

Calculation protocols and decision identifiability

Ivan B. Kurpishev / Иван Борисович Курпишев

01 / Findings and scope

Continuation of B3.5. A local protocol intake checker compares actual bytes, structure and atom counts, input/output ENCUT, energy definitions and run-log associations. The engineering implementation is ready; original pilot protocols have not been obtained.

A fresh JVASP-114724 record [S4] matches the B3.4 retained SHA-256: bc8d103827a6926e74b1fa20279096fecf37f01c9e96731612a6496d04cb40e0. The download_files field contains a quoted empty string. This does not establish that files are absent from all provider storage.

Observation and channel	Original-run binding
#11 / DFT-OptB88vdW	UNVERIFIED
#12 / DFT-TBmBJ	UNVERIFIED
#13 / DFT-OptB88vdW-Voigt-bulk	UNVERIFIED

Filename searches for OUTCAR, vasprun and 114724 found no primary calculation files in the accessible stores. Pilot mentions returned earlier dossiers. The four-entry attached owner ZIP was also inspected by filename. This bounded discovery does not cover every possible archive or replace original records.

Three levels are separated: byte integrity, consistency of declared relations, and independent provenance verification. PROV-DM distinguishes entities, activities and responsible agents [S1]; this implementation uses that distinction without claiming full PROV compliance.

science=ABSTAIN; DHC_chem=OPEN-NOT-CONFIRMED; EXT-RUN=NOT_RUN; truth_layer_promotion=0; T6-RA5=NOT_ISSUED; axiom_promotion=DISABLED.

02 / Two exact mathematical statements

Identifiability on observation fibers

Let P be a set of protocols, $\pi_i: P \rightarrow Y$ the available-information map, and $C: P \rightarrow \{0,1\}$ the target decision. A function $c: \pi_i(P) \rightarrow \{0,1\}$ satisfying $C = c \circ \pi_i$ exists if and only if C is constant on every fiber of π_i .

$$\pi_i(p) = \pi_i(q) \Rightarrow C(p) = C(q).$$

Proof. If $C = c \circ \pi_i$, equal images give equal decisions. Conversely, define $c(y) = C(p)$ for any p with $\pi_i(p) = y$. Constancy on the fiber makes this independent of the representative. This is the standard factorization criterion for a function, applied to protocol intake, not a claim of a new universal theorem.

Normalization counterexample: the same observed series with radius $r=0.002$ and tolerance $\varepsilon=0.001$ gives FAIL at $n=1$ and PASS at $n=2$. If π_i does not distinguish the normalizations, the observations cannot determine the decision. These numbers are synthetic; the example does not establish both protocols as physically admissible for the pilot.

Robust decisions with incomplete normalization

Let N be a nonempty finite set of positive integer divisors, $r \geq 0$ and $\varepsilon > 0$. Set $C(n) = [r \leq n\varepsilon]$. Then ALL_PASS is equivalent to $r \leq \varepsilon \min(N)$; ALL_FAIL is equivalent to $r > \varepsilon \max(N)$; otherwise the decision is AMBIGUOUS.

$$N = \{2,4\}: \text{ALL_PASS}; N = \{1\}: \text{ALL_FAIL}; N = \{1,2,4\}: \text{AMBIGUOUS}.$$

Proof follows from monotonicity of $n \mapsto n\varepsilon$. A decisive result persists under any nonempty restriction of N . An ambiguous result may become decisive when additional evidence excludes candidates. An empty N is rejected to prevent a vacuous PASS.

The conclusion is conditional: provenance, admissibility and exhaustiveness of N require separate justification. Even ALL_PASS establishes neither the actual normalization, asymptotic convergence nor a property-error bound. The finite radius refers to the first supplied point, not a sequence limit.

03 / Executable contract and rejection scenarios

The consistent synthetic packet contains 13 files and three ENCUT points. Seven common bindings identify structure, KPOINTS, potential identifiers, procedure, run log, extraction description and terms. Each point separately references INCAR, OUTCAR and a run identifier. No pseudopotential bytes are included.

Checks cover each file digest and size; atom counts in POSCAR/OUTCAR; ENCUT in INCAR/OUTCAR; output version; energy; and exact input/output digest associations in the run log. Distinct settings cannot reuse one output reference. Hashes do not authenticate authorship, time or execution.

Synthetic scenario	Outcome
Complete consistent packet	Provenance review required
Missing output	FILE_NOT_FOUND
Altered bytes	FILE_DIGEST_MISMATCH
#11 combined with TBmBJ	CHANNEL_METHOD_MISMATCH
Incorrect atom count	STRUCTURE_NATOMS_MISMATCH
One output for multiple ENCUT values	OUTPUT_REUSED_ACROSS_SETTINGS
Substituted log association	RUN_LOG_BINDING_MISMATCH
Extra stopping stage failed	BOTH_STOPPING_CHECKS_REQUIRED

A positive internal check returns `CONSISTENT_DECLARATIONS_REVIEW_REQUIRED`; `scientific_admission` always remains `BLOCKED`. The `RESEARCH` label itself authenticates nothing. Stopping flags are checked as declarations; the underlying tests still require review.

Supported adapters are restricted POSCAR with explicit species and coordinates [S2], an unambiguous ENCUT, and the final free-energy TOTEN in OUTCAR [S3]. Zero-temperature energy, vasprun XML and other formats are not silently converted: each needs an adapter grounded in actual samples.

04 / Reproduction, limits and next step

Checks: 35/35 unit tests and eight reproducible synthetic packets. Tests cover exact tolerance boundaries, joint unit/tolerance scaling, file/log substitution, channel mixing, path escape, symlinks, duplicate JSON and invalid numbers. Inputs stay unchanged; repeated reports are identical.

```
python3 -m unittest discover -s tests -v
```

Example: `python3 protocol_binding.py --packet fixtures/complete/packet.json --root fixtures/complete --out my_report.json`. The output file must be new. Python 3.11+ is required; the checker uses the standard library and executes no packet file.

Limits: method semantics, KPOINTS and potential correctness, log authenticity, successful execution, license scope and property convergence are not certified. Thirteen files do not establish scientific truth. Independent KLT-structure correspondence remains open.

B3.7: obtain the three original channel packets and independently review provenance; establish energy definitions and band-gap/elastic-property extraction; extend adapters only against obtained samples. No provider messages were sent. All 37 historical blockers remain; none removed. Training CLOSED; real forecasts NOT_RUN.

The active chemistry TXT is pinned in B3.5; another search for legacy A7 is unnecessary for continuation. RESEARCH is an additive bundle with selected evidence and the immutable B3.5 reference; it does not duplicate the full chemistry compilation. ReportLab generated the PDFs; XeLaTeX sources are supplied but were not compiled.

S1 · Moreau L., Missier P. (eds.). PROV-DM: The PROV Data Model. W3C Recommendation, 2013.

S2 · VASP Wiki. POSCAR: structure format and atom counts. Accessed 24 September 2026.

S3 · VASP Wiki. OUTCAR: output information. Accessed 24 September 2026.

S4 · NIST JARVIS-DFT. JVASP-114724: retained published record.

S5 · JARVIS-Tools. Pinned VASP workflow, commit 2c2241dbd47c7d3ec5f0f7a78e7231e7705b43f7.