
The Cost of Soundness: f-Contour Bounds and the Inductive-Bias Limits of Learned Heuristics

Eduardo Perez Denadai
Optimality.ai

Abstract

A verification-centric system lets a **learned proposer** direct a **pinned, exact authority** without ever being trusted with the guarantee: the authority owns the guard, the proposer only *restricts*. Two questions follow. What does that certification cost in search effort? And, within that cost, can a learned heuristic ever beat a classical one?

We answer both with a small theory and a controlled empirical programme. The cost of certification is geometric — a certifying search expands exactly the region where the heuristic’s estimate falls below the true cost, so its price is the **volume of the f-contour** — and a learned heuristic can beat a classical one only by reducing that heuristic’s error (Theorem 3). We then test the question empirically, one axis at a time. The first result is a **negative**: a certified “accelerator” that appeared to beat the classical baseline at $0.72\times$ nodes was, at **matched weight**, worth $1.00\text{--}1.06\times$ — the speedup was the greedy weight, not the learned heuristic. The second is a **positive**: the apparent wall is not that global functions are unlearnable, but that they are not learnable by architectures whose **inductive bias is local** — a differentiable relaxation whose recurrence *is* the layer crosses the tight classical heuristic at $0.93\times$, stays at $\approx 0.90\times$ across sizes $9\rightarrow 21$, and, on a second global predicate, crosses at $0.96\times$ only when its bias is rebuilt to match that predicate. We state these as theorems with constructive proofs, name the design rule they hand to practice — **choose the bias for the predicate** — and pose the open problem they leave: a characterization of *for which (architecture, predicate, regime) triples* a class can match a predicate. We then measure the **regime axis** of that open problem on a second substrate — certified search over continuous traffic — where the law *without* the regime fails its own crossing bar ($1.02\times$) and the regime qualifier carries it cell by cell at $\rho = +0.79$: part (iii) of the theorem moves from assertion to empirical necessity. We name the class of systems: **sound-by-verification**.

1 Introduction

This is the fourth paper in a series on the Ω -TD neurosymbolic architecture. The first three established the architecture [Denadai, 2026c], its belief-space generalization [Denadai, 2026a], and an empirical characterization of the learned components [Denadai, 2026b]. This paper supplies the theory those results require: what certification *costs* — the f-contour bound — and what a learned heuristic can buy inside it, the inductive-bias law that governs the crossing, and its extension to transfer across scales and predicates.

Verification-centric architectures are built on a division of labour: a **pinned authority** Ω — deterministic, parameter-free, returning a typed verdict — a **proposer** π that may be a learned model, a search, or an adversary, and a **guard** owned by the authority. Execution selects from the guard’s admissible set plus an abstention; the proposer’s only power is to choose *within* it. The safety of such a system is structural and, as we recall in §2, does not depend on the proposer at all.

What is *not* structural is its **efficiency**. A certifying search must still search: it expands the states whose estimated cost could still undercut the true optimum, and the learned heuristic is used only to *order* that search. Two questions therefore define the field. First, **what does certification cost** — how many nodes must be expanded before the authority is satisfied? Second, **can learning pay for itself** inside that cost — is there a learned heuristic that beats the best classical one?

The second question is easy to ask badly. Reported “learned accelerators” in this space have repeatedly measured the *weight* of a weighted search rather than the heuristic (e.g. §4.1), and the honest measurement requires the **matched-weight control** — hold the certified suboptimality budget fixed and compare only the heuristic. Our contribution is to take that control seriously, and then to take the resulting negative seriously: the question is not whether *a* learned heuristic helps, but *what property* of a model makes one possible.

We give three results, in the order they were discovered:

1. **The apparent speedup was the weight.** At matched weight the learned heuristic’s marginal value is $1.00\text{--}1.06\times$ (§4.1). The negative is not a dead end; it is a refocusing of the question onto *error reduction*.
2. **The wall is inductive bias, not learnability.** Holding capacity fixed, three model families told one story: the local and generic arms cannot reduce the error of a tight classical heuristic, while a differentiable relaxation — whose recurrence *is* the layer — crosses at $0.93\times$ (§4.2). The crossing survives a scale sweep (§4.3).
3. **The bias must match the predicate.** On a second global predicate, the *same* relaxation fails ($1.58\times$) until its bias is rebuilt for that predicate, at which point it crosses ($0.96\times$) (§4.4).

From these we state the general law as a theorem (Thm 4) and hand practice the design rule *choose the bias for the predicate* (§5), leaving one open problem (§6).

2 Preliminaries: containment and the price of certification

Containment. With exact evidence channels, a parameter-free guard, and always-available abstention, the proposer cannot leave the safe set — for an *arbitrary* proposer (Theorem 1; composition in Theorem 2). Safety is thus structural, and the interesting question moves to cost.

The f-contour. Consider an exact search (A^*) to a goal of optimal cost C^* under a heuristic h . A^* expands exactly the states with $g(n) + h(n) < C^*$, so the cost of the search is the **volume of the f-contour** $\mathcal{F}(h)$ — a purely geometric quantity determined by h . With relative error ε , expansions are bounded by the volume of $\{f \leq (1 + \varepsilon)C^*\}$ (Dechter–Pearl; Korf).

Theorem 3 (No speed without error reduction). Let h_c be an admissible classical heuristic and h_θ a learned heuristic used only to *order* a certifying search. If h_θ has no smaller error than h_c , the certified search with h_θ expands no fewer nodes than with h_c .

The proof is one line of monotonicity: $\mathcal{F}$ shrinks only as the pointwise under-estimate of h grows, and a larger error means a smaller estimate. Two corollaries follow. The weight of a weighted search is classical — any admissible h may be used — so an apparent speedup that compares weighted-learned against unweighted-classical is measuring the weight. And because h^* is a *global* function (s-t-connectivity-like, L/NL -complete), reducing error below a tight classical heuristic is a **learnability** question about the model class, not about data or capacity.

3 Method: change one axis, watch which way the ceiling moves

The corollaries are testable, and so is the boundary they describe. Our method is to **change one axis at a time** — capacity, certificate, inductive bias — and observe which one moves the ceiling. A limit that survives every axis-swap is fundamental; one that falls is a design choice in disguise. Each experiment registers its bars and falsifiers before the run, and each carries a control matched on the axis it is supposed to isolate: the **matched-weight control** (§4.1), fixed-capacity arms (§4.2), and the predicate ablation (§4.4). The substrate is a family of procedurally generated mazes (uniform

Table 1: The matched-weight control (control_ch_mg.json, slack_low). Where the classical heuristic is tight, the learned h adds $\approx 1.0\times$.

regime	apparent win vs unweighted	learned h 's marginal at matched w
open	0.558	1.000
sparse	0.672	1.059
dense	0.899	1.064
packed	0.947	1.022
wfc	0.824	1.356

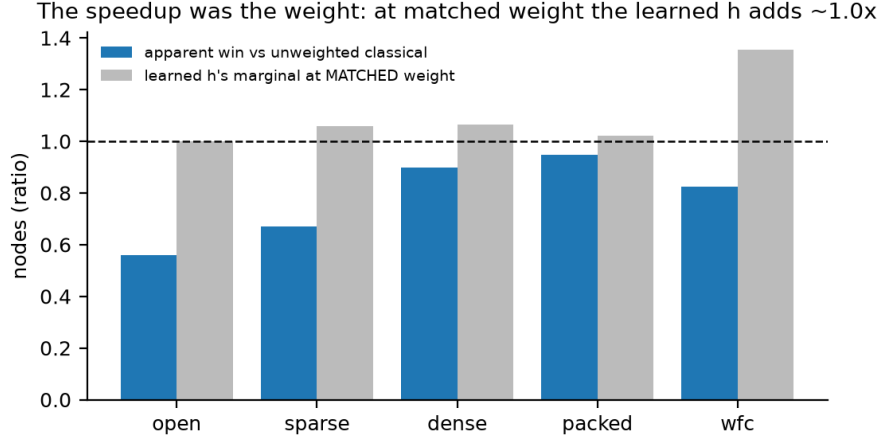


Figure 1: The certified win is the *weight*, not the learned heuristic. Blue bars are the apparent speedup against the unweighted classical baseline; grey bars are the learned h 's true marginal value at **matched** weight.

random, WFC, and hand-tuned regimes), so that “tight classical heuristic” means a real ALT baseline with measurable slack, not a straw man.

4 Results

4.1 The apparent accelerator was the weight

A certified search that beat the unweighted classical baseline at $0.72\times$ nodes is the motivating case. Held to matched weight, the learned heuristic’s marginal value is $1.00\text{--}1.06\times$ wherever the classical heuristic is already tight (Table 1). The $0.72\times$ was the greedy weight w , and Theorem 3 predicted it: without error reduction, the weight is the only thing that moved.

The negative is informative rather than terminal: it relocates the question from “does learning help?” to “what property of a model reduces error against a *tight* heuristic?” — the question Theorem 3’s second corollary poses, and the subject of §4.2.

4.2 The wall is inductive bias, not learnability

Predicting distance-to-go on hard grid mazes, we fixed capacity and varied only the inductive bias. The local patch and a message-passing network of depth far beyond the graph diameter cannot reduce the error of a tight ALT heuristic; a differentiable relaxation — a Bellman–Ford recurrence implemented *as* the layer — crosses (Table 2, Figure 2).

The wall the programme spent a chapter mapping is therefore not “global functions are unlearnable” but “not learnable by architectures whose **inductive bias is local**” — a corollary of Theorem 3’s second statement, now with a constructive witness.

Table 2: The inductive-bias frontier (`bias_frontier.json`, rows). Fixed capacity; only the bias changes.

arm	inductive bias	held-out search vs the tight classical heuristic
local MLP (patch)	none	$1.42\times$ (fails)
message passing, depth $\gg$ diameter	locality	$2.16\times$ (fails)
differentiable relaxation	the algorithm is the layer	$0.93\times$ (crosses)

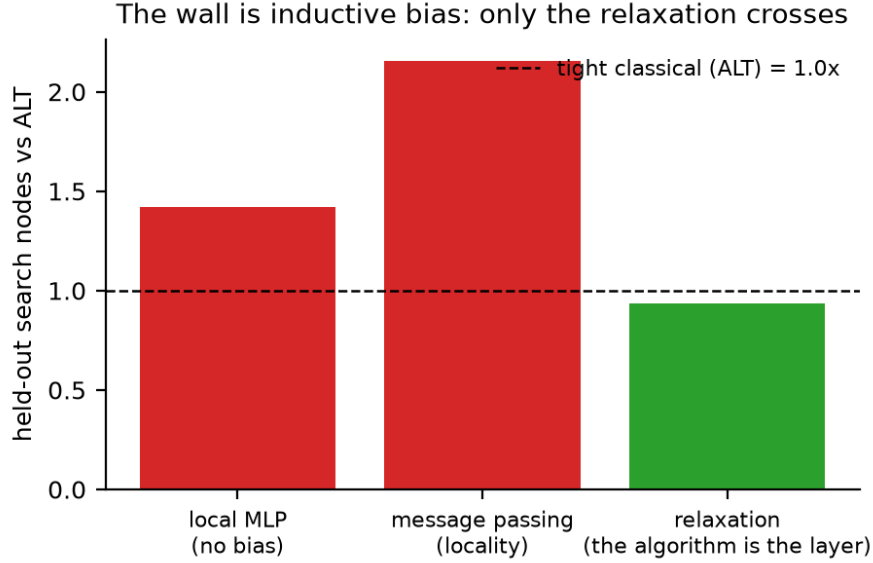


Figure 2: The wall is inductive bias, not learnability. Held-out search nodes vs the tight classical heuristic ($ALT = 1.0\times$): the local MLP and message passing fail, while the relaxation crosses at $0.93\times$.

4.3 And it scales

A crossing that appears only in a small regime is a curiosity. Trained once at size 9 and tested at 9, 13, 17, and 21, the relaxation arm crosses at **every** size, holding $\approx 0.90\times ALT$, while the local and message-passing arms degrade monotonically (Table 3, Figure 3). The learned seed transfers and the bias scales with the problem: the crossing is a method, not an accident of the small regime.

4.4 And the bias must match the predicate

If an algorithmic bias were sufficient *in the abstract*, the relaxation that crossed distance-to-go should cross any global predicate. It does not. On a second predicate — reach the goal **through** a designated waypoint — the plain relaxation fails at $1.58\times$, while a relaxation rebuilt for that predicate (seeded at the waypoint *and* the goal, combined through the predicate) crosses at $0.96\times$ (Table 4, Figure 4). The falsifier for a universal bias is measured, and it fails.

This is the sharpest of the results: it converts “the bias matters” into “the bias must *match* the predicate,” a methodology rather than a curiosity.

4.5 And the regime decides when matching helps

The theorem’s part (iii) says matching is a property of the *triple* $(\mathcal{H}, P, R)$ — the evidence regime — because h_P^* can be global in one regime and local in another. Until now that component was asserted, not measured. The safety substrate’s search problem provides the test: the certified search over

Table 3: Crossing generalization (crossing_generalization.json).

test size	local MLP	message passing	relaxation
9 (trained here)	1.080	1.649	0.898
13	1.452	2.235	0.923
17	1.528	2.409	0.899
21	1.515	2.427	0.902

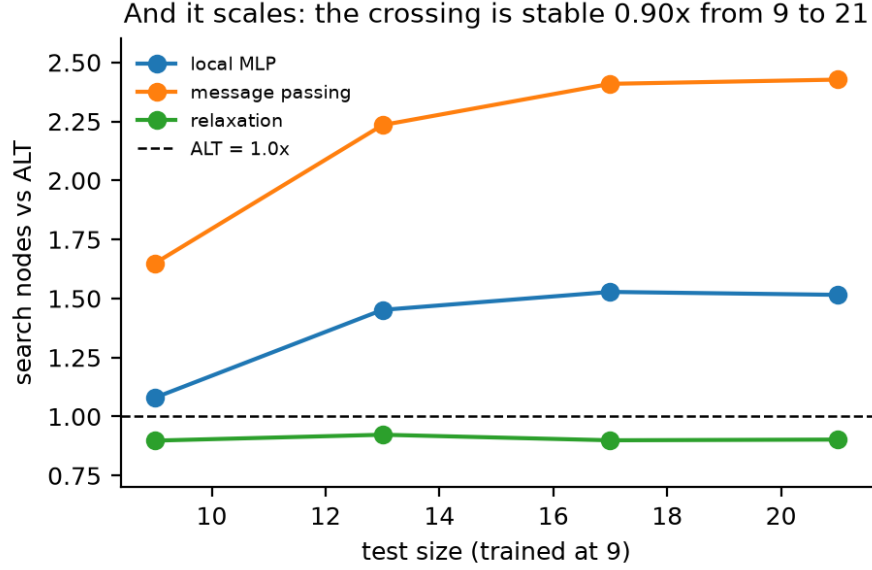


Figure 3: And it scales. Trained at size 9, the relaxation stays at $\sim 0.90\times$ the tight classical heuristic at every test size, while the local arms degrade monotonically.

discrete driving maneuvers, where the classical baseline is the hand-coded speed-first order and the predicate is collision-freedom under the authority’s kinematics. Three orderings, scored at identical decision states (paired): **classical** (speed-first), **matching** (a relaxation ranked by predicted clearance under the authority’s own dynamics), and **local** (a frozen-leader gap model that ignores closing speed — a plausible bias that does *not* match the predicate).

The registered claim — matching crosses the f-contour at $\leq 1.0\times$ classical where local does not — **fails on the aggregate**: local fails to cross as predicted ($1.20\times$), matching clearly beats local (1.42 vs 1.67 mean expansions), but matching misses the bar at $1.02\times$ classical. Reported, not moved. The per-cell measurement then explains the failure. Sweeping all seven scenarios $\times$ five densities and recording, for each cell, the **FASTER-admissibility** — the fraction of decision states where classical’s first candidate certifies — the classical–matching gap tracks 1 – admissibility at $\rho = +0.79$ (cell) and $+0.79$ (scenario aggregate). Where the regime makes a predicate-blind bias optimal (roundabout, 97% admissible), the matching relaxation is pure overhead ($+1.55$ expansions); where it does not (highway, 36%), matching wins decisively (-0.63). The aggregate that failed simply averaged opposing regimes.

This is the necessity result for part (iii), not merely consistency: **without the regime in the theorem, the (architecture, predicate)-only law is false on this substrate’s own data** — the same bias that must match the predicate (§4.4) can only *show* its gain where the regime leaves it something to fix. The design rule therefore refines to: **choose the bias for the predicate, in the regime you will run in** — and the regime is measured, not assumed.

Table 4: The predicate axis (`predicate_axis.json`). The same relaxation that crosses distance-to-go fails the constrained predicate until its bias is rebuilt.

arm	inductive bias	vs the tight constrained ALT
local MLP (patch)	none	$2.043\times$ (fails)
message passing	locality	$2.910\times$ (fails)
plain relaxation	distance-to-go	$1.584\times$ (fails)
predicate-matching relaxation	seeded at w and g	$0.957\times$ (crosses)

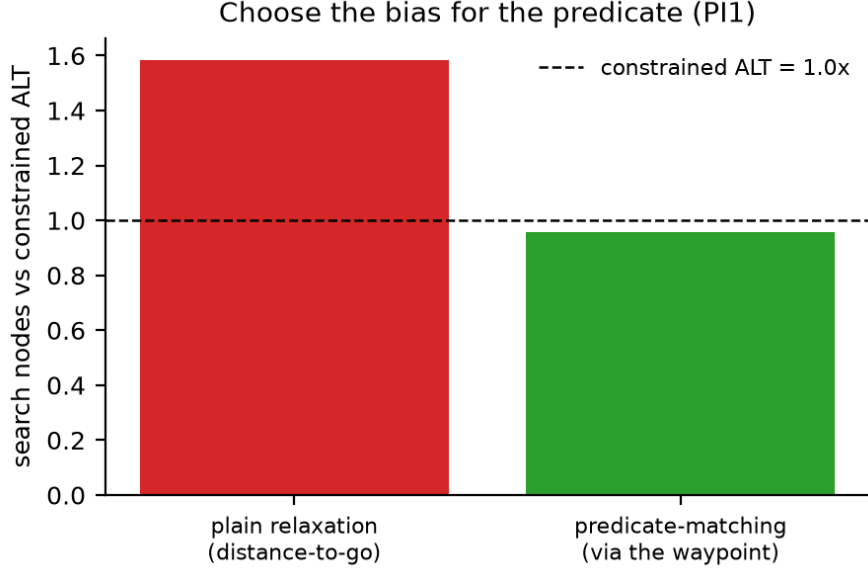


Figure 4: Choose the bias for the predicate (PI1). The plain relaxation that crossed distance-to-go ($0.93\times$) fails the constrained-reachability predicate at $1.58\times$; the predicate-matching relaxation crosses at $0.96\times$.

5 The law, and what it hands to practice

The findings above are one law. Make the predicate explicit: for a task predicate P with value function h_P^* , a heuristic class $\mathcal{H}$ **matches** P at error ε if it contains a member with relative error $\leq \varepsilon$ against h_P^* .

Theorem 4 (Bias–predicate matching). Let h_c be a classical admissible heuristic for P with error ε_c . (i) If $\mathcal{H}$ matches P at $\varepsilon \leq \varepsilon_c$, the certified search with that member expands no more nodes than with h_c — and if $\mathcal{H}$ contains a monotone transform of h_P^* (error 0), it expands the minimum. (ii) If $\mathcal{H}$ misses P at $\varepsilon_0 > \varepsilon_c$, no member expands fewer nodes than h_c . (iii) Matching is a property of the triple $(\mathcal{H}, P, R)$ — the predicate’s *evidence regime* R — because h_P^* can be global in one regime and local in another.

Theorem 4 is Theorem 3 generalized to an arbitrary predicate, with part (iii) explaining why the same bias can match one predicate and miss another. Part (iii) is no longer only asserted: §4.5 measures the regime axis on a second substrate and shows it is *load-bearing* — drop R and the law is falsified by its own aggregate. Its practical content is a method for localizing a ceiling — change one axis, watch which one the ceiling moves with — and the design rule it hands to practice is one sentence: **choose the bias for the predicate, in the regime you will run in.**

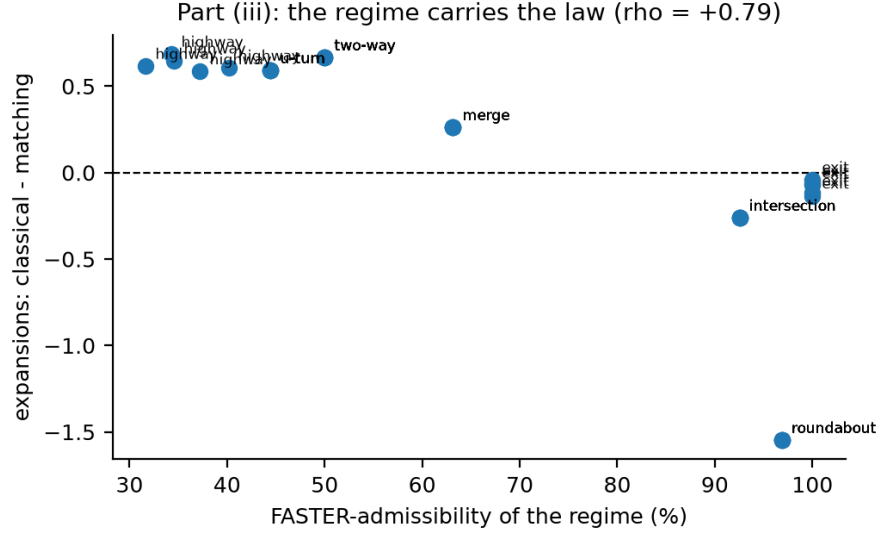


Figure 5: The regime qualifier, measured. Per scenario $\times$ density cell, the classical–matching expansion gap against FASTER-admissibility ($\rho = +0.79$, `bias_regime.json`). Without R the law is falsified by its own aggregate ($1.02\times$); with R every cell falls on the predicted line.

6 Limitations and open problems

The theorems are exact under their assumptions; the empirical content rests on a maze substrate and a handful of predicates — with one second substrate (continuous traffic) for the regime axis alone (§4.5) — so the *scope* of “matches” is a question, not a result. Three items are open.

The characterization of matching. Theorem 4 makes matching sufficient and necessary given the error comparison, so the open problem is to predict that comparison from the *structure* of $\mathcal{H}$ and the predicate’s computation. Our instances are the first data points — local and message-passing classes miss global predicates at every scale; a relaxation that *is* the predicate’s recurrence matches. The falsifier is a single counterexample: an $\mathcal{H}$ that crosses P while missing P ’s computation.

The regime index. Part (iii) makes matching depend on R , but we have only measured two regimes (the high- and low-observation limits of a witness predicate elsewhere in the programme). A systematic map of R is missing.

Scope. We do not claim that no learned heuristic can reduce error — the relaxation shows one can. We claim only that it must reduce error to speed a search, that such a reduction is a learnability question set by the model’s inductive bias, and that the bias must match the predicate being asked of it.

7 Conclusion

The safety of a verification-centric system is structural and free; its *efficiency* is not. The price of certification is the volume of the f-contour, and a learned heuristic pays only by reducing error (Thm 3). Measured honestly, an apparent accelerator was the weight; measured with the right bias, a differentiable relaxation crosses the tight classical heuristic at $0.93\times$, holds $\approx 0.90\times$ under scale, and crosses a second predicate only when its bias is rebuilt for it. The general statement (Thm 4) turns this into a methodology — *choose the bias for the predicate* — and a single open problem: characterize which classes match which predicates in which regimes.

References

Eduardo P. Denadai. Exact-authority belief-space planning: Where composition emerges and where it fails, 2026a. Optimality.ai.

Eduardo P. Denadai. Learnable boundaries, unlearnable competence: A characterization of learned components in a symbolic search stack, 2026b. Optimality.ai.

Eduardo P. Denadai. Containment rather than calibration: A sound architecture for learning under partial observability, 2026c. Optimality.ai.

8 Provenance

Every empirical claim is reproduced from the isolated subsystems `../..certified_heuristic/`, `../..expressivity_bound/`, and — for §4.5 — `../..env_substrates/highway/`; the figures and tables are generated **verbatim** from their artifacts by `make_figures.py` (nothing hard-coded): Table 1 and Figure 1 from `control_ch_mg.json` (`slack_law.frontier`); Table 3 and Figure 3 from `crossing_generalization.json`; Table 4 and Figure 4 from `predicate_axis.json`; Table 2 and Figure 2 from `bias_frontier.json`; Figure 5 and §4.5’s numbers from the highway substrate’s `bias_regime.json` (bars H-6/H-6-D registered in `env_substrates/highway/docs/PLAN.md` before the runs; H-6’s crossing FAIL is reported where it occurred). Containment and the witness results live in `../..minigrid_refit/` and `../paper_model_free/`, `../paper_model_based/`. Bars and falsifiers were registered before each run; the failed bars are reported where they occurred (§4.1 in particular).