Algorithmic Limits in Optimal Control: A Computational and Statistical Study

Faranak Rajabi

Department of Computer Science,
University of California, Santa Barbara

M.S. Thesis Defense
June 2026

Presentation Overview

Limit-Cycle Oscillators: Self-Sustained Rhythms

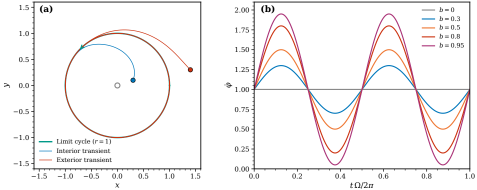
Definition: A closed loop in state space that trajectories settle onto.

Contrast with a pendulum:

- Pendulum: infinitely many amplitudes
- Limit cycle: *one preferred amplitude*; perturb it, it returns

Models:

- Neurons firing
- Heartbeats
- Circadian rhythms
- Power-grid inverters



(a) Stuart–Landau trajectories spiral onto the unit-circle limit cycle.

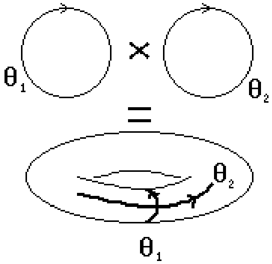
(b) Angular velocity $\dot{\phi}$ on the cycle, averaging to $\Omega < \omega$ for $b > 0$.

Coupled Oscillators: Phase Relationships

Two oscillators that influence each other.
Each alone settles onto its own cycle.

Three possible joint behaviors:

- **In-phase synchrony:** both peak at the same time
- **Anti-phase:** one peaks while the other troughs
- **Desynchronized:** no fixed phase relationship



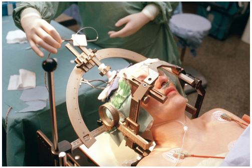
Two phase circles $S^1 \times S^1$ form the 2-torus $\mathbb{T}^2$: the natural state space for the two-oscillator phase pair.

The control problem (this thesis)

Start near in-phase.
Steer to anti-phase.
Minimize control energy.

On anti-phase synchronization in oscillator networks: Joseph, Pakrashi (2020) *Sci. Rep.* 10:10178.

Coupled Oscillator Control Has Real Stakes



DBS / Parkinson's

Power Grid Inverters

Circadian Rhythm

- ***Parkinson's:*** pathological neural synchrony causes tremor; desynchronization is therapeutic
- ***Power grids:*** grid-forming inverters as coupled oscillators; desynchronization causes blackouts
- ***Circadian rhythms:*** coupled neural oscillators set the body's internal clock

Diffusive Coupling: Interaction Proportional to Difference

Form (substituted into Stuart–Landau dynamics)

$$\dot{x}_i = F_x(x_i, y_i) + c(x_j - x_i) + u_i$$
$$\dot{y}_i = F_y(x_i, y_i)$$

F_x, F_y: intrinsic limit-cycle dynamics; c: coupling strength.

Same form as heat flow: hot to cold, proportional to temperature *difference*.

Two key properties:

- If $x_1 = x_2$: coupling vanishes (*promotes synchronization*)
- If $x_i \to 0$: coupling on oscillator i vanishes (*the trapdoor we will exploit*)

We Study Stuart–Landau-Like Oscillators

Stuart–Landau: the universal normal form for systems near a Hopf bifurcation (the canonical way a fixed point gives birth to a limit cycle).

Polar form

$$\dot{r} = ar(1 - r^2), \qquad \dot{\phi} = \omega + br^2 \sin 2\phi$$

- $\dot{r} = 0$ at $r = 1$: unit-circle limit cycle
- b controls *nonisochronicity* (amplitude-dependent rotation rate)
- $\sin 2\phi$ modification makes anti-phase a stable attractor (small basin $\Rightarrow$ nontrivial control)

Two coupled oscillators in Cartesian form

$$\begin{aligned}\dot{x}_i &= ax_i - \omega y_i - ax_i^3 - (a + 2b)x_i y_i^2 + c(x_j - x_i) + u_i \\ \dot{y}_i &= \omega x_i + ay_i - ay_i^3 + (2b - a)x_i^2 y_i\end{aligned}$$

$i \in \{1, 2\}$, $|u_i| \leq u_{\max}$. Canonical: $a = \omega = 1$, $b = 0.8$, $u_{\max} = 1$.

Optimal Anti-Phase Control: The Cost

Energy-optimal anti-phase control

$$J(z_0) = \int_0^{T_f} R(u_1^2 + u_2^2)\, dt + \frac{\gamma}{2} d^2(z(T_f))$$

Terminal cost targets the anti-phase orbit:

$$d^2(z) = (x_1 + x_2)^2 + (y_1 + y_2)^2 + (r_1 - 1)^2 + (r_2 - 1)^2$$

- Vanishes for any anti-phase pair at $r = 1$ (phase-free target)
- $(r_i - 1)^2$ anchors the radii: blocks the trivial $z = 0$ solution

Why anti-phase?

- Maximally desynchronized state (clinical relevance for Parkinson's)
- Real attractor in our system, but with a small basin (nontrivial control)

Settings: $R = 10$, $\gamma = 1000$, $T_f = 2T$.

Two Approaches: Direct Optimization vs Dynamic Programming

Direct approach (control-space)

Parameterize $u(t)$, e.g. piecewise-constant on N intervals:

$$u^* = \arg \min_{\mathbf{u} \in [-u_{\max}, u_{\max}]^N} J(\mathbf{u}; z_0)$$

$\Rightarrow$ finite-dim optimization. Gives *one* trajectory from z_0.
Risk: local minima.

Dynamic programming (state-space)

Define the *value function*:

$$\mathcal{V}(z,t) = \min_{u(\cdot)} \int_t^{T_f} R(u_1^2 + u_2^2)\, ds + \tfrac{\gamma}{2} d^2(z(T_f))$$

$\Rightarrow$ infinite-dim PDE on full state space. Gives feedback law $u^*(z,t)$ *everywhere*.

The Hamilton–Jacobi–Bellman Equation

The value function satisfies a first-order nonlinear PDE. Switch to backward time $\tau = T_f - t$:

HJB equation

$$\frac{\partial \mathcal{V}}{\partial \tau} + H(z, \nabla_z \mathcal{V}) = 0, \qquad \mathcal{V}(z, 0) = \tfrac{\gamma}{2} d^2(z)$$

with Hamiltonian (drift + optimized control):

$$H(z, p) = -f(z) \cdot p + H^*(p_{x_1}) + H^*(p_{x_2})$$

$$H^*(p) = \begin{cases} p^2/(4R), & |p| \leq 2Ru_{\max} \\ -Ru_{\max}^2 + u_{\max}|p|, & \text{otherwise} \end{cases}$$

Optimal feedback in closed form:

$$u_i^*(z, t) = \mathrm{sat}(-\partial_{x_i} \mathcal{V}(z, t)/2R,\ u_{\max})$$

No local-minimum escape hatch: the gradient of $\mathcal{V}$ *is* the policy.

Four Dimensions Is the Feasibility Ceiling

Two 2D oscillators $\Rightarrow$ ***4-dimensional state space***.

At $G = 64$ grid points per axis:

$$64^4 \approx 1.7 \times 10^7 \text{ grid points}$$

Single A100 GPU (40 GB)	$\sim$5 min
Comparable CPU	10–20 hours (est.)
$d = 6$ at same G	6.9×10^{10} pts (infeasible)

Two coupled oscillators is the largest system where grid-based HJB gives a definitive answer.

Solver: WENO5 + Local Lax–Friedrichs + SSP-RK3

Spatial discretization:

- **WENO5**: 5th-order weighted essentially non-oscillatory scheme
  - High order in smooth regions, no oscillations at kinks
- **LLF**: local Lax–Friedrichs numerical Hamiltonian
  - Monotone, consistent, Lipschitz

Time integration:

- **SSP-RK3**: strong-stability-preserving 3rd-order Runge–Kutta
- Floquet exponent $\lambda = -2a$ sets the natural time scale

GPU implementation:

- Extends CASL-HJX from 2D to 4D, $\sim$7,100 lines C++/CUDA
- One thread per grid point, coalesced memory access

Jiang, Shu (1996); Shu, Osher (1988); Barles, Souganidis (1991).

Solver Pseudocode: One SSP-RK3 Step on the GPU

Initialization ($\tau = 0$)

$\mathcal{V}^0(z) \leftarrow \frac{\gamma}{2}\, d^2(z)$ on grid G^4, $G = 64$

For each backward step $\tau_n \to \tau_{n+1}$ (CFL-limited $\Delta\tau$)

SSP-RK3 stages, in parallel over all G^4 grid points:

$$\begin{aligned}
\mathcal{V}^{(1)} &= \mathcal{V}^n + \Delta\tau\,\mathcal{L}(\mathcal{V}^n)\\
\mathcal{V}^{(2)} &= \tfrac{3}{4}\mathcal{V}^n + \tfrac{1}{4}\mathcal{V}^{(1)} + \tfrac{1}{4}\Delta\tau\,\mathcal{L}(\mathcal{V}^{(1)})\\
\mathcal{V}^{n+1} &= \tfrac{1}{3}\mathcal{V}^n + \tfrac{2}{3}\mathcal{V}^{(2)} + \tfrac{2}{3}\Delta\tau\,\mathcal{L}(\mathcal{V}^{(2)})
\end{aligned}$$

Spatial operator $\mathcal{L}(\mathcal{V}) = -H_{\mathrm{LLF}}(z, \nabla^{\mathrm{WENO5}}\mathcal{V})$
Each grid point: WENO5 reconstruction on each face $\to$ LLF flux $\to$ RHS.

One CUDA thread per grid point. Memory coalesced over x_1-axis.
Production resolution: $\sim$3,000 backward steps, $\sim$5 min on A100.

The Solver Is Verified Three Ways

Barles–Souganidis (1991)

Any numerical scheme converges to the viscosity solution of HJB if it is ***monotone***, ***consistent***, and ***stable***.
Our scheme satisfies all three.

Empirical verification (each test does a different job):

- **Linear advection** (4D, periodic BCs): measured order 4.84–4.99
  $\rightarrow$ confirms *fifth-order* accuracy (smooth solution, order matters)
- **LQR vs analytical Riccati**: relative error 6.3×10^{-5} at $G = 64$
  $\rightarrow$ end-to-end pipeline check against an analytical control solution
  not a fifth-order test: $\mathcal{V}$ is quadratic, any high-order scheme captures it
- **Stuart–Landau grid convergence**: cost variation $< 0.5\%$ for $G \in \{64, 80, 96\}$
  $\rightarrow$ confirms convergence on the actual problem

The Solver: `casl-hjx4d-gpu-accelerated`

Implementation:

- ~7,100 lines C++/CUDA
- Extends CASL-HJX from 2D to 4D
- Templated solver: one Hamiltonian device function per problem
- Four boundary conditions supported

Five validated problems:

- Coupled Stuart–Landau (main result)
- Coupled Van der Pol (validation)
- 4D LQR (vs analytical Riccati)
- Linear advection (WENO5 convergence)
- Hodgkin–Huxley (exploratory)

Memory and timing on A100:

G	VRAM	Wall time
32	0.08 GB	30 s
64	1.3 GB	5 min
96	6.8 GB	25 min
128	21 GB	OOM 16 GB

Add a new problem in ~300 lines:

1. Copy any `casl_gpu_*.cu`
2. Replace `H()` with your Hamiltonian
3. Replace `maxAbsH*()` wave speeds
4. Write a main with sweep + BCs

github.com/faranakR/casl-hjx4d-gpu-accelerated

HJB Reveals Two Qualitatively Different Strategies

The optimal mechanism depends sharply on coupling strength c.

Weak coupling ($c = 0.10$)

- Stays near limit cycle
- Small periodic nudges
- Gently steers phase to $180°$
- $r_{1,\min} > 0.85$

Conventional wisdom: confirmed.

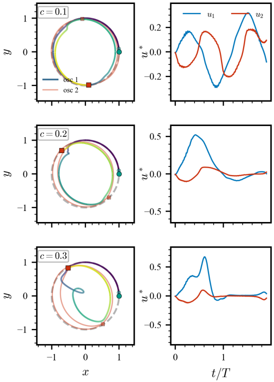
Strong coupling ($c = 0.30$)

- Drives oscillator 1 to origin
- $r_{1,\min} = 0.11$
- Holds briefly
- Lets it rebuild into anti-phase

*We call this **amplitude collapse**.*

c is dimensionless (we set $a = \omega = 1$). “Weak/strong” refers to which basin holds the global optimum, with crossover at $c^* = 0.227$ (next slides).

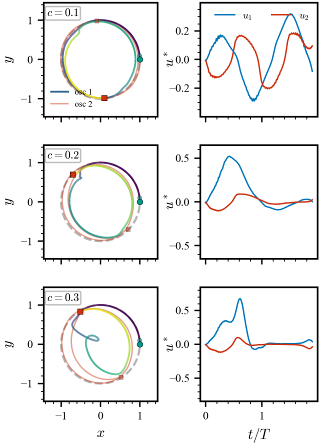
Trajectories and Controls at Three Coupling Strengths



Left column: phase portraits. Right column: optimal control signals.
At $c = 0.30$ oscillator 1 spirals to the origin; controls go bang-bang then quiescent.

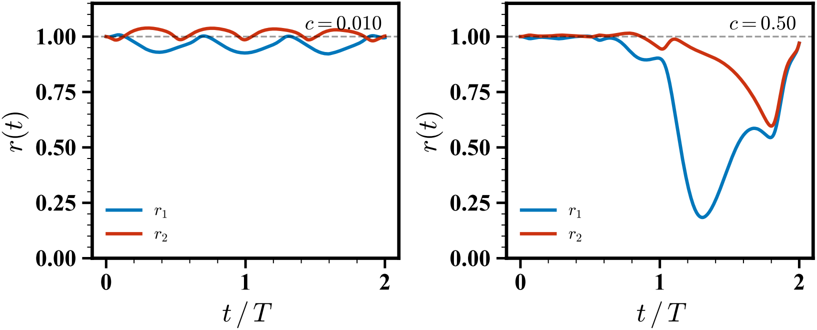
Bang-bang: from Pontryagin's maximum principle. When H is linear in u on the constraint set $[-u_{\max}, u_{\max}]$, the optimal u saturates at the boundary, switching sharply. Here: pulse during collapse, $u \approx 0$ during hold (origin: natural dynamics + vanished coupling do the work), pulse

HJB Dashboard: Live Animation at $c = 0.30$



Coupled Stuart-Landau, $c = 0.30$.

Amplitude Time Histories Tell the Same Story



At $c = 0.30$*:* $r_1(t)$ *collapses to* ~ 0.1*, holds, rebuilds.*
$r_2(t)$ *stays near unity throughout.*

Time-unit reference (canonical $a = \omega = 1$, $b = 0.8$**):** natural period $T = 2\pi/\Omega \approx 10.47$, horizon $T_f = 2T \approx 20.94$, Floquet exponent $\lambda = -2a = -2$ per time unit (rebuild $0.11 \to 0.99$ takes $\sim \ln(89)/2 \approx 1.1$ time units, matches figure).

The Transition Is Sharp: $c^* = 0.227 \pm 0.003$

Dense sweep over 30 coupling values:

- At $c = 0.2247$: $r_{1,\min} = 0.728$ (on-cycle)
- At $c = 0.2299$: $r_{1,\min} = 0.097$ (collapse)
- Transition window: $\Delta c < 0.006$

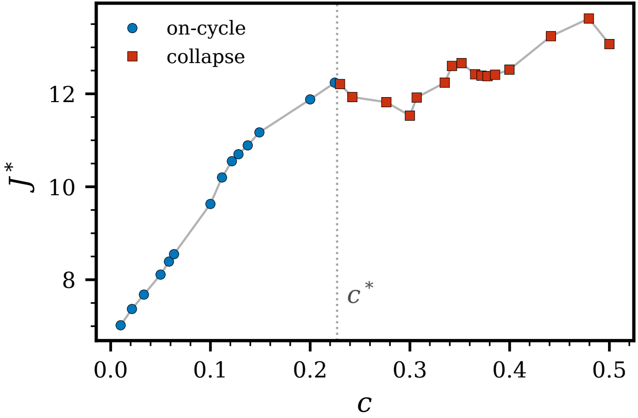
The cost landscape has two disconnected basins

Think of all possible control trajectories as a landscape; cost is height.

- One basin: “stay near cycle, nudge”
- Other basin: “collapse, hold, rebuild”
- Separated by a cost *ridge*: intermediate strategies cost *more*

At c^* the collapse basin becomes deeper.
The optimizer jumps; it doesn’t interpolate.

Non-Monotonic Cost: The Diagnostic Signature



$J^*(c)$ peaks at $c \approx 0.22$, then ***decreases*** *despite stronger coupling.*
Only possible if the controller switched to a more efficient mechanism.

Two Geometric Properties Make Collapse Work

Polar dynamics: $\dot{\phi} = \omega + br^2 \sin 2\phi$.

(1) Coupling elimination at the origin:

- Diffusive coupling on oscillator 1 is $c(x_2 - x_1)$
- At $x_1 \approx 0$: oscillator 1 stops fighting on the cycle

(2) Free phase advance (phase-velocity surplus):

- At $r \approx 0$: $\dot{\phi} = \omega$ (bare frequency)
- On cycle: $\dot{\phi}$ averages to $\Omega = \sqrt{\omega^2 - b^2} < \omega$
- Free phase per unit time: $\delta\omega = \omega - \Omega = b^2/(\omega + \Omega)$
- At $b = 0.8$: $\delta\omega = 0.4$

Hold cost $\sim Rc^2\tau/2$ vs on-cycle cost $\sim Rc^2 T_f$.
Brief hold + free phase wins.

The Free-Phase Story Checks Out Quantitatively

Sweeping nonisochronicity b at fixed $c = 0.30$:

b	τ (s)	$\delta\omega\tau$	$\omega\tau$	$\Omega\tau$	J^*
0.0	4.68	0°	268°	268°	19.93
0.2	4.57	5°	262°	257°	18.10
0.4	4.65	22°	266°	244°	15.48
0.6	5.03	58°	288°	230°	14.16
0.8	6.00	138°	344°	206°	11.53
0.9		*(no collapse)*			19.76

- Predicted $\delta\omega\tau$ matches measured value to $< 1^\circ$
- Optimizer *extends* hold time ($4.7 \to 6.0$ s) as return grows
- 42% cost reduction $b = 0 \to b = 0.8$ comes from 138° of free phase

Collapse Occurs Across the (c, b) Plane

b	c=0.05	c=0.10	c=0.20	c=0.30	c=0.40	c=0.50
0.0	0.928	0.885	0.806	**0.181**	**0.218**	**0.234**
0.2	0.908	0.857	**0.112**	**0.191**	**0.219**	**0.223**
0.4	0.900	0.858	**0.113**	**0.190**	**0.207**	**0.196**
0.6	0.901	0.862	**0.102**	**0.177**	**0.165**	**0.272**
0.8	0.897	0.857	0.773	**0.110**	**0.131**	**0.184**
0.9	0.870	0.845	0.801	0.761	0.337	0.374

$r_{1,\min}$. *Bold: collapse.*

- Collapse occurs at all b, including isochronous $b = 0$
- Energetic advantage scales with $\delta\omega \propto b^2$
- $b = 0.9$: long period $T = 14.4$ gives on-cycle steering enough time

L-BFGS-B: The Standard Gradient-Based Optimizer

Limited-memory Broyden–Fletcher–Goldfarb–Shanno with Box constraints

Workhorse for smooth, bounded nonlinear optimization.

Quasi-Newton update:

$$\mathbf{u}^{k+1} = \mathbf{u}^k - H_k^{-1} \nabla J(\mathbf{u}^k), \qquad |u_i| \leq u_{\max}$$

- **Limited-memory:** stores only recent gradient updates, not full Hessian
- **Box:** handles bound constraints $|u| \leq u_{\max}$
- Superlinear local convergence near a minimum

Standard for smooth, well-behaved cost landscapes.

Byrd, Lu, Nocedal, Zhu (1995) *SIAM J. Sci. Comput.* 16:1190.

CMA-ES: The Standard Global Stochastic Optimizer

Covariance Matrix Adaptation Evolution Strategy

Derivative-free, population-based. Specifically designed to escape local optima.

Each generation

- Sample candidates from $\mathcal{N}(\mathbf{m}_k,\, \sigma_k^2 C_k)$
- Evaluate cost on each
- Update mean $\mathbf{m}$, step size σ, covariance C based on best samples

Covariance C *learns the shape* of the local cost landscape.

If anything could find a hidden global basin, it's this.

Hansen, Ostermeier (2001) *Evol. Comput.* 9:159.

Both Direct Optimizers Get Stuck On-Cycle

Same problem at $c = 0.30$, 100-interval piecewise-constant $u(t)$

Method	J^*	$r_{1,\min}$	Evals	Strategy
HJB ($G{=}64$)	**11.53**	**0.110**	–	**collapse**
L-BFGS-B (zero init)	21.23	0.746	–	on-cycle
L-BFGS-B (warm start)	15.75	0.715	–	on-cycle
L-BFGS-B (5 random)	27.50	0.695	–	on-cycle
CMA-ES (400D)	24.24	0.964	25k	on-cycle
CMA-ES (100D)	24.75	0.991	196k	on-cycle

- L-BFGS-B from any of 5 initializations: on-cycle
- CMA-ES with 1.96×10^5 evals: on-cycle, $r_{1,\min} = 0.99$
- Failure is ***structural***: cost ridge separates the basins

Phase Reduction: The Domain-Specific Standard Tool

Idea: when weakly perturbed, collapse the full (x, y) state to a single phase variable ϕ.

Why this works:

- Amplitude decays fast back to the cycle (Floquet exponent $-2a$)
- Only phase has neutral dynamics the controller can manipulate

Reduced equation

$$\dot{\phi}_i = \omega + Z_x(\phi_i) \cdot u_i(t)$$

$Z_x(\phi)$: *infinitesimal phase response curve* (iPRC).

- **Uncoupled PRC baseline:** design control independently per oscillator
- **Coupled PRC baseline:** include interaction function $H(\psi)$

The price: amplitude is enslaved to phase.
Off-cycle moves are not representable.

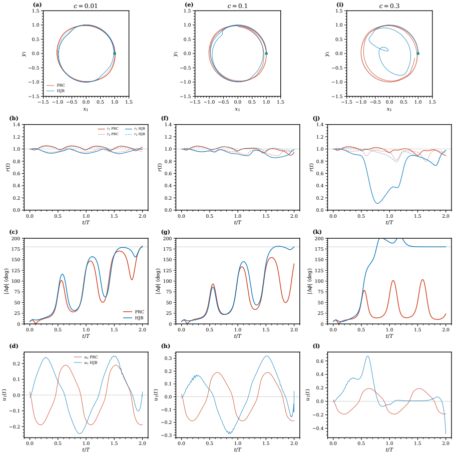
Phase Reduction Fails at the Same Boundary

	$\Delta\phi(T_f)$			J_u		
c	Uncoup.	Coupled	HJB	Uncoup.	Coupled	HJB
0.01	178°	175°	180°	7.47	8.14	7.02
0.10	140°	179°	180°	7.47	10.95	9.63
0.20	49°	172°	180°	7.47	15.07	11.88
0.30	24°	171°	180°	7.47	20.01	11.53
0.50	11°	151°	180°	7.47	30.54	13.07

- Uncoupled iPRC: collapses by $c = 0.10$
- Coupled phase reduction: saturates at $\sim 171^\circ$, 1.7–$2.3\times$ HJB cost

*For strong coupling, phase reduction **cannot reach the target** because the optimal mechanism isn't on the cycle.*

Phase Reduction Cannot Represent the Off-Cycle Trajectory



Weak coupling: the methods agree. Strong coupling:
HJB leaves the cycle, phase reduction structurally cannot.

Van der Pol Is a Controlled Falsification Test

Reasonable objection: "Maybe collapse only happens because Stuart–Landau has special algebraic structure."

Van der Pol oscillator:

- Classic non-Stuart–Landau limit-cycle oscillator
- Limit cycle is *egg-shaped*, not circular
- Not a Hopf normal form
- Crucially: *isochronous*, $\delta\omega = 0$

Why this is the right test

Shares property (1): limit cycle + diffusive coupling
Lacks property (2): no nonisochronicity, no phase-velocity surplus

Separates the two claims cleanly.

Van der Pol parameters: $\mu = 1$, effective amplitude ≈ 2 in x (≈ 2.7 in y, egg shape), period $T \approx 6.66$, $u_{\max} = 0.5$. Off-cycle deviations reported next slide are absolute, not normalized.

Van der Pol Confirms the Mechanism

c	J^*	Off-cycle 1	Off-cycle 2	$\Delta\phi$
0.05	25.27	0.52	0.51	180.0°
0.10	12.09	0.39	0.39	179.9°
0.20	27.96	0.62	0.80	179.5°
0.25	37.76	1.09	0.86	180.0°
0.30	48.18	1.20	0.91	179.5°

Two findings, both predicted:

- Significant off-cycle excursions at strong coupling
  - Off-cycle strategy *generalizes* beyond Stuart–Landau (property 1)
- Cost monotonic in c (no non-monotonicity)
  - No free phase, no energetic advantage (property 2 absent)

The Mechanism Distilled: Two Geometric Prerequisites

(1) Coupling elimination

Must be a state where the coupling force vanishes.
For diffusive coupling, that state is $r_i \approx 0$.
Makes collapse possible.

(2) Phase-velocity surplus

Phase velocity at the origin must exceed the on-cycle average.
Makes collapse energetically advantageous.

Van der Pol confirms (1) without (2): collapse happens, no cost benefit.
Same conclusion at $b = 0$ for Stuart–Landau.

Thesis Contributions

Computational artifact:

- GPU-accelerated 4D HJB solver (extends CASL-HJX), $\sim$7,100 LoC C++/CUDA
- Full HJB solutions in $\sim$5 min on a single A100 ($G = 64^4$, 17M grid points)
- Verified: 5th-order WENO convergence, LQR vs Riccati at 10^{-5}, grid convergence

Scientific finding:

- Sharp critical coupling $c^* = 0.227 \pm 0.003$ where mechanism changes
- Amplitude collapse: globally optimal above c^*, missed by every standard method
- Quantitative mechanism: $\delta\omega = b^2/(\omega + \Omega)$ predicts free-phase budget to $< 1^\circ$
- Generalizes: confirmed on Van der Pol (different limit cycle, no nonisochronicity)

Code: github.com/UCSB-CASL/CASL-HJX

The Broader Argument

Choosing a reduced model up front and optimizing within it can *qualitatively* miss the global optimum.
The optimal mechanism lies in a region the reduced model cannot represent.

Why this matters for computational science:

- “Local optimum” assumes the reduced model contains the global one
- Population-based search has the same blind spot if basins are disconnected
- Full-state methods are the only verification when both methodologies fail

For coupled oscillator control under strong coupling,
verify against full-state HJB before trusting reduced-order results.

Limitations and Future Work

Honest limitations:

- Grid-based HJB doesn't scale beyond $\sim$4D
- Three coupled oscillators (6D) is infeasible with current hardware
- Single oscillator family treated in depth

Future directions:

- Higher-N via tensor-train decomposition or neural HJB approximation
- Stochastic HJB extension (in progress; IMEX-ADI + cuSPARSE)
- Conformal prediction for calibrated method selection
- Experimental validation (electrochemical, in-vitro neural)
- Clinical translation: reduced DBS duty cycle

Acknowledgments: Dedication

To my ***parents***, who have never been able to attend my defenses, and who will not have access to the internet to attend this one. I know you are the ones who wanted to be here most.

Matin Khademi

*In memory of **Matin Khademi**,*
and for the 50000 innocent Iranian people massacred in two days.

WE WILL NOT FORGET, WE WILL NOT FORGIVE.

Thank You

Questions? Comments?