EchoKey Ethical Supplement

Reversibility, Optimization, and Chemistry — No Comfort. No Hype. No Erasure.

This document is a full ethical deep dive accompanying the EchoKey unified reversible coin framework.

0. Contract of the Reader

We will not comfort you. We define every term we use, state each claim and its scope, quantify tradeoffs, and give procedures to reproduce all ethical metrics. If you dislike a symbol, rename it. *Everything is explicit and reproducible in code.*

1 Scope, Boundaries, and Definitions

1.1 What we built

A single measurement–free, reversible "Bennett sandwich" core (compute → use → uncompute), paired with an exact inhomogeneous coin over a finite configuration set. We instantiate it in two domains:

- **Optimization (TSP):** anchored tours, adjacent–swap neighborhood, exact three–edge ΔL; EchoKey–7 features with **INT**$= -\Delta L$ primary.
- **Chemistry (sandbox):** finite conformer/graph state set, reversible local moves, per–move **INT**$= -\Delta E$ from a classical oracle/surrogate suitable for tiny, auditable examples.

1.2 What we explicitly do *not* claim

(a) No polynomial speedups, no asymptotic advantages, no claims beyond small–N truth machines.

(b) No claim that logical reversibility zeroes physical energy use. Logical $\neq$ physical. Our simulator consumes energy; we report it.

(c) No black–box classical oracles hidden inside U. Any baseline is evaluation–only and outside the unitary.

(d) No people–routing or coercive decision support. Our TSP demo is for emissions–proxy distance only; human logistics is out of scope.

(e) No chemistry synthesis or hazardous pathway discovery.

1.3 Definitions

- **Small–N truth machine:** A deliberately tiny, exact construction used to demonstrate correctness, reversibility, and accounting.
- **Invalid mass:** Probability on padded states introduced by register embedding. Our shift keeps padded subspace invariant; no leakage.

- **Ancilla–zero rate:** Fraction of runs where all ancillas are measured zero immediately after the full sandwich.
- **Reversibility fidelity:** Overlap $F = \langle\psi| U^\dagger U |\psi\rangle$; equals 1 ideally, measured up to floating–point.

2 Ethical Principles and Theoretical Grounding

2.1 Landauer, Bennett, Feynman — logic vs physics

Landauer bounds the *physical* cost of *erasing* information. Bennett shows classical computations can be made *logically* reversible by preserving information and uncomputing. Feynman observed that quantum dynamics is unitary and time–reversible in the ideal. Our core respects the logical side: compute a delta into a work register via ROM–style XOR, *use only its sign*, and *return* the work register and chain ancillas to $|0\ldots0\rangle$.

2.2 Unitarity and measurement discipline

All stages are unitary; there is no mid–circuit measurement or erasure. The only net effect of the Bennett sandwich is on a marker and a flag; the work and selection ancillas are clean.

2.3 Energy honesty

Logical reversibility is not free energy. We run on classical hardware; we report wall–time and a rough energy estimate and convert to CO_2e with stated grid factors. Physical reversibility requires specialized devices — which we do not build, claim, or simulate faithfully.

3 Societal Impact by Track

3.1 Track 1: Reversible computing

Benefits: literacy in reversible logic; concrete, auditable exemplar of Bennett compute→use→uncompute; reusable testing harness.

Risks: greenwashing ("reversible = free energy"); miscommunication of limits.

Mitigations: energy appendix; explicit claims table; negative results included. We show ancilla–zero, $\Pr[\Delta{=}0^b]$ (before/after), and $F \approx 1$.

3.2 Track 2: Optimization (TSP)

Benefits: toy routing that is emissions–proxy aligned; demonstrates constraint hooks (no–go zones, fairness caps). Transparent optimality gaps.

Risks: dual–use for people routing or surveillance; overclaiming "quantum advantage".

Mitigations: repository `USE_POLICY.md`: forbid people–routing; require consent, impact assessment, and institutional review for any human–centric use. No advantage claims. Constraints and penalties are explicit in code.

3.3 Track 3: Chemistry

Benefits: clear reversible predicate with ΔE; honest tiny sandbox aligned to sustainability exemplar (battery/pollutant proxy).

Risks: pathway misuse; misinterpretation as predictive chemistry at scale.

Mitigations: only benign molecules; explicit scope limits; property errors vs baseline; no synthesis tools; document surrogate limits.

4 Risk Register and Mitigations

Risk	Harm	Lik.	Mitigation	Residual
Greenwashing (reversible → free energy)	Public misperception, policy distortion	Med	Energy appendix, explicit physical vs logical distinction	Low
Dual–use logistics	Rights harms (coercion, profiling)	Med	Use policy forbids people–routing; constraint hooks demonstrated	Low–Med
Chemistry misuse	Safety/toxicity exploration	Low–Med	Benign sandbox, no synthesis; scope statement	Low
Overclaiming quantum	Misallocation, hype	Med	"Small–N truth machine" framing; publish negative results	Low
Compute footprint	CO_2e	Med	Minimal runs; report wall–time/energy; prefer green power when possible	Low

5 Governance, Guardrails, and Licensing

5.1 Use policy (excerpt for `USE_POLICY.md`)

- **Allowed:** pedagogy, research on reversible logic, toy routing on synthetic instances, benign chemistry sandboxes.
- **Prohibited:** people–routing, surveillance optimization, weapons or toxin design, any human deployment without consent and independent ethics review.
- **Review triggers:** any move from toy to real data; any coupling to decision systems; any chemical synthesis coupling.

5.2 License

We publish under CC0 to maximize verifiability and remix. We pair it with a prominent `USE_POLICY.md` and `Ethics.md`. Freedom to use does not nullify responsibility to use ethically.

6 Transparency, Reproducibility, and Logging

- **Reversibility proofs printed:** ancilla–zero rate, $\Pr\left[\Delta{=}0^b\right]$ drop after compute and return after sandwich, $F = \langle\psi| U^\dagger U |\psi\rangle$.
- **Resource accounting:** gate counts, depth, qubit registers, padded mass.
- **Optimization metrics:** optimality gap, hit rate modulo reversal, constraint satisfaction.
- **Chemistry metrics:** property error vs baseline; surrogate provenance and limits.
- **Artifacts:** CSV/JSON logs with seeds, CLI flags, commit hash, host specs.

7 Environmental Footprint Methodology

7.1 Measurement plan

We report for each run:

$$E_{\text{est}} \approx P_{\text{CPU}}^{\text{avg}}\, t + P_{\text{GPU}}^{\text{avg}}\, t, \quad \text{and} \quad \text{CO}_2\text{e} = E_{\text{est}}\, \frac{1\,\text{kWh}}{3.6 \times 10^6\,\text{J}}\, g_{\text{grid}}, \tag{1}$$

where g_{grid} is a stated grid intensity (kg CO_2e per kWh). We disclose measurement uncertainty and that this is a rough upper bound.

7.2 Reporting

We include wall–time, seeds, machine model, CPU/GPU utilization (if available), and g_{grid} reference. We avoid deceptive "per–op" normalizations.

8 Human Subjects, Consent, and Data

No human data is used. For any future real–world routing, *prior, informed consent* and independent review would be mandatory. Our demo is synthetic.

9 Dual–Use and Export Control

We do not provide synthesis, design, or optimization pathways for hazardous agents. Any chemistry content remains in benign toy domains. Users are responsible for compliance with applicable laws.

10 Communication Ethics (Video, README)

The video will contain: (i) what we built in one line, (ii) reversible core animation, (iii) TSP toy with constraints, (iv) chemistry toy with error bars, (v) ethics 'n energy slide, (vi) limitations slide. No advantage claims; explicit small–N framing.

11 Audit Checklist (Pre–Submission)

1. README quickstart + flags table; Ethics.md + Use_Policy; LICENSE; citations.
2. Scripts: identical CLI; seeds and logs saved; commit hash printed.
3. Reversibility logs present: ancilla–zero, $\Pr\left[\Delta{=}0^b\right]$, F.
4. Optimization metrics: gap, constraints, symmetry hits.
5. Chemistry metrics: baseline error; surrogate notes.
6. Energy appendix: wall–time, E_{est}, grid factor.
7. Video recorded to 5–10 min; captions; no hype.

A Metric Definitions

- $\Pr[\Delta{=}0^b]$: probability work register is all zeros at probe time.
- Ancilla–zero: all selection/chain ancillas measure to 0 immediately after the sandwich.
- Reversibility fidelity F: numerical overlap after $U^\dagger U$.
- Padded mass: probability on P_{pad} subspace.

B CLI Examples (Reproducibility)

```
# TSP
python tsp_reversible_coin.py --N 6 --steps 8 --beta 0.25 --stay_gap 3.0 --seed 7 --repeats

# Chemistry (toy sandbox)
python chem_reversible_coin.py --steps 8 --beta 0.25 --seed 7 --repeats 3
```

C References (indicative)

R. Landauer (1961). Irreversibility and Heat Generation in the Computing Process. *IBM J. Res. Dev.*
C. H. Bennett (1973). Logical Reversibility of Computation. *IBM J. Res. Dev.*
R. P. Feynman (1982). Simulating Physics with Computers. *Int. J. Theor. Phys.*