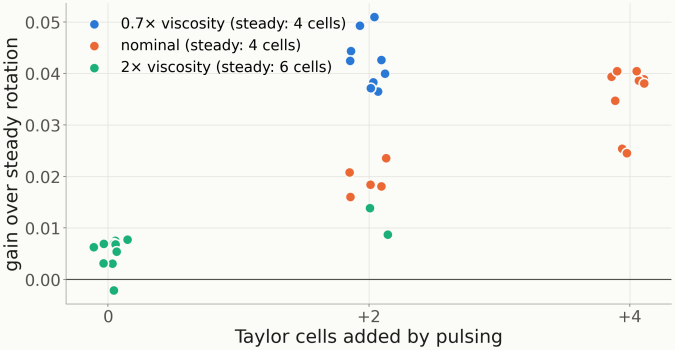


Backup: the gain is a higher plateau, not a restart spike



From the wall-reaction traces of the best waveform's original run (wall-flux budget, last 26 s).

Backup: gain needs cells added over steady rotation



At $2\times$ viscosity, steady rotation already holds 6 cells. Most pulses add none, and the gain shrinks to about $+0.01$.

Backup: electrical power model

López-Guajardo et al. (2024), §3.4, Eqs. 18–23

$$T_{\mathrm{mot}} = \underbrace{J_{\mathrm{tot}}\,\dot{\omega}}_{\text{inertia}} + \underbrace{T_0\,\mathrm{sign}(\omega)}_{\text{bearings}} + \underbrace{\beta_2\,\omega^2\,\mathrm{sign}(\omega) + \beta_1\,\omega}_{\text{drag}}$$

$$i_{\mathrm{mot}} = \frac{T_{\mathrm{mot}}}{K_t}, \qquad e_{\mathrm{mot}} = L\,\frac{di_{\mathrm{mot}}}{dt} + R\,i_{\mathrm{mot}} + K_e\,\omega$$

$$P_{\mathrm{mot}} = e_{\mathrm{mot}}\,i_{\mathrm{mot}}, \qquad P_e = \begin{cases} P_{\mathrm{mot}}/\eta, & P_{\mathrm{mot}} \geq 0 \\ P_{\mathrm{mot}}\,\eta, & P_{\mathrm{mot}} < 0 \end{cases}$$

Evaluated on the commanded $\omega(t)$; reward uses $\overline{P_e}/P_{\max}$, $P_{\max} = 31.94\,\mathrm{W}$ (2500 rpm).

J_{tot}	1.66×10^{-4} kg m^2
T_0	0.1 N m
β_2	7.19×10^{-8} N m s^2
β_1	1.51×10^{-5} N m s
$K_t = K_e$	93.1 mN m/A
L, R	0.28 mH, 0.178 Ω
η	0.90

At 300 rpm the bearing term alone is ≈ 3.5 of the 3.7 W, so power is set by the mean speed.