Reversibilization of Free Will: Reversible Local Boundary Conditions, KL/Bregman Choice Operators, and Reversible Ledgers in Reversible Cellular Automata

Anonymous Author

Version: 1.10

Abstract

We propose a methodology for rigorously implementing "free will" as **reversible local boundary conditions** (RLBC) on finite domains of **reversible cellular automata** (RCA), and placing both the **information fidelity** characterization of "why choose this not that" and the **choice operator** in the same reversible ledger. The core approach is: design each "decision" as a **bijective update** on the boundary strip, carrying reversible **evidence/randomness registers** when necessary to preserve flattening and sampling paths, so that it **composes as an overall bijection** with one-step evolution of internal RCA; the decision itself given by **minimal-KL/I-projection soft selection**, degenerating to **hard selection** (argmax) in the **Γ-limit**, with all information overhead precisely accounted for in **Bennett reversible embedding** and reversibly recoverable. Reversibility and decidability are guaranteed by CHL representation theorem, Garden-of-Eden theorem, block/partition reversible cellular automaton structure, and linear-boundary matrix criteria; the undecidability of general neighborhoods in two and higher dimensions is circumvented by selecting **block permutation** or **linear reversible block** verifiable subclasses. This paper also connects the boundary-decision paradigm with the unified **windowed readout–phase–relative density of states–group delay** calibration, providing end-to-end error and stability statements.

1 Notation & Axioms / Conventions

1. **RCA and configuration space.** Finite alphabet A, $d \geq 1$, full space $X = A^{\mathbb{Z}^d}$. One-time-step global update $F : X \to X$ continuous under Cantor topology and commuting with shifts if and only if it is given by finite-range local rule (Curtis–Hedlund–Lyndon, CHL). If F is bijective, called RCA; on full lattice $\mathbb{Z}^d$, Garden-of-Eden theorem gives **surjective** $\Leftrightarrow$ **pre-injective (no twins)** $\Leftrightarrow$ **no orphans (no predecessor patterns)**, and **injective** $\Rightarrow$ **surjective**; thus bijective if and only if both injective and surjective [**?**].

2. **Block/partition reversibility and linear reversibility.** Adopt Margolus partition: if **intra-block transformation is permutation**, then global is reversible; reverse evolution implemented by inverse permutation and reverse-order partition. Reversibility of linear cellular automata under various boundary conditions (including "intermediate boundary") reduces to invertibility of rule matrix and Kronecker decomposition criterion [**?**].

3. **General undecidability and verifiable subclasses.** Reversibility problem of general-neighborhood CA in two and higher dimensions is **undecidable** (Kari); therefore this

paper focuses on **block permutation** and **linear-boundary matrix** two **verifiable** subclasses [?].

4. **Calibration identity (WSIG–EBOC convention).** On windowed scattering calibration, adopt the identity

$$\rho_{\rm rel}(E) = \frac{1}{2\pi i}\frac{\mathrm{d}}{\mathrm{d}E}\log\det S(E) = \frac{1}{2\pi}\operatorname{tr} Q(E) = \frac{1}{2\pi}\,\varphi'(E),$$

where $Q(E) = -i\,S^\dagger(E)\,\partial_E S(E)$, $\det S(E) = e^{i\varphi(E)}$. Birman–Kreĭn formula connects scattering phase and spectral shift function, serving as common mother scale for energy/time readout and "submission" [?].

5. **Information geometry and choice.** "Minimal-KL under linear consistency constraints gives exponential family/softmax and satisfies Pythagorean identity and Fenchel duality"; "TV–KL" Pinsker bound used for temperature-perturbation-jitter stability estimation [?].

6. **Reversible ledger.** Only information **erasure** dissipates (Landauer lower bound); **logical reversibility** can achieve **arbitrarily low dissipation** in the limit; random sampling and flattening evidence accounted for through reversible registers (Bennett) [?].

2 Model: Reversible Local Boundary Conditions (RLBC) on Finite Domain

Let $\Lambda \Subset \mathbb{Z}^d$ be a connected finite domain, take boundary strip $\partial\Lambda$ of sufficient thickness. One-step internal evolution given by some RCA F_Λ. Define **boundary layer operator**

$$\mathsf{B}:\ A^{\partial\Lambda}\times\mathcal{E}\ \longrightarrow\ A^{\partial\Lambda}\times\mathcal{E},$$

where $\mathcal{E}$ is **reversible auxiliary register** (evidence/randomness/flattening labels, etc.); this paper defaults $\mathcal{E}$ alphabet **finite**, and localizes configurations by boundary-touching blocks to ensure overall finite-alphabet CA framework (applicable to CHL/GOE). **One-round evolution** defined as

$$\mathcal{U}F_\Lambda\circ\mathsf{B}\quad\text{(boundary first, then internal).}$$

Call B **reversible local boundary condition** (RLBC) if satisfying:

- **(R1) Locality:** B can **read** finite outer shell beyond $\partial\Lambda$ (compress observations into $\mathcal{E}$ if necessary), but **only writes** $\partial\Lambda$ and $\mathcal{E}$, not touching Λ°;
- **(R2) Bijectivity:** $(b,\epsilon)\mapsto(b',\epsilon')$ is a permutation;
- **(R3) Evidence accounting:** All randomness, flattening indices, **selected action index** $a^\star$ **(or minimal sufficient evidence to reconstruct** $a^\star$**)** and observation evidence used for decision are written into $\mathcal{E}$ for inversion recovery (Bennett embedding) [?].

In partition implementation, let all "boundary-touching blocks" each undergo intra-block permutation; by **parallel direct product of block permutations = permutation**, B automatically satisfies (R2). Reverse-order Margolus partition and inverse permutations give B^{-1} [?].

3 Cascade Reversibility and Decidability

Proposition 3.1 (Finite Domain Cascade as Bijection)**.** *Let global F be RCA, and take its local implementation F_Λ on Λ. If B is RLBC,* ***and*** *B* ***and*** *F_Λ* ***adopt partition two-phase (e.g., Margolus) implementation such that boundary-touching blocks in each phase are pairwise disjoint, and execute synchronously according to "boundary phase → internal phase" two-phase schedule****, then*

$$\mathcal{U} = F_\Lambda \circ \mathsf{B}$$

is a ***finite-domain bijection*** *on state $(x|_\Lambda, b, \epsilon)$, with $\mathcal{U}^{-1} = \mathsf{B}^{-1} \circ F_\Lambda^{-1}$. When Λ and B are applied synchronously to the full lattice according to the above* ***partition tiling and two-phase schedule****, the resulting global evolution is RCA.*

Proof. $\mathcal{U}$ is bijective with $\mathcal{U}^{-1} = \mathsf{B}^{-1} \circ F_\Lambda^{-1}$. When Λ and B are applied synchronously to full lattice via partition tiling, by CHL the corresponding global map is continuous and commutes with shifts, and its inverse likewise [**?**]. □

Proposition 3.2 (Decidable Subclasses for Boundary-Internal Unity)**.** *In the following two implementation types, reversibility of $\mathcal{U}$ is* ***decidable*** *and inversion constructible:*

- ***Block/partition RCA:*** *If and only if each block rule is a permutation [**?**].*
- ***Linear-intermediate boundary:*** *Reversibility reduces to invertibility of rule matrix and its Kronecker decomposition; efficient algorithms available for multi-dimensional and intermediate boundaries [**?**].*

Remark 3.3 (General Undecidability)**.** *Reversibility of general-neighborhood CA in two and higher dimensions is undecidable (Kari), hence this paper's RLBC selects from verifiable subclasses [**?**].*

4 Choice Operator: KL/Bregman Fidelity (Soft→Hard)

Let boundary feasible action set $\mathcal{A}(b)$. Given baseline $q(\cdot \mid b)$ and moment constraints, introduce feature map $\phi : \mathcal{A}(b) \to \mathbb{R}^m$.

Assumption (Feasibility): $b^\star \in \mathrm{conv}\{\phi(a) : a \in \mathcal{A}(b)\}$.

Definition 4.1 (Soft Selection / I-Projection)**.**

$$p^\star(\cdot \mid b) \in \arg\min_{p\in\Delta(\mathcal{A}(b))} \Big\{ D_{\mathrm{KL}}(p\|q) :\ \sum_a p(a \mid b)\phi(a) = b^\star \Big\},$$

whose KKT condition gives

$$p^\star(a \mid b) \ \propto\ q(a \mid b) \exp\big(\langle \lambda, \phi(a)\rangle\big),$$

*i.e., exponential family/softmax; satisfies information geometry's Pythagorean identity and Fenchel-Legendre duality [**?**].*

Proposition 4.2 (Robustness and TV-KL Bound)**.** *When temperature/regularization parameter change introduces KL error δ, total variation deviation controlled by Pinsker bound*

$$\|p_1 - p_2\|_{\mathrm{TV}} \le \sqrt{\tfrac{1}{2} D_{\mathrm{KL}}(p_1\|p_2)},$$

*serving as "temperature–jitter" upper bound; Bretagnolle–Huber bound available for refinement when necessary [**?**].*

Theorem 4.3 (Γ-Limit: Soft$\to$Hard, via Entropy/KL Regularization)**.** *Let $q \in \Delta(\mathcal{A})$, cost $c : \mathcal{A} \to \mathbb{R}$. For $\tau > 0$, let*

$$p_\tau \in \arg\min_{p\in\Delta(\mathcal{A})} \left\{\langle c, p\rangle + \tau D_{\mathrm{KL}}(p\|q)\right\}$$

then $p_\tau(a) \propto q(a)\exp(-c_a/\tau)$, and as $\tau \downarrow 0$, $p_\tau \Rightarrow \delta_{a^\star}$ where $a^\star \in \arg\min_a c_a$; convergence unique if minimizer unique. ***(With linear moment constraint)*** *If adding $\sum_a p(a)\phi(a) = b^\star$ with **feasibility** ($b^\star \in \operatorname{conv}\phi(\mathcal{A})$), then there exists dual variable $\lambda(\tau)$ such that*

$$p_\tau(a) \propto q(a)\exp\left(\tfrac{\langle\lambda(\tau),\phi(a)\rangle - c_a}{\tau}\right).$$

*If minimizer in feasible set is **unique and a point mass** (exists $a^\star \in \mathcal{A}$ with $\phi(a^\star) = b^\star$), then as $\tau \downarrow 0$, $p_\tau \Rightarrow \delta_{a^\star}$ [**?**].*

Proof sketch. As $\tau \downarrow 0$, entropy/KL term weight decreases, linear objective dominates; in exponential family expression $-c_a/\tau$ exponent difference amplifies, concentrating probability mass on minimizer; Γ-convergence and large deviation principle give rigorous limit; moment constraint case via Lagrange multiplier scale analysis yields same conclusion. Uniqueness and selection stability given by Proposition **??**. □

5 Reversible Implementation: Bennett Embedding and "Reversible Sampling"

Theorem 5.1 (Reversible Decider)**.** *Any soft/hard selection on finite action set admits a **reversible extension** writing **randomness, flattening evidence, and sampling path** into $\mathcal{E}$:*

$$(b, \epsilon) \;\mapsto\; (b' = \mathrm{Sel}(b;\epsilon),\; \epsilon'),$$

*making this extension a permutation on (b, ϵ); inversion recovers sampling tree and flattening order from ϵ' and erases evidence, hence **no irreversible dissipation**.*

Proof. By Bennett logical reversibility: as long as intermediate information is not erased, can reverse erase. Implement sampling as controlled permutation of **prefix-tree/alias method**: Knuth–Yao DDG-tree gives entropy-optimal binary sampling framework; Walker/Vose alias method gives equivalent discrete sampling structure in constant amortized time. Writing tree/table indices and coin-flip sequences into $\mathcal{E}$ yields the result [**?**]. □

Remark 5.2 (Implementation Note (Verifiable Operator Family))**.** *On Margolus partition's **boundary-touching blocks**, apply in parallel via "selection result $\to$ intra-block permutation" to obtain B; its sufficiency and necessity as permutation and inversion construction directly follow from block reversibility [**?**].*

6 Formal Definitions and Main Theorems

Definition 6.1 (RLBC)**.** *In one-round evolution, boundary layer update*

$$\mathsf{B}(b, \epsilon) = \big(b', \epsilon'\big)$$

*satisfies (R1)–(R3), and in block implementation is direct product of **disjoint in one phase** boundary-touching block permutations.*

Theorem 6.2 (RLBC $\otimes$ Local RCA $\Rightarrow$ Finite Domain Bijection)**.** *If F_Λ is RCA and B is RLBC, then $\mathcal{U} = F_\Lambda \circ \mathsf{B}$ is finite-domain bijection; $\mathcal{U}^{-1} = \mathsf{B}^{-1} \circ F_\Lambda^{-1}$. If applying isomorphic B and corresponding F_Λ* ***via partition two-phase schedule*** *to every translate copy of the full lattice, with boundary-touching blocks* ***pairwise disjoint*** *in any phase, then obtain global RCA, with inverse given by reverse phase and inverse permutation playback [**?**].*

Theorem 6.3 (Choice = I-Projection; Soft→Hard)**.** *Let $\sum_a p(a)\phi(a) = b^\star$ be moment constraint with* ***feasibility*** *($b^\star \in \operatorname{conv}\phi(\mathcal{A})$), then*

(i) $p^\star = \arg\min D_{\mathrm{KL}}(p\|q)$ unique and exponential family;

(ii) For regularized family

$$p_\tau \in \arg\min_{p\in\Delta(\mathcal{A})} \Big\{ \langle c, p\rangle + \tau D_{\mathrm{KL}}(p\|q) \;:\; \sum_a p(a)\phi(a) = b^\star \Big\},$$

there exists *dual variable $\lambda(\tau)$; and* ***if*** *$b^\star$ is in* ***relative interior*** *of $\operatorname{conv}\phi(\mathcal{A})$ (Slater condition), then $\{\lambda(\tau)\}$ is* ***bounded*** *(has cluster point). In general feasible but non-interior case, $\lambda(\tau)$ may diverge while the following still holds:*

$$p_\tau(a) \propto q(a) \exp\left(\tfrac{\langle \lambda(\tau), \phi(a)\rangle - c_a}{\tau}\right).$$

If minimizer unique and a point mass *(exists $a^\star$ with $\phi(a^\star) = b^\star$), then as $\tau \downarrow 0$, $p_\tau \Rightarrow \delta_{a^\star}$;*

*(iii) Jitter stability controlled by Pinsker/Bretagnolle–Huber bounds [**?**].*

Theorem 6.4 (Reversible Ledgerization)**.** *Let Sel be the soft/hard selection of Theorem **??**. There exists a family of boundary block-permutations $\{\Pi^{(a)}_{block}\}_{a\in\mathcal{A}}$ and reversible register updates such that*

$$\mathsf{B}_\theta = \prod_{\textit{boundary-touching blocks}} \Pi_{block}(a^\star_\theta)$$

*constitutes RLBC, with all sampling/flattening information written into $\mathcal{E}$ and erased back in B_θ^{-1}, no Landauer cost [**?**].*

Theorem 6.5 (Linear-Boundary Reversibility Criterion)**.** *In linear CA and "intermediate boundary" settings, reversibility of $\mathcal{U}$ is equivalent to invertibility of corresponding rule matrix; matrix can be reduced-dimension tested via Kronecker decomposition [**?**].*

7 Connection with Windowed Readout–Phase–Density of States–Group Delay Calibration

Take the "evidence aggregation/consistency constraint" at boundaries from a class of windowed spectral readouts: in the absolutely continuous spectrum region, with the unified calibration

$$\rho_{\mathrm{rel}}(E) = \frac{1}{2\pi i}\frac{\mathrm{d}}{\mathrm{d}E}\log\det S(E) = \frac{1}{2\pi}\operatorname{tr} Q(E) = \frac{1}{2\pi}\varphi'(E),$$

where $Q = -i\, S^\dagger\, \partial_E S$, $\det S(E) = e^{i\varphi(E)}$, express "readout" as integral functional of phase derivative/relative density of states/group delay trace; Birman–Kreĭn formula gives equivalence of spectral shift and scattering phase, enabling explicit binding of "submission/decision" consistency constraint $\sum_a p(a)\phi(a) = b^\star$ to energy-phase ledger. Sensitivity of boundary-selection **soft→hard** transition to "readout jitter" controlled by Pinsker-type inequalities and linearized response estimates [**?**].

8 Paradigm Construction: Margolus-Boundary Reversible Decider

On two-dimensional Margolus partition, let outer ring all be "boundary-touching blocks". For each boundary-touching block, given finite action set $\mathcal{A} \subset \mathfrak{S}(A^{2\times 2})$ (intra-block permutation family) and feature map $\phi : \mathcal{A} \times \partial\Lambda \to \mathbb{R}^m$. According to §?? I-projection, for temperature $\tau > 0$ define soft selection distribution

$$p^\star_{\tau,\theta}(a \mid b) \;\propto\; q_\theta(a \mid b) \exp\big(\langle \lambda_\theta, \phi(a,b)\rangle/\tau\big),$$

whose hard limit is

$$a^\star_\theta(b) \in \arg\max_{a\in\mathcal{A}} \langle \lambda_\theta, \phi(a,b)\rangle.$$

Define

$$\mathsf{B}_\theta = \prod_{\text{boundary-touching blocks}} \Pi_{\text{block}}\big(a^\star_\theta(b)\big),$$

with reversible register update

$$(b,\epsilon) \;\mapsto\; \Big(b' = \Pi_{\text{block}}\big(a^\star_\theta(b)\big)(b),\; \epsilon' = \epsilon \oplus \text{code}\,\big(a^\star_\theta(b)\big)\Big),$$

where $\text{code}(\cdot)$ is reversible encoding of action index. Soft case uses reversible sampling to draw a from $p^\star_{\tau,\theta}(\cdot \mid b)$, writing sampling path and $\text{code}(a)$ into $\mathcal{E}$ to guarantee replay; hard case first computes $a^\star_\theta(b)$ based on b, writes $\text{code}\,\big(a^\star_\theta(b)\big)$ then applies permutation. Inverse process reads encoding, applies $\Pi_{\text{block}}(a^\star)^{-1}$ and erases encoding, thus B_θ is permutation on (b,ϵ). After cascading with internal block RCA, $\mathcal{U}$ maintains reversibility; ϵ accounting guarantees inversion can recover all random/evidence paths [?].

9 End-to-End Verifiable Checklist (Theory-Only, Experiment-Free)

1. **Reversibility verification:** Block-permutation or linear-boundary matrix method; undecidable region of two-dimensional general neighborhoods not entered into implementation aperture [?].

2. **Choice-fidelity:** Solve I-projection (or its convex dual); soft/hard mutually accessible under temperature parameter tuning; jitter-error via Pinsker/Bretagnolle–Huber bounds [?].

3. **Reversible ledger:** Reversibilization of Knuth–Yao/DDG-tree or alias method; randomness and indices written into $\mathcal{E}$; inversion erase-back [?].

4. **Calibration binding:** Via $\rho_{\text{rel}} = \frac{1}{2\pi}\,\varphi' = \frac{1}{2\pi}\text{tr}\,Q$ and Birman–Kreĭn formula, embed "readout→constraint" in unified energy/phase ledger [?].

Appendix A: Garden-of-Eden and RLBC Consistency

On Euclidean lattice, Garden-of-Eden theorem gives **local pre-injectivity⇔global surjectivity**. RLBC's block permutation implementation makes boundary layer **locally injective** in its action domain; internal RCA is also globally bijective; their cascade maintains injectivity and surjectivity, hence overall remains RCA. This argument relies on CHL's continuity-equivariance closedness and GOE's **surjectivity-pre-injectivity equivalence** (and **injectivity⇒surjectivity**) [?].

Appendix B: Γ-Limit for Soft→Hard Selection

Let feasible set be simplex $\Delta(\mathcal{A})$, cost $c : \mathcal{A} \to \mathbb{R}$, baseline distribution $q \in \Delta(\mathcal{A})$. Define functional

$$\Phi_\tau(p) = \langle c, p\rangle + \tau D_{\mathrm{KL}}(p\|q).$$

Then $p_\tau \in \operatorname{argmin} \Phi_\tau$ gives $p_\tau(a) \propto q(a) \exp(-c_a/\tau)$. As $\tau \downarrow 0$, KL term weight tends to zero, Φ_τ **Γ-converges** to linear functional $\Phi_0(p) = \langle c, p\rangle$, minimized at simplex vertices (point masses). If further assuming equi-tightness and unique minimizer $a^\star \in \arg\min_a c_a$, **then minimizer sequence p_τ converges in weak topology** to point mass: $p_\tau \Rightarrow \delta_{a^\star}$. Large deviation principle guarantees probability mass concentration rate at exponential scale $1/\tau$. **Moment constraint case (assuming feasibility):** Adding $\sum_a p(a)\phi(a) = b^\star$ (assuming $b^\star \in \operatorname{conv} \phi(\mathcal{A})$), **there exists** dual variable $\lambda(\tau)$; **if** $b^\star$ is relative interior (Slater condition), $\{\lambda(\tau)\}$ is **bounded** (has cluster point), otherwise its norm may diverge while hard limit still equivalent to constrained linear programming. Exponential family solution

$$p_\tau(a) \propto q(a) \exp\left(\tfrac{\langle\lambda(\tau),\phi(a)\rangle - c_a}{\tau}\right).$$

As $\tau \downarrow 0$, limit problem equivalent to $\min\{\langle c, p\rangle : \sum_a p(a)\phi(a) = b^\star\}$; if minimizer $a^\star$ unique and a point mass, then $p_\tau \Rightarrow \delta_{a^\star}$ [**?**].

Appendix C: Reversible Sampler Construction Outline

- **DDG-tree (Knuth–Yao):** Optimal discrete sampling with random bits as source; writing visit path (left-right branches) and leaf number into $\mathcal{E}$ yields reversible implementation [**?**].
- **Alias (Walker/Vose):** Constant-time sampling after preprocessing two tables; writing table index and threshold comparison result into $\mathcal{E}$, replaying in inversion [**?**].
- Both compatible with Bennett's "save-erase-back" strategy, hence no irreversible thermal lower bound [**?**].

References

[1] Hedlund. *Endomorphisms and Automorphisms of the Shift Dynamical System* (CHL characterization). `https://link.springer.com/article/10.1007/BF01691062`

[2] *Reversible cellular automaton.* Wikipedia. `https://en.wikipedia.org/wiki/Reversible_cellular_automaton`

[3] Kari. *Reversibility of 2D cellular automata is undecidable.* `https://ui.adsabs.harvard.edu/abs/1990PhyD...45..379K/abstract`

[4] Smith, F.T. *Lifetime Matrix in Collision Theory.* Phys. Rev. 1960. `https://link.aps.org/doi/10.1103/PhysRev.118.349`

[5] Csiszár, I. *I-Divergence Geometry of Probability Distributions and Minimization Problems.* 1975. `https://pages.stern.nyu.edu/~dbackus/BCZ/entropy/Csiszar_geometry_AP_75.pdf`

[6] Bennett, C.H. *Logical reversibility of computation.* IBM J. 1973. `https://dl.acm.org/doi/10.1147/rd.176.0525`

[7] Chang et al. *Reversibility of linear cellular automata with intermediate boundary condition.* https://www.aimspress.com/article/doi/10.3934/math.2024371

[8] *Pinsker's inequality.* Wikipedia. https://en.wikipedia.org/wiki/Pinsker's_inequality

[9] *An Introduction to Γ-Convergence.* Springer. https://link.springer.com/book/10.1007/978-1-4612-0327-8

[10] Knuth, Yao. *The complexity of nonuniform random number generation.* https://www.semanticscholar.org/paper/The-complexity-of-nonuniform-random-number-Knuth-Yao/58f10efb7c76b41a6ddc26ff9ff94f7faa1e2e35

[11] *Alias method.* Wikipedia. https://en.wikipedia.org/wiki/Alias_method

[12] *The Garden of Eden theorem: old and new.* arXiv:1707.08898. https://arxiv.org/abs/1707.08898