

Supplementary Information

Boundary-damped wave dynamics enables local equilibrium propagation with finite-time gradient guarantees

Dino Vlahsek^{1*}, Dijana Oreški¹ and Alen Kišić^{2,3}

¹Faculty of Organization and Informatics, University of Zagreb, Varaždin, Croatia

²Solektra d.o.o, Čakovec, Croatia

³VERN University, Zagreb, Croatia

*Corresponding author: dvlahsek@foi.hr

Proofs of Theorems 1, 2, 4 and 5

Theorem numbering and symbols follow the main article, while supplementary equations use the S-prefix. For convenience, the potential and the dynamics are restated here. At nudging strength β ,

$$V_{\theta,\beta}(q) = E_{\theta}(q; x) + \beta C(q, y), \quad (\text{S1})$$

and the state follows

$$M\ddot{q} + B^{\top}RB\dot{q} + \nabla_q V_{\theta,\beta}(q) = 0, \quad (\text{S2})$$

where $M \in \mathbb{R}^{n \times n}$ is symmetric positive definite, $B \in \mathbb{R}^{m \times n}$ extracts the m selected boundary velocities, and $R \in \mathbb{R}^{m \times m}$ is symmetric and positive definite. Equations (S1) and (S2) are equations (1) and (2) of the main article.

Theorem 1 (Boundary-flux identity)

Statement. For every classical solution of equation (S2), the augmented energy

$$\mathcal{H}_{\beta}(q, \dot{q}) = \frac{1}{2} \dot{q}^{\top} M \dot{q} + V_{\theta,\beta}(q) \quad (\text{S3})$$

satisfies

$$\frac{d\mathcal{H}_{\beta}}{dt} = -\dot{q}^{\top} B^{\top}RB\dot{q} = -\left\|R^{1/2}B\dot{q}\right\|^2 \leq 0. \quad (\text{S4})$$

Proof. Differentiate $\mathcal{H}_{\beta}$ and insert equation (S2):

$$\dot{\mathcal{H}}_{\beta} = \dot{q}^{\top} M \ddot{q} + \nabla_q V_{\theta,\beta}^{\top} \dot{q} \quad (\text{S5})$$

$$= \dot{q}^{\top} \left(-B^{\top}RB\dot{q} - \nabla_q V_{\theta,\beta} \right) + \nabla_q V_{\theta,\beta}^{\top} \dot{q}, \quad (\text{S6})$$

which gives equation (S4). $\square$

Theorem 2 (No-dark-mode condition)

Statement. Let q_β satisfy $\nabla_q V_{\theta,\beta}(q_\beta) = 0$ and let $K_\beta = \nabla_{qq}^2 V_{\theta,\beta}(q_\beta) \succ 0$. The equilibrium of the linearized system $M\ddot{u} + B^\top R B \dot{u} + K_\beta u = 0$ is exponentially stable if and only if there is no nonzero generalized normal mode v such that

$$K_\beta v = \omega^2 M v, \quad B v = 0. \quad (\text{S7})$$

Proof. Let $D = B^\top R B \succeq 0$. A mode on the imaginary axis has the form $u(t) = v \exp(i\omega t)$ and therefore satisfies

$$(-\omega^2 M + i\omega D + K_\beta) v = 0.$$

Multiplying by v^* and taking the imaginary part gives $\omega v^* D v = 0$. Since $K_\beta \succ 0$ excludes a nonzero zero-frequency mode and $R \succ 0$, an imaginary-axis mode must satisfy $B v = 0$. Its real part then gives $K_\beta v = \omega^2 M v$. Conversely, every nonzero v that satisfies these two equations generates an undamped periodic solution. The finite-dimensional dissipative system has no spectrum in the open right half-plane. Excluding all imaginary-axis modes is therefore equivalent to exponential stability. $\square$

Theorem 4 (Centered equilibrium gradient)

Statement. Assume that E_θ and C are sufficiently smooth and that $\nabla_{qq}^2 E_\theta(q_0; x)$ is nonsingular, where $q_0(\theta)$ obeys $\nabla_q E_\theta(q_0; x) = 0$ and $\mathcal{L}(\theta) = C(q_0(\theta), y)$. The local equilibrium branch q_β of equation (S1) satisfies

$$\nabla_\theta \mathcal{L} = \left. \frac{d}{d\beta} \nabla_\theta E_\theta(q_\beta; x) \right|_{\beta=0}. \quad (\text{S8})$$

Consequently, the symmetric estimator

$$g_\beta = \frac{\nabla_\theta E_\theta(q_{+\beta}; x) - \nabla_\theta E_\theta(q_{-\beta}; x)}{2\beta} \quad (\text{S9})$$

obeys $\|g_\beta - \nabla_\theta \mathcal{L}\| \leq M_3 \beta^2 / 6$ when the third derivative of $\nabla_\theta E_\theta(q_\beta; x)$ with respect to β is bounded by M_3 .

Proof. Differentiate the stationary equation $\nabla_q E_\theta(q_\beta; x) + \beta \nabla_q C(q_\beta, y) = 0$ at $\beta = 0$ and use the implicit-function theorem. Let

$$H_0 = \nabla_{qq}^2 E_\theta(q_0; x), \quad G_0 = \nabla_{q\theta}^2 E_\theta(q_0; x), \quad c_0 = \nabla_q C(q_0, y).$$

Differentiation of the stationary equation with respect to β gives

$$\left. \frac{dq_\beta}{d\beta} \right|_{\beta=0} = -H_0^{-1} c_0.$$

Consequently,

$$\left. \frac{d}{d\beta} \nabla_\theta E_\theta(q_\beta; x) \right|_{\beta=0} = -G_0^\top H_0^{-1} c_0.$$

Implicit differentiation of $\nabla_q E_\theta(q_0; x) = 0$ with respect to θ gives

$$\frac{dq_0}{d\theta} = -H_0^{-1} G_0,$$

and therefore

$$\nabla_{\theta}\mathcal{L} = \left(\frac{dq_0}{d\theta}\right)^{\top} c_0 = -G_0^{\top} H_0^{-1} c_0.$$

The two expressions coincide by symmetry of mixed derivatives, proving equation (S8). A centered Taylor expansion gives the stated remainder. $\square$

Theorem 5 (Finite-time certificate)

Statement. Let $\hat{q}_{\pm\beta}(T)$ be states produced by finite physical relaxation and let

$$g_{\beta,T} = \frac{\nabla_{\theta}E_{\theta}(\hat{q}_{+\beta}(T);x) - \nabla_{\theta}E_{\theta}(\hat{q}_{-\beta}(T);x)}{2\beta}.$$

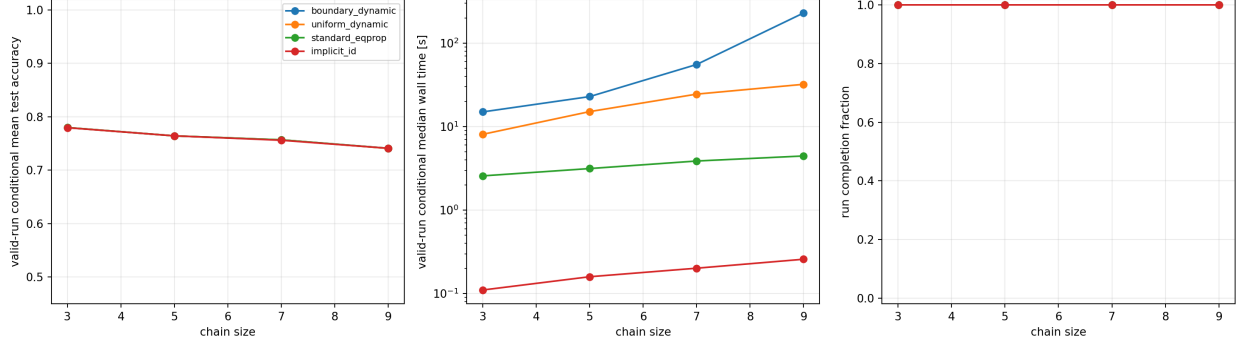
Suppose $q \mapsto \nabla_{\theta}E_{\theta}(q;x)$ is L_{θ} -Lipschitz. If $\varepsilon_{\pm}(T) = \|\hat{q}_{\pm\beta}(T) - q_{\pm\beta}\|$, then

$$\|g_{\beta,T} - \nabla_{\theta}\mathcal{L}\| \leq \frac{M_3}{6}\beta^2 + \frac{L_{\theta}}{2|\beta|}(\varepsilon_{+}(T) + \varepsilon_{-}(T)). \quad (\text{S10})$$

If both nudged potentials are μ -strongly convex, define the final force residuals by $r_{\pm} = \|\nabla_q V_{\theta,\pm\beta}(\hat{q}_{\pm\beta})\|$. Then

$$\|g_{\beta,T} - \nabla_{\theta}\mathcal{L}\| \leq \frac{M_3}{6}\beta^2 + \frac{L_{\theta}}{2\mu|\beta|}(r_{+} + r_{-}). \quad (\text{S11})$$

Proof. Add and subtract the exact centered estimator g_{β} , apply its Taylor bound, and use Lipschitz continuity for the two endpoint differences. Strong convexity gives $\|\hat{q} - q_{\beta}\| \leq \|\nabla_q V(\hat{q})\|/\mu$. $\square$



Supplementary Figure S1: Controlled synthetic benchmark over five datasets and five seeds. Mean test accuracy is shown on the left, while median wall time on a logarithmic scale is shown in the middle. The completion fraction is shown on the right. All 400 requested runs produced valid endpoints. Boundary dynamics preserves the predictive results, although it relaxes more slowly than the compared methods.

Supplementary Table S1: Higher-capacity gradient-usability control. Accuracies are means and standard deviations over five deterministic seeds. Gradient diagnostics compare the boundary-generated centered gradient with exact centered EqProp at the same nudging strength. The final two rows compare the fixed 48-feature spiral configuration with its one-factor 96-feature extension.

Dataset	Boundary accuracy	Standard EqProp	Implicit	Mean relative error	Maximum relative error	Minimum cosine
Two moons	98.38 $\pm$ 0.71%	98.38 $\pm$ 0.71%	98.38 $\pm$ 0.71%	0.0793%	0.1904%	0.999998
Circles	99.88 $\pm$ 0.28%	99.88 $\pm$ 0.28%	99.88 $\pm$ 0.28%	0.1109%	0.3562%	0.999994
XOR	96.13 $\pm$ 0.52%	96.13 $\pm$ 0.52%	96.13 $\pm$ 0.52%	0.1566%	0.4757%	0.999989
Intertwined spirals						
48 RBF	73.88 $\pm$ 4.60%	73.88 $\pm$ 4.60%	74.00 $\pm$ 4.61%	0.6376%	1.7859%	0.999841
Intertwined spirals						
96 RBF	83.00 $\pm$ 3.38%	83.00 $\pm$ 3.38%	83.00 $\pm$ 3.38%	0.4884%	1.5408%	0.999882