

# The Causal Verification Framework: Distinguishing Laws from Artifacts in the Physical and Engineering Sciences

Ahmad Justanieah

[ajustanieah@kau.edu.sa](mailto:ajustanieah@kau.edu.sa)

King Abdulaziz University

---

## Article

**Keywords:** Data Correlations, Causal Verification Framework, Hierarchical Functional Causal Model, Causal Adjudication, Intervention Testing, Physical Laws, Statistical Phantom, Causal Stability, Multi-Process Correlation, Causal Skeleton, Enthalpy-Entropy Compensation

**Posted Date:** October 13th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-10919965/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# The Causal Verification Framework: Distinguishing Laws from Artifacts in the Physical and Engineering Sciences

**Ahmad M. Justanieah**  
*Chemical & Materials Engineering Department*  
*King Abdulaziz University*  
*Jeddah, Saudi Arabia*  
*ajustanieah@kau.edu.sa*

## Keywords

Data Correlations, Causal Verification Framework, Hierarchical Functional Causal Model, Causal Adjudication, Intervention Testing, Physical Laws, Statistical Phantom, Causal Stability, Multi-Process Correlation, Causal Skeleton, Enthalpy-Entropy Compensation

## ABSTRACT

Adjudicating whether observed global correlations in multi-level physical systems represent genuine causal laws or emergent statistical artifacts is a fundamental challenge across engineering and applied sciences. Such correlations underpin scaling laws, kinetic models, and materials design rules, yet this ambiguity compromises reliable prediction and design. This work introduces the Causal Verification Framework (CVF), a novel analytical methodology that resolves this ambiguity by testing observed correlations against known, low-level physical laws. The CVF inverts standard causal inference from identification to verification. Formalized as a Hierarchical Functional Causal Model (HFCM), the CVF employs targeted analytical intervention tests of stability (invariance under compositional change) and autonomy (invariance under parametric perturbation) to diagnose statistical phantoms — correlations that are contingent, non-autonomous artifacts of aggregation. The framework delivers a causal resolution to the long-standing Enthalpy-Entropy Compensation (EEC) debate, demonstrating the global correlation to be such a phantom. A protein denaturation case study illustrates the CVF's diagnostic power. The CVF provides a principled methodology for causal verification with applications across catalyst screening, process optimization, and materials discovery.

## Introduction

The challenge of distinguishing causal physical laws from statistical artifacts in aggregated data manifests acutely in engineering and applied science practices. From catalyst screening and scaling laws<sup>1, 2</sup>, kinetic modeling and compensation effects<sup>3-9</sup>, thermodynamic property prediction<sup>10-12</sup>, materials property correlations<sup>13</sup>, to astrophysical correlations where apparent trends may reflect

artifacts of data aggregation rather than underlying physical principles<sup>14</sup>, scientists and engineers routinely identify correlations in aggregated datasets to extract design rules. However, the leap from correlation to causation remains perilous — a fact evident from careful consideration of the aforementioned literature across disciplines. A pattern observed in aggregated data may reflect a unifying physical principle, or it may be a statistical phantom: a correlation that is contingent, non-autonomous, and arises as an artifact of data aggregation rather than a genuine causal law. Resolving this ambiguity requires more than statistical diagnostics; it demands a formal causal adjudication framework that tests observed patterns against known laws.

This challenge lies at the heart of causal inference: distinguishing genuine causal relationships from mere associations. Causal inference has been instrumental in uncovering unknown causal mechanisms in physics and other sciences. Existing causal inference methods address the problem of identification (discovery) — 'discovering equations, laws, and principles that are invariant, robust, and causal'<sup>15</sup>. Statistical diagnostics, by contrast, address the problem of association: assessing whether a correlation is statistically robust. However, neither addresses the complementary problem of verification: given a known causal skeleton  $F$ , can we determine whether an observed statistical summary  $O$  is a genuine autonomous law or merely a derivative phantom?

The Verification Paradigm introduced here fundamentally inverts the logic of standard causal inference: rather than searching for unknown causal structures in data, it begins with known causal skeletons and asks whether an observed pattern is autonomous from those skeletons — or merely a derivative phantom. This is the missing step that precedes discovery: first adjudicate whether the pattern is even a candidate for being a law.

The Causal Verification Framework (CVF) operationalizes this paradigm through a conceptual synthesis: use the known to adjudicate the observed. It poses the verification question — "Given a known causal skeleton  $F$ , is an observed correlation  $O$  a genuine autonomous law or a derivative phantom?" — and answers it through a formal architecture (the Hierarchical Functional Causal Model,  $HFCM \langle U, V, F, O \rangle$ ) and analytical intervention-based tests of Stability and Autonomy.

The CVF architecture is built, generalized through abstraction, then illustrated through a constructive demonstration using the Enthalpy-Entropy Compensation (EEC) phenomenon — a long-standing controversy in thermodynamics defined by the apparent linear correlation  $\Delta H = \beta_0 + \beta_1 \Delta S$  observed in aggregated thermodynamic datasets. The causal skeleton for the EEC is established by thermodynamics: for a family of transformations  $\psi_k$  (induced change between initial and final states, see Appendix B), the structural law is  $dh_k/ds_k = \tau$ , where  $\tau$  is temperature, which provided a geometric resolution: global linearity in the  $(\Delta S, \Delta H)$  plane is geometrically constructible. This geometric diagnosis establishes the null model<sup>16</sup>, but it cannot adjudicate whether a specific observed correlation is a phantom or a law. The CVF provides the missing causal logic to test that hypothesis and deliver a structural verdict. Crucially, the CVF does not presume the outcome. It provides formal principles and tools to test whether a given pattern is autonomous from the known skeleton, delivering a categorical verdict: H0 (phantom) or H1 (genuine law). As the culminating part of this constructive demonstration,

71 application to the EEC delivers a causal verdict. Whether the verdict confirms a physical law or  
 72 demonstrates the correlation to be a statistical phantom, the result validates the framework's capacity  
 73 to resolve such intricate ambiguities.

## 74 Problem Formulation

75 Aggregated data in scientific analysis naturally possesses multi-level structures that translate to global  
 76 dataset versus population subsets (distinct processes). This feature is fully exploited in subsequent  
 77 analysis. For thermodynamic datasets, Appendix B formalizes this structuring along continuous  
 78 thermodynamic coordinates (e.g., temperature, pressure) which shape the underlying thermodynamic  
 79 subspace (e.g., the  $\Delta S$ - $\Delta H$  plane for the EEC), in addition to discrete coordinates such as molecular  
 80 structural differences embodied within transformation processes. Building upon this foundational  
 81 structure, the corresponding problem statement is now formulated:

### 82 Problem Statement P1

83 Consider a population composed of  $K$  known subpopulations  $S_1, S_2, \dots, S_K$ . For each subpopulation  $S_k$ ,  
 84 we have a set of triplets  $(s_{k,i}, h_{k,i}, \tau_i)$  obtained by sampling the continuous confounder  $\tau \in [\tau^l, \tau^u] >$   
 85 0 at discrete points  $\tau_i$ . For each subpopulation  $S_k$ , the following axioms hold:

86 **Axiom 1 (Structural Law):** For each subpopulation  $S_k$ , there exists a function  $h_k = f_k(s_k)$  such that:

$$87 \quad \frac{dh_k}{ds_k} = \tau$$

88 This implies each  $f_k$  is strictly increasing and differentiable, and consequently, that a strong positive  
 89 correlation between  $h$  and  $s$  exists within each subpopulation  $S_k$ .

90 **Axiom 2 (Confounding):** For each subpopulation  $S_k$ , there exists a continuous, subpopulation-specific  
 91 structural function  $c_k(\tau)$  such that:

$$92 \quad h_k(\tau) = h_{k,0} + \int_{\tau_{k,0}}^{\tau_i} c_k(\tau) d\tau$$

$$93 \quad s_k(\tau) = s_{k,0} + \int_{\tau_{k,0}}^{\tau_i} \frac{c_k(\tau)}{\tau} d\tau$$

94 where  $\tau_{k,0}$ ,  $h_{k,0}$ , and  $s_{k,0}$  are subpopulation-specific constants.

95 **Axiom 3 (Consistency):** For each subpopulation  $S_k$ , the structural function  $c_k(\tau)$  and constants satisfy  
 96 the Structural Law (Axiom 1). This is automatically satisfied via the chain rule:

$$97 \quad \frac{dh_k}{ds_k} = \frac{dh_k/d\tau}{ds_k/d\tau} = \frac{c_k(\tau)}{c_k(\tau)/\tau} = \tau$$

98 **Axiom 4 (Pooling Operation):** The global dataset is  $\mathcal{D} = \bigcup_{k=1}^K \{(s_{k,i}, h_{k,i}, \tau_i)\}$ .

99 **Axiom 5 (Observed Global Correlation):** The pooled dataset  $\mathcal{D}$  yields a pronounced global linear  
100 pattern. This pattern is reflected in a Pearson correlation coefficient substantially different from zero,

101 
$$\rho_{\text{global}} = \frac{\sum_{k=1}^K \sum_{i=1}^{n_k} (s_{k,i} - \bar{s})(h_{k,i} - \bar{h})}{\sqrt{\sum_{k=1}^K \sum_{i=1}^{n_k} (s_{k,i} - \bar{s})^2 \cdot \sum_{k=1}^K \sum_{i=1}^{n_k} (h_{k,i} - \bar{h})^2}}$$

102 and, correspondingly, in an ordinary least-squares regression model

103 
$$\Delta H = \beta_0 + \beta_1 \Delta S$$

104 with a fitted slope  $\beta_1$  substantially different from zero.

105 The central challenge is to **adjudicate the causal status** of this observed pattern  $(\rho_{\text{global}}, \beta_1)$ :

106 Does it represent a **genuine structural relationship** — a latent uniform causal structure that governs  
107 the global  $s$ - $h$  relationship consistently across all subpopulations — or is it **demonstrably a statistical**  
108 **phantom** — a spurious artifact emerging solely from the aggregation of distinct subpopulations, each  
109 obeying the structural law  $dh_k/ds_k = \tau$  but lacking any unifying cross-population causal link?

110 The formulation of the MLM serves as the cornerstone for a causal resolution. By utilizing its multi-level  
111 construct, it translates the phenomenological notion of "compensation" into a rigorous causal target:  
112 the global correlation  $\rho_{\text{global}}$ . This translation reframes the entire controversy into a single, testable  
113 question: does  $\rho_{\text{global}}$  correspond to a genuine physical relationship, or is it an inevitable byproduct of  
114 aggregating data that conforms to known thermodynamic laws?

115 To assess the status of  $\rho_{\text{global}}$ , one might naturally turn to the standard statistical toolkit. Methods such  
116 as hierarchical regression, mixed-effects models, and partial correlation are designed to diagnose  
117 whether an observed correlation could be explained by confounding or aggregation artifacts. These are  
118 the appropriate instruments for diagnosing whether an observed pattern might be spurious. They  
119 operate at Pearl's Rung 1 (Association): they assess statistical robustness, not causal legitimacy.  
120 Applied to this formulation, they answer a specific diagnostic question: "Is this dataset's pattern  
121 consistent with an aggregation artifact — does it merely reflect systematic differences between the  
122 subpopulations  $S_k$ , rather than a real hidden relationship within them?" However, this diagnostic  
123 exercise can only test consistency with the artifact hypothesis; it cannot adjudicate whether the  
124 observed correlation is a genuine causal law or a statistical phantom. Consequently, these tools cannot  
125 answer the foundational question: "Can any dataset structured by the axioms of the MLM — where each  
126 point obeys known structural laws — ever produce a  $\rho_{\text{global}}$  that constitutes a new physical law?"

127 The gap between these questions is not merely methodological but ontological — it concerns the  
128 fundamental nature of the correlation itself. A dramatic instance of it is embodied in Simpson's Paradox  
129 (SP), the statistical phenomenon where within-group trends reverse upon pooling. SP does not explain

the gap but exposes it in its most striking form: it is a possible empirical outcome of the causal underdetermination inherent in the system. Critically, SP is not an anomaly — it is one symptom of the same underlying condition. The degeneracy and arbitrary realizability established in the PTF<sup>16</sup> permit any linear pattern in the pooled data. A spuriously positive global correlation (EEC), a spuriously negative one (AEEC), or a complete reversal (SP) are all statistical phantoms born from the same geometric reality. In fact, EEC and AEEC can manifest within selected subpopulations of a grand population and not the whole thermodynamic universe. Thus, SP serves as the definitive warning that statistical tools, which can only measure the manifestation of this degeneracy in a given dataset, are insufficient to judge whether a given pattern transcends it to achieve legitimate causal status.

Crucially, the MLM elucidates a universal structure that encompasses the core challenge discussed in the Introduction. In physical science, the existence of axiomatic constructs similar to Axioms 1, 2 or both is almost inevitable. This makes subsequent development generic and extensible as will be demonstrated in the sections that follow.

## The Causal Verification Paradigm

The preceding discussion emphasizes the absence of a verification framework capable of adjudicating questions like the EEC. To close this gap, a framework that begins not with the data, but with the enforceable causal skeleton is needed — one that provides formal criteria to test whether an observed statistical pattern  $\rho_{\text{global}}$  can legitimately claim causal autonomy within that causal universe. The following section introduces such a framework: the **causal verification framework**, a novel extension of causal inference designed to deliver the missing verdict.

| Primary Goal                   | Discover                                                       | hidden | causal             | Adjudicate the causal legitimacy of an observed correlation                                 |
|--------------------------------|----------------------------------------------------------------|--------|--------------------|---------------------------------------------------------------------------------------------|
|                                | mechanisms                                                     |        |                    |                                                                                             |
| Theoretical Foundation & Input | Data hypotheses; theory provides candidate models to test.     |        | generates provides | Theory provides a fixed causal skeleton; data provides the pattern to be judged against it. |
| Causal Question                | What latent structure could have produced this pattern?        |        |                    | "Is this pattern a genuine independent mechanism, or a consequence of the skeleton?"        |
| Level of Analysis              | Seeks a single, consistent causal model for the entire dataset |        |                    | Distinguishes local (causal) from global (artifact) levels                                  |
| Outcome                        | A proposed causal graph                                        |        |                    | A true/false verdict on causal legitimacy                                                   |

|                                 |                                                                                   |
|---------------------------------|-----------------------------------------------------------------------------------|
| <b>Abstract</b>                 | The Pearl's <i>M</i> model definition <b><i>HFCM</i> (Definition 1 hereunder)</b> |
| <b>Mathematical Formulation</b> | (Pearl 2009, §7.1.1, p203) <sup>17</sup>                                          |

151 *Table 1: Identification (General) vs. Verification (This work) Paradigms of Causal Inference Domain of*  
152 *Applications*

153 Thus far, analysis has shown that the observed correlation aligns with the known behavior of  
154 underdetermined data, and its potential reversal (Simpson's Paradox) confirms that anti-compensation  
155 is another manifestation of the same underlying condition. However, a **formal theoretical dismissal** of  
156 a 'genuine origin' — a proof that the pattern cannot function as a causal law — **has not been rigorously**  
157 **established.**

158 The required logic represents an **inversion of the standard causal paradigm**. Conventional causal  
159 inference seeks to *discover* latent mechanisms from patterns in observational data. Here, I begin  
160 with **known, enforceable causal mechanisms** — the thermodynamic laws and structural equations  
161 definitively established in Problem Statement P1.

162 I then employ Pearl's foundational causal logic — the principles of modularity and intervention<sup>17</sup> — not  
163 for discovery, but for **verification**. The central question becomes: does a robust statistical pattern  
164 observed at an aggregate level (the global correlation  $\rho_{\text{global}}$ ) possess genuine causal status, or is it an  
165 emergent artifact — a *statistical phantom* arising from specific selections within a degenerate  
166 thermodynamic subspace?

167 This 'inverse' verification paradigm is uniquely enabled by the exact, deterministic nature of physical  
168 laws, which provide a fixed causal skeleton. In this framework, the 'association' rung of Pearl's ladder is  
169 bypassed; starting with known structural relationships. Table 1 summarizes this shift from the  
170 standard **identification paradigm** to the proposed **verification paradigm**.

171 To address the challenge stated in the Introduction and exemplified by the EEC Multi-Level Model (P1),  
172 I now establish the complete Causal Verification Framework (CVF).

## 173 The Causal Verification Framework (CVF)

174 The Causal Verification Framework operationalizes the paradigm through a formal architecture and  
175 intervention-based logic. The following subsections define the framework's components, workflow,  
176 execution modes, domain positioning, and scope conditions.

### 177 The Causal Model Within the Verification Paradigm

178 To construct the causal verification framework, the analysis begins with the foundational principle upon  
179 which causal logic is built: **Autonomy**—the conception that individual causal mechanisms are  
180 independently alterable. To establish mathematical rigor, this principle is first formalized as  
181 the **Modularity Postulate**. It is then clarified how this postulate is not an external assumption but a

182 direct and necessary consequence of the Multi-Level Model's (MLM) structure, which inherently  
183 separates deterministic laws from the statistical patterns they generate.

184 **Modularity (Autonomy) Postulate:**

185 *Causal mechanisms are modular. In a functional causal model, each structural equation represents an*  
186 *autonomous mechanism; it is conceivable to change one such equation without necessitating a change*  
187 *in the others.*

188 This postulate formalizes Pearl's foundational assumption that "each parent–child relationship in the  
189 network represents a stable and autonomous physical mechanism" (Pearl, 2009, p. 22)<sup>17</sup>. Within the  
190 Multi-Level Model (Problem P1), however, modularity is not an external assumption but an inherent  
191 property of its construction.

192 The model is built upon the existence of Structural Law (e.g., Axiom 1 within the MLM), which causally  
193 governs the local (intra-subpopulation) relationship. This is a contract provided by underlying physics.  
194 The global correlation  $\rho_{global}$  is a separate, emergent statistical property (statistical summary)  
195 calculated from the pooled outputs of these local laws. This creates a **causal hierarchy**: a set of  
196 autonomous laws, denoted  $F$  hereafter, operates at the level of individual processes  $S_k$  (populations),  
197 while a set of any observed global pattern, denoted  $O$  hereafter, exists at the aggregate level.

198 Consequently, the applicability of Pearl's causal logic (e.g., interventions) to the framework is not  
199 predicated on assuming modularity; it is a **direct and necessary consequence** of the model's  
200 integrated, two-tiered architecture. This inherent modularity is precisely what provides the formal basis  
201 for the verification paradigm, allowing us to treat any global correlation  $O_j$  as a distinct entity whose  
202 causal status can be tested against the fixed local law  $F$ .

203 With the above established property of the MLM, formal causal definitions for the Multi-Level Model are  
204 now introduced:

205 **Definition 1. Hierarchical Functional Causal Model (HFCM)**

206 *A **Hierarchical Functional Causal Model (HFCM)** is a quadruplet*

208 
$$HFCM = \langle U, V, F, O \rangle,$$
  
207 *where:*

209 *(i)  $U$  is a set of **background (exogenous) variables**, determined by factors external to the model.*

210 *(ii)  $V = \{V_1, V_2, \dots, V_n\}$  is a set of **endogenous variables**, each determined by other variables*  
211 *within the model—that is, by variables in  $U \cup V$ .*

212 *(iii)  $F = \{f_1, f_2, \dots, f_n\}$  is a set of **structural laws (functions)**. Each  $f_i$  is a mapping*

216 
$$f_i: PA_i \cup U_i \rightarrow V_i,$$
  
213 *where  $PA_i \subseteq V \setminus \{V_i\}$  are the parents of  $V_i$  within  $V$ , and  $U_i \subseteq U$ . The entire set  $F$  forms*  
214 *a **deterministic mapping** from the combined domain of exogenous variables and parent sets to*  
215 *the endogenous variables:*

218 
$$F: PA \cup U \rightarrow V,$$
  
217 *where  $PA = \bigcup_{i=1}^n PA_i$  is the union of all parent sets, and  $dom(F) = PA \cup U$ .*

219 *(iv)  $O = \{O_1, O_2, \dots, O_m\}$  is a set of **statistical summaries**:*

220

$$O_j = SS(PA_{O_j}; \Theta_j), PA_{O_j} \subseteq V$$

221

where  $PA_{O_j}$  is the subset of endogenous variables used for the  $j$ -th statistical summary,  $\Theta_j$  is the set of parameters that define the summary (e.g., regression coefficients, correlation coefficients, fitted curve parameters), and  $SS(\cdot; \Theta_j)$  is a function that computes the statistical summary from the realized values of  $V_{O_j} \subseteq V$  parameterized by  $\Theta_j$ .

222

223

224

225

226

227

228

229

230

231

232

233

The definition of the *HFCM* creates the foundational construct for implementing causal logic within the verification paradigm. It separates the true causal laws ( $F$ ) from the statistical summaries (observed  $O$ ) **by design**. This abstraction establishes a clear distinction from standard causal identification models and provides a generalized framework for handling both continuous and discrete/underdetermined coordinates in systems governed by underlying causal relationships. The core logic becomes straightforward: genuine,  $F$ -parallel relationships within  $O$  must be consistent with—and operate within the boundaries defined by—the established physical laws  $F$ . This leads thereby to the **first fundamental principle** of the verification framework, the principle of **Structural Compliance (SC)**:

234

#### **Structural Compliance Principle:**

235

236

237

238

A statistical summary  $O_j \in O$  can be considered a **genuine**, operating in parallel to  $F$  (**F-parallel structural relation**) only if (necessity) it is modular and **structurally compliant** with  $F$ . Compliance requires that  $O$  demonstrates autonomous, law-like behavior while remaining consistent with the constraints imposed by  $F$ .

239

240

241

242

243

244

245

246

247

248

249

250

251

Crucially, the SC demands two necessary conditions for  $O$  to claim causal status: (i) Modularity; (ii) Consistency with  $F$ . The necessity of modularity arises because causal relationships possess greater stability than probabilistic associations due to their ontological nature (Pearl, 2009, §1.3.2)<sup>17</sup>. As Pearl states, “Therefore, causal relationships should remain unaltered as long as no change has taken place in the environment, even when our knowledge about the environment undergoes changes.” The necessity of consistency with  $F$  arises because of the causal status of  $F—O$  should never violate  $F$ . **This epistemic robustness—the intersection of these two necessities established by the SC—is interpreted within the verification paradigm as causal stability:** a genuine cause-effect relationship must be stable not only under changes in knowledge but also in its very existence when the underlying structural laws are directly probed. This rationale leads to the **formalization of the second core principle of the verification framework, the Stability (Invariance) Principle**, testing the behavior of  $O$  under permissible interventions that evaluate the necessary conditions established by the SC principle:

252

#### **Stability (Invariance) Principle:**

253

254

255

256

257

258

Consider the *HFCM* (definition 1) and let contextual intervention:  $do(U_{sel} = u')$  be interventions that alter the exogenous context  $U$  but leave the functional component of  $F$  unchanged and structural interventions:  $do(I = i')$  be interventions that alter a fundamental input variable of  $F$  and thereby alter the functional response of  $F$  itself. Then, a statistical summary  $O_j$  can be considered causally stable structural relation only if it remains invariant under both types of interventions—the necessities of the SC principle.

259 **Structural Intervention** targets autonomy of  $O$ —invariance under skeletal evolution. It acts as  
 260 a *mechanistic probe of the skeleton*. This could be achieved by performing a series of  $do(I = i')$ —  
 261 changing the confounding variable in the system (e.g. temperature or pressure in a thermodynamic  
 262 system) while observing the status of  $O$ . It is to be noted here that nature of  $I$  is intentionally left  
 263 underdetermined as an input of  $F$  for generality; it may be an exogenous variable in  $U$  or an endogenous  
 264 variable in  $V$ , depending on the specific structure of the skeleton. The only requirement is that  $I$  induce  
 265 a functional response of  $F$ . **Contextual Intervention**, in contrast, acts as a *contextual probe*, targeting  
 266 the universality of  $O$ —invariance under ensemble changes of exogenous contexts (e.g., the selection of  
 267 transformation series  $S_k$ ). Since both are necessary, simultaneous execution of both types is also  
 268 permissible.

269 The two principles are complementary: the **Structural Compliance Principle** establishes the  
 270 necessary relationship a correlation must have with the foundational law  $F$ , while the **Stability**  
 271 **(Invariance) Principle** provides the operational toolkit to test for that relationship. This toolkit is built  
 272 upon the classification of interventions, which directly reflects the hierarchical nature of the Multi-Level  
 273 Model. This provides a rigorous path from the abstract *HFCM* to concrete causal verdicts. It cleanly  
 274 separates the investigation of autonomous physical mechanisms from the exposure of spurious,  
 275 aggregation-dependent correlations. In essence, the framework operates on a principle of causal  
 276 charity: any observed pattern  $O_j$  is treated as a legitimate candidate for causal status, but  
 277 must **earn** that status by demonstrably satisfying the necessary conditions imposed by the principles  
 278 of Structural Compliance and Stability.

## 279 The CVF Workflow

280 This section describes the proposed CVF workflow. It begins by formalizing two competing  
 281 hypotheses—H0 (phantom) and H1 (genuine law). It then states the inference rule and its quantification.  
 282 The CVF logical verdict is grounded in the previously established principles grounded in modularity and  
 283 interventional testing of causal inference.

284 Two competing hypotheses are defined for any statistical summary  $O_j \in O$ , where  $O_j =$   
 285  $SS(PA_{O_j}; \Theta_j)$  and  $PA_{O_j} \subseteq V$ .

286 **H0 (Statistical Phantom):**  $O_j$  is a derivative echo of the known skeleton  $F$ . The system is described  
 287 solely by the map  $F: U \cup PA \rightarrow V$ . In H0,  $O_j = SS(PA_{O_j}; \Theta_j)$  is determined by  $F$  and has no independent  
 288 causal status.

289 **H1 (Genuine Structural Law):**  $O_j$  is a genuine cross-population structural law, operating in parallel to  $F$ .  
 290 It is autonomous—modularly independent from  $F$  and consistent with it. If H1 holds, the system is  
 291 described by an additional causal map  $O_j: PA_{O_j} \cup \Theta_j \rightarrow V_{O_j}$ , where  $PA_{O_j}$  are additional parents  
 292 of  $V$  operating through  $O_j$  at the aggregate level, and  $\Theta_j$  are the parameters of the statistical summary,  
 293 which, within H1, function as a cross-population causal structure. The combined system is then:

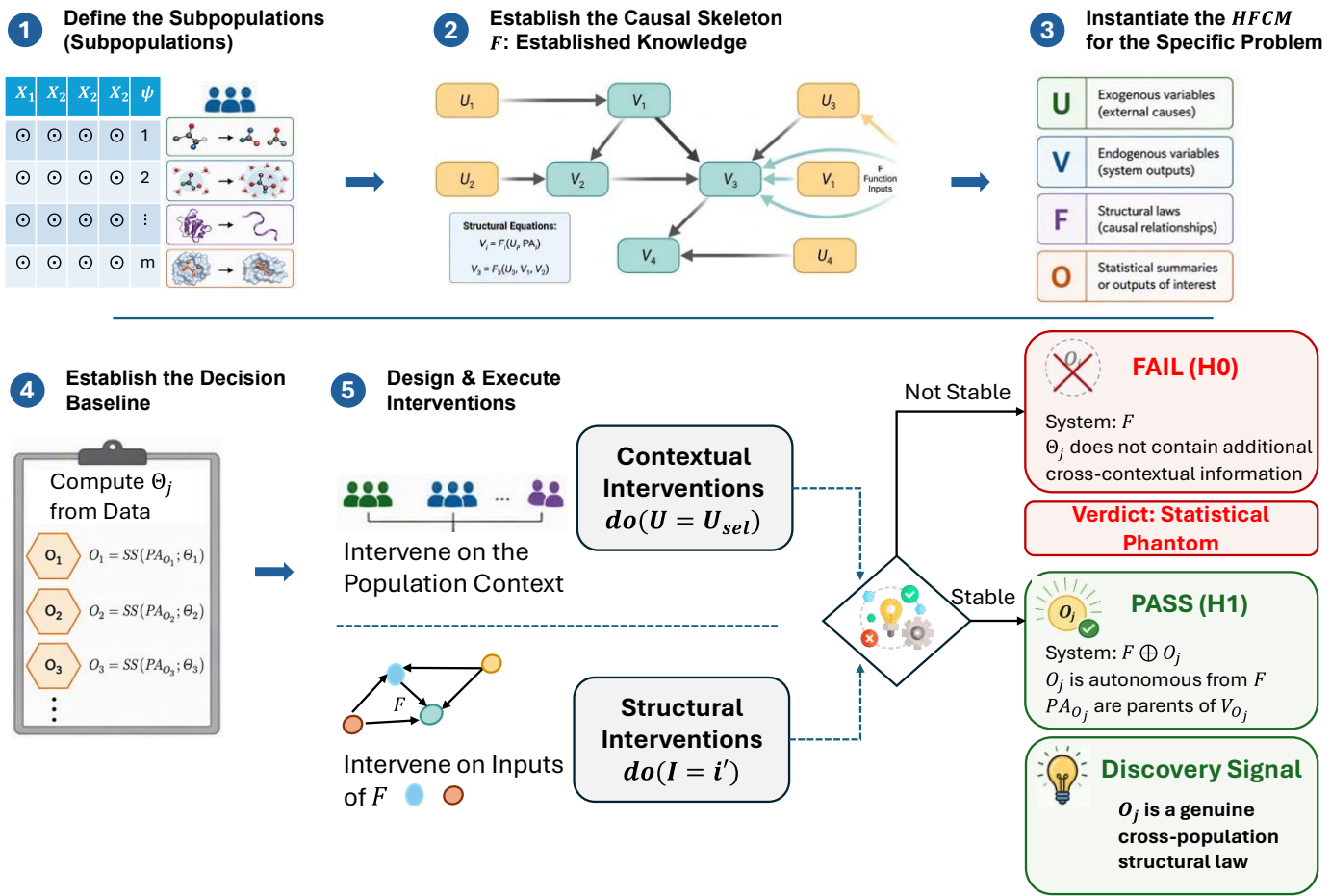
$$295 \quad F \oplus O_j: U \cup PA \cup PA_{O_j} \rightarrow V$$

294 where  $\oplus$  denotes modular independence.

296 Accepting or rejecting the hypothesis is completely contingent on executing the interventions.

297 If  $O_j$  fails to remain invariant under structural interventions ( $do(I = i')$ ) or contextual interventions  
 298 ( $do(U_{sel} = u')$ ), then  $O_j$  is not autonomous from  $F$  and is therefore a statistical phantom—H1 is rejected  
 299 and H0 holds. If  $O_j$  survives both interventions, it is structurally compliant with  $F$  and is a candidate  
 300 genuine law—H1 is accepted. The success or failure of interventions can be quantified by comparing  
 301 the behavior of  $\Theta_j$  under different intervention settings. A **structurally significant change** in  $\Theta_j$ —one  
 302 that exceeds the expected bounds of experimental uncertainty—indicates failure of invariance. No such  
 303 change indicates success: for structural interventions, compare  $\Theta_j(do(I = i'_1))$  versus  $\Theta_j(do(I = i'_2))$ ;  
 304 for contextual interventions, compare  $\Theta_j(do(U_{sel} = u'_1))$  versus  $\Theta_j(do(U_{sel} = u'_2))$ .

305 When H1 holds, this constitutes a Discovery Signal:  $O_j$  is accepted as a structural law and the system  
 306 is described by  $F \oplus O_j$  as in H1 above. Clearly, the verdict is conditional on the known skeleton  $F$  and  
 307 the parameterization of  $O_j$  through  $\Theta_j$ . If new knowledge supersedes  $F$ , the CVF can be re-applied. The  
 308 complete CVF workflow is depicted in Figure 1.



309  
 310 **Figure 1 The CVF Workflow.** The figure summarizes the CVF workflow, from defining subpopulations  
 311 and establishing the causal skeleton  $F$ , to executing interventions and delivering a structural verdict:  
 312 phantom (H0) or genuine law (H1).

## 313 Analytical Execution of Interventions

314 The CVF's interventions — structural ( $do(I = i')$ ) and contextual ( $do(U_{sel} = u')$ ) — are executed  
 315 analytically by exploiting the known functional form of  $F$  for structural interventions and the functional

316 form of  $O_j = \text{SS}(\text{PA}_{O_j}; \theta_j)$  for contextual interventions. This is the strongest form of interventional  
317 execution. The associated verdicts are universal with respect to similar datasets governed by  $F$  and the  
318 execution assumptions. Crucially, the dataset size does not matter when the dataset is assumed to  
319 comply with  $F$  within experimental uncertainty or numerical errors for numerically generated datasets.

320 The analytical execution of interventions yields structural verdicts — categorical conclusions that are  
321 robust relative to the known skeleton  $F$ . The verdict does not depend on dataset size, sampling variation,  
322 or statistical thresholds. It follows from the logical structure of the CVF Inference Rule and the known  
323 functional form of  $F$  and parameterization of  $O_j$ . The incorporation of experimental uncertainty is  
324 handled through the criterion of "structurally significant change" in  $\theta_j$ , as defined in the CVF Workflow  
325 section.

326 The distinction establishes why the CVF does not require statistical sensitivity analyses: verdicts are  
327 structural, not empirical. The CVF operates at Pearl's Rung 2 (Intervention), where interventions are the  
328 ultimate test of modular autonomy. Statistical sensitivity analyses belong to Rung 1 (Association); they  
329 assess robustness, not autonomy.

330 Crucially, the CVF answers a Rung 2 question by deploying a Rung 2 logic. In essence, the logic of the  
331 Intervention level is used to vet a claim made at the Association level.

## 332 The CVF within the Causal Inference Domain

333 The CVF is a causal inference methodology for causal verification. It complements existing statistical  
334 methods for causal identification by addressing a distinct problem: adjudicating whether an observed  
335 correlation is autonomous from a known causal skeleton  $F$ .

336 To distinguish methodologies, one must consider three aspects: the question being investigated, the  
337 method of investigation, and the rung on Pearl's Causal Hierarchy (PCH). PCH defines three layers—  
338 Association (Rung 1), Intervention (Rung 2), and Counterfactuals (Rung 3)<sup>18</sup>—each corresponding to a  
339 distinct type of question.

340 **Statistical identification methods**—such as Invariant Causal Prediction (ICP), the PC algorithm with  
341 Momentary Conditional Independence (PCMCI), and the Linear Non-Gaussian Acyclic Model  
342 (LiNGAM)—search for causal structure<sup>15</sup>. ICP searches for direct causal parents by testing whether the  
343 conditional distribution  $Y \mid X_S$  is invariant across different environments. PCMCI and LiNGAM  
344 algorithmically infer the time-series graph or DAG from data using conditional independence tests  
345 (with time-order) and non-Gaussianity, respectively. All three methods answer Rung 2 causal  
346 questions ("What causes  $Y$ ?") by leveraging strong, testable assumptions to discover structures from  
347 observational data without physically performing new experiments.

348 **Statistical diagnostics**—hierarchical regression, mixed-effects models, and partial correlation—  
349 operate at Rung 1 (Association). They assess whether a correlation is robust under sampling variation,  
350 model specification, or covariate adjustment. Even methods that seek invariant representations  
351 across environments, such as Invariant Risk Minimization (IRM), remain at Rung 1: they identify

352 statistical invariances that support robust prediction, not causal guarantees. These methods do not  
 353 perform interventions; they operate on data alone.

354 **The CVF** addresses a different Rung 2 question: "*Is  $O_j$  a phantom or a law?*" Rather than searching for  
 355 unknown structures, it executes actual interventions (*do*-operations) on the known causal  
 356 skeleton  $F$  and the statistical summary  $O_j$ . Its contract is physical:  $F: U \cup PA \rightarrow V$  is known,  
 357 enforceable, and provided by physics.

358 From a statistical perspective, the CVF addresses a limitation of invariance-based methods such as  
 359 ICP and IRM. These methods identify causal parents by testing invariance across environments.  
 360 However, invariance can arise as a mathematical consequence of known structural relationships—not  
 361 as evidence of causality. The CVF complements these methods by testing the autonomy of a  
 362 statistical summary from the known structural laws themselves, providing a new criterion for  
 363 distinguishing genuine laws from artifacts.

364 The epistemological contrast is summarized in Table 2: identification methods discover unknown  
 365 causal structures (algorithmic search, no intervention, empirical contract); statistical diagnostics  
 366 assess robustness (Rung 1); the CVF adjudicates known candidates (analytical *do*-operations,  
 367 physical contract, Rung 2).

368 The CVF is a complementary addition to the field of statistical methodology. It challenges a perceived  
 369 statistical summary  $O_j$  by leveraging a known skeleton  $F$ , qualifying modularity through structural and  
 370 contextual interventions grounded in  $F$ . It shares the foundational principle of autonomy with methods  
 371 such as ICP<sup>19</sup> and IRM<sup>20</sup>, but qualifies and operationalizes it through structural and contextual  
 372 interventions grounded in  $F$ . This represents a paradigm expansion to the field, addressing a  
 373 challenging statistical problem not articulated before as a problem of causal verification.

| <b>Examples</b> | ICP, PCMCI, LiNGAM                  | Hierarchical Regression, ANCOVA, Mixed-Effects, Omitted-Variable, IRM           | This work: the CVF                              |
|-----------------|-------------------------------------|---------------------------------------------------------------------------------|-------------------------------------------------|
| <b>Question</b> | "What causes Y?" (Rung 2)           | "Is this association robust?" or "What generalizes?" (Rung 1)                   | "Is $O_j$ a phantom or a law?" (Rung 2)         |
| <b>Outcome</b>  | Causal graph / Structural equations | Coefficient estimates, p-values, sensitivity metrics, invariant representations | Structural verdict ("phantom" or "genuine law") |

|                                          |                                                                                                                |                                                                                             |                                                         |
|------------------------------------------|----------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|---------------------------------------------------------|
| <b>Operation</b>                         | Algorithmic search over data                                                                                   | Statistical fitting, model comparison, optimization                                         | Analytical execution of <i>do</i> -operations           |
| <b>Intervenes?</b>                       | No                                                                                                             | No                                                                                          | Yes — on $I$ and $U$                                    |
| <b>Enabler</b>                           | Preestablished conditions (e.g. Markov/faithfulness, non-Gaussianity, or invariance) contingent on assumptions | Empirical / Model-based (assumes correct specification, exchangeability, or representation) | Physical Skeleton (known $F: U \cup PA \rightarrow V$ ) |
| <b>Can adjudicate causal legitimacy?</b> | No — discovers structural candidates from passive data                                                         | No — assesses robustness or predicts, does not adjudicate                                   | Yes                                                     |

*Table 2 Identification, Statistical Methods, and Verification—A Paradigmatic Comparison. The table contrasts identification, statistical methods, and verification across question type, outcome, operation, intervention status, enabler, and ability to adjudicate causal legitimacy.*

## The CVF Scope and Defining Conditions

The CVF operates in domains where a known, enforceable causal skeleton  $F$  exists. In the physical and engineering sciences—thermodynamics, kinetics, catalysis, materials science—such skeletons are almost always available and mathematically tractable. This is the enabling condition of the framework as stated previously.

The CVF rests on and is facilitated by the following conditions:

**Known Necessity of  $F$ :** The structural laws  $F$  are known, enforceable, and necessary partial generators of the measured variables  $V$ . This means  $F$  is satisfied for any valid data point.

**Structured Data:** Aggregated data correlations require the dataset  $\Gamma$  to be organized by well-defined subpopulations  $S_k$ , each corresponding to a distinct transformation process  $\psi_k$  (induced changes in physicochemical properties). The principled structuring of the data facilitates the tracking of global confounders specific to a given field (e.g., temperature and pressure in thermodynamics), which offer the most direct handle for structural interventions. However, even when such variables are not explicitly reported, the underlying skeleton is active and necessary—unexplicit reporting does not eliminate its skeletal contributions.

**Structural Interventions:** Structural intervention requires that the structural laws  $F$  are deterministic functions mapping inputs to outputs. This is characteristic of exact physical laws (e.g., thermodynamic identities, conservation principles). It is preferable that  $F$  describes the relationship between the known global confounders and the measured variables. This is not a requirement, however.

396 **Contextual Interventions:** Contextual interventions require field-specific realizability and  
397 manipulability of the context. For example, the distinct processes or experimental conditions  
398 constituting  $U_{\text{sel}}$  are identifiable and can be systematically intervened upon.

399 Crucially, the strength of the causal verdicts is tied to the strength of  $F$  and its validity. In domains  
400 where  $F$  is exact and unassailable (e.g., thermodynamics), the verdict is causally definitive. With  
401 approximations or assumed  $F$  (e.g., structural models in social science), however, the verdict is relative  
402 to  $F$ . The latter type is beyond the current scope of correlations in physical sciences—the focus of this  
403 work.

404 **Application: Enthalpy-Entropy Compensation**

|                            |                                                                                                             |                                                                                                                                                                                                                                                                                |                                                                                                                                                                            |
|----------------------------|-------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                            |                                                                                                             |                                                                                                                                                                                                                                                                                |                                                                                                                                                                            |
| <b>U (Exogenous)</b>       | Background variables determined outside the model                                                           | $\tau \in [\tau^l, \tau^u]$ : global continuous confounder (temperature)<br>$S_k$ : data subset corresponding to series member $k$<br>$\{h_{k,0}, s_{k,0}\}$ : series-specific parameters<br>$c_k(\tau)$ : group-specific functional dependency (the "heat capacity" function) | Represents all external factors and series specific structures (properties)                                                                                                |
| <b>V (Endogenous)</b>      | Variables determined within the model                                                                       | $s_k(\tau)$ : independent variable, specific to $S_k$<br>$h_k(\tau)$ : response variable, specific to $S_k$                                                                                                                                                                    | Core measured variables, whose functional relationship is the object of study. They are <i>generated</i> by the system ( $F$ ) from the context ( $U$ ).                   |
| <b>F (Structural Laws)</b> | Set of autonomous mechanisms $\{f_1, f_2, \dots, f_n\}$ where each $f_i$ maps from $PA_i \cup U_i$ to $V_i$ | The universal structural constraint:<br>$\frac{dh_k}{ds_k} = \tau$ This single law, parameterized by $U(S_k, \tau)$ , generates the specific functional forms for $h_k(\tau)$ and $s_k(\tau)$ within each subpopulation:                                                       | Encodes the universal physical law that generates the context-dependent variables in $V$ . All other associated thermodynamic relationships are implicitly reflected. (for |

|                                  |                                                                                               |                                                                                                                                                                                                                                                                                                                |                                                                                                    |
|----------------------------------|-----------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|
|                                  |                                                                                               | $h_k(\tau) = h_{k,0} + \int_{\tau_{k,0}}^{\tau_i} c_k(\tau) d\tau$ $s_k(\tau) = s_{k,0} + \int_{\tau_{k,0}}^{\tau_i} \frac{c_k(\tau)}{\tau} d\tau$                                                                                                                                                             | example, Gibbs equation, $s_k$ & $h_k$ functional forms...etc.)                                    |
| <b>O (Observed Correlations)</b> | Set of statistical patterns $\{O_1, O_2, \dots, O_m\}$ calculated from realized values of $V$ | $O_1$ : Global linear correlation $\rho_{global}$ from pooled data (alternatively, representable by $\beta_0$ (the intercept) and $\beta_1$ (compensation temperature))<br>$O_2$ : The correlation between subpopulation <i>means</i> , i.e., the ecological correlation (e.g., $Corr(\bar{s}_k, \bar{h}_k)$ ) | Represents the statistical summaries computed from the realized, context-dependent values of $V$ . |

Table 3 Instantiation of the Hierarchical Functional Causal Model for the EEC Multi-Level Problem (P1)

The objective here is to test the causal legitimacy of the Enthalpy-Entropy Compensation Effect using the formalism in Appendix B. Specifically, I employ: (i) the transformation processes  $\psi_i$  and the aggregate family  $\Psi = \{\psi_{k=1\dots m}\}$ —which defines the distinct populations within the CVF; (ii) the structured datasets  $\Gamma = \bigcup_{k=1}^m \Gamma_k$ ; and (iii) the Linear Regression Model (LRM),  $\Delta H = \beta_0 + \beta_1 \Delta S$ —which formally defines the EEC correlation under investigation.

Hence, the analysis begins by **instantiating the abstract HFCM**, mapping its elements directly onto the Multi-Level Model of Problem Statement P1. This contextual mapping, summarized in Table 3, provides the concrete foundation for applying the causal principles of **Structural Compliance** and **Stability**. The set  $O = \{O_1, O_2\}$  defines the two global correlations whose causal legitimacy is under investigation:  $O_1$  is the standard linear regression on the pooled data, and  $O_2$  is the *ecological correlation*, calculated from subpopulation averages and thus bypassing individual-level data entirely. It is to be noted that the within-group correlation is a deterministic manifestation of the structural law  $F$ , fully resolved fully resolved in the Predictive Thermodynamic Framework (PTF) as incidental geometric consequence of  $\frac{dh_k}{ds_k} = \tau^{16}$ , and is therefore not considered a subject for causal verification here—only aggregate level correlations are considered.

## Design of Interventions

With the *HFCM* instantiated for the EEC, the core challenge—determining whether a global correlation is genuine or spurious—must be addressed. As established in the The CVF Workflow section, this is achieved by crossing from Pearl’s first rung (Association) to the second (Intervention), beginning now with the design of Interventions.

The observed global correlation  $\rho_{global}$  (or the fitted parameter  $\beta_1$ ) is an association, a statistical pattern in the pooled data  $\mathcal{D}$ . The structural law  $F$  provides the causal skeleton. To complete this analysis

requirement, **Structural and Contextual Interventions** are defined. From these definitions, formal domain-specific, testable criteria for structural compliance are derived:

**Exclusion-Inclusion Intervention (EII):** The EII is a contextual intervention that directly alters the instantiations of the exogenous variables  $U$  by selectively including or excluding subpopulations  $S_k$  from the analyzed dataset. This intervention tests the foundational claim that a genuine correlation represents a **universal, cross-contextual causal phenomenon**. Thermodynamically, a global correlation purporting to be a genuine physical law must remain **invariant** to the composition of the specific series of transformations  $\Psi$  considered; its parameters (slope  $\beta_1$ , and intercept  $\beta_0$ ) should not change substantively if transformations are added to or removed from the pool. More fundamentally, for a correlation to represent a universal relationship, its **pairwise logic must be internally consistent**: if the correlation holds for the full set  $\Psi$ , it should also hold, to a significant degree, between any individual pair of processes  $(\psi_i, \psi_j)$  drawn from that set. A correlation that fails this test of compositional and pairwise invariance is exposed as context-dependent and cannot be considered a universal law.

**The logical implication is as follows:** If  $O$  changes under this intervention, H1 is rejected and H0 is confirmed—it is exposed as a contingent property of a specific data composition, not a cross-population structural law. Specifically:

- It demonstrates **context-dependence**, proving the correlation is not invariant across the population.
- It reveals the correlation as an **aggregation artifact**, whose existence is tied to a particular ensemble of groups rather than a fundamental causal mechanism.
- It fails the most basic test of **causal universality**, confirming its status as a statistical phantom.

**Global Confounder Intervention (GCI):** The GCI tests **autonomy from the foundational mechanism  $F$** . The Structural Compliance Principle specifies that for a correlation  $O_j$  to be considered a genuine, parallel causal law, its behavior under perturbations that alter the input of the functional component of  $F$  must be **governed by its own stable rule—captured by the parameterization  $\Theta_j$** .

**The logical implication is as follows:** If perturbing the global confounder  $\tau$  (e.g., temperature) destroys  $O_j$  or causes its behavior to violate the deterministic mandates of  $F$ , it proves  $O_j$  **lacks autonomous existence—i.e.** H1 is rejected and H0 is confirmed. It is not an additional causal law but a **fragile, contingent pattern**—a statistical phantom. Specifically, such a failure demonstrates that  $O_j$  is:

- **Not robust**, failing the basic test of autonomy.
- **Functionally dependent** on a specific window of  $F$ 's inputs.
- An **epiphenomenon**—a side-effect of the specific instantiation of the exogenous variables  $U$  that condition  $F$ .

These interventions therefore provide the means to test the CVF hypothesis space whether the observed EEC line  $O_1$ —**can be trusted as a causal structural relationship for the system it purports to describe—i.e. does it embody additional cause-effect information beyond what is given by  $F$ ?** If  $O_1$  were a genuine causal law (H1), it could, in principle, support predictions like: "if a process is manipulated to have a specific  $\Delta S$ , it will exhibit a corresponding  $\Delta H$  as dictated by the compensation line." Compactly, the **EII** qualifies the claim of EEC universality through **stability under compositional change**, while the **GCI** qualifies the nature of an apparently stable correlation by testing its **autonomy under mechanistic perturbation**. This duality is summarized in Table 4.

| Contextual | Exclusion-Inclusion (EII):<br>$do(U_{sel} = k')$ | Selective addition or removal of subpopulations $S_k$ (or linear combinations thereof) from the dataset. | $O_1, O_2$ | A genuine universal law must be invariant to the specific ensemble of processes used to observe it.                     | <b>Stable (H1):</b> Candidate genuine law.<br><b>Unstable (H0):</b> Statistical phantom (aggregation artifact).   |
|------------|--------------------------------------------------|----------------------------------------------------------------------------------------------------------|------------|-------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| Structural | Global Confounder (GCI):<br>$do(\tau = \tau')$   | Controlled perturbation of the intensive confounder $\tau$ (temperature).                                | $O_1, O_2$ | Determines if a correlation is an <b>autonomous law</b> or a <b>redundant manifestation</b> of the structural law $F$ . | <b>Autonomous &amp; Robust (H1):</b> Genuine parallel law.<br><b>Redundant/Fragile (H0):</b> Statistical phantom. |

Table 4 Intervention Designs for Causal Verification of the EEC

With respect to the GCI, it is beneficial to observe that the EEC relevant skeleton is very strong and restrictive. This is clearly observed through the confounding effect of the temperature that operates across the populations generating an obvious backdoor—a direct causal relationship between the pair  $s_k(\tau)$  and  $h_k(\tau)$  exists:  $\frac{dh_k}{ds_k} = \tau$ . This is also clear in MLM (Axiom 2) and common in systems described by physics. This makes the selection of  $\tau$  for structural interventions a natural and powerful choice for GCI analytical execution. This observation serves as a cautionary note: an obvious global confounder should never be ignored.

Notwithstanding the above observation, confounding is not a general requirement in the *HFCM* and every instantiation should be problem and/or domain specific.

## 481 Analytical Execution of EII: Contextual Vulnerability

482 The next step in the CVF's workflow is the execution of interventions to deliver the causal verdict,  
483 confirming either H1 or H0. The execution of interventions for the EEC is performed analytically. For EII,  
484 this is achieved by exploiting geometrical properties of planar datasets while contextualizing the action  
485 of EII within these geometric constraints. In contrast, GCI is executed by exploiting the functional  
486 properties of the skeleton and its mathematical implications.

487 **Exclusion-Inclusion Intervention (EII):** Before analyzing EII outcome for the EEC consider the following  
488 mathematical result with respect to Ordinary Least Square (OLS) of planar datasets:

### 489 **Theorem 1 (Geometric Vulnerability of OLS on Planar Data).**

490 *When ordinary least squares (OLS) is applied to any planar dataset whose convex hull is*  
491 *non-degenerate, the fitted correlation is geometrically vulnerable: its slope  $\hat{\beta}_1$  can be systematically*  
492 *destroyed or reversed by strategic inclusion or exclusion of points.*

### 493 **Proof:**

494 By **Lemma A.1**, there exists a segment inside the convex hull whose slope has sign opposite to  $\hat{\beta}_1$ .  
495 Adding many points along that segment (inclusion) or removing points except those near such a  
496 segment (exclusion) forces the OLS slope toward the opposite sign. ■

497 Theorem 1 constitutes the mathematical basis for analytical execution and interpretation of EII. It  
498 establishes a generic geometric guarantee that the OLS slope is vulnerable to manipulation. To this end,  
499 a specific qualification of this general theorem to thermodynamic state-space is required.

500 Crucially, the convex hull is physically and **thermodynamically realizable—thereby the line**  
501 **segments within the hull also**. To see this, consider two transformations  $\psi_i: A \rightarrow A'$  and  $\psi_j: B \rightarrow B'$ .  
502 Because enthalpy and entropy changes are extensive state functions, any point on a segment  
503 connecting two vertices can be realized by a well-defined process  $\psi_{ij} = \lambda\psi_i + (1 - \lambda)\psi_j, \lambda \in [0,1]$ ,  
504 whose initial and final states are convex combinations of the corresponding states of the endpoint  
505 processes  $\psi_i$  and  $\psi_j$ . Physically, this can be viewed as a mixing operation— $\lambda$  stoichiometric amount of  
506  $A$  plus  $(1 - \lambda)B$  stoichiometric amount of  $B$  are transformed to same stoichiometric ratios of  $A'$  and  $B'$   
507 (mass conservation):

$$508 \quad \psi_{ij}: \lambda A + (1 - \lambda)B \rightarrow \lambda A' + (1 - \lambda)B'$$

509 Thermodynamically and for any extensive state function  $X$ :

$$510 \quad \Delta_{\psi_{ij}} X = \lambda \Delta_{\psi_i} X + (1 - \lambda) \Delta_{\psi_j} X$$

511 Thus, the convex hull of the measured data points in the plane corresponds to the set of all physically  
512 realizable thermodynamic changes accessible by compositional mixing. The OLS slope computed from  
513 any dataset within this region is sensitive to the subset of points sampled, demonstrating that it is not a  
514 property of any single process but depends on the composition of the dataset.

515 Importantly,  $\psi_{ij}$  does not introduce any new mechanistic or molecular information; it is a  
516 thermodynamic interpolation of the existing processes. In fact, if the linear regression model were a  
517 genuine physical law, the data for  $\psi_{ij}$  would necessarily fall on the same global regression line as its  
518 parent processes—a condition that is easily violated under an Exclusion-Inclusion Intervention,  
519 revealing the regression’s fragility.

520 This realizability translates the abstract geometric vulnerability into concrete intervention strategies.  
521 An **inclusion attack** constructs an augmented dataset  $\mathcal{D}' = \mathcal{D} \cup \{(x_i, y_i)\}_{i=1}^m$  by adding  $m$  new points  
522 along the opposing-slope segment  $PQ$ . The OLS slope  $\hat{\beta}'_1$  is a weighted average of the original slope and  
523 the slope  $m_{PQ}$ ; choosing  $m$  sufficiently large forces  $\hat{\beta}'_1$  to approach  $m_{PQ}$ , thereby significantly altering of  
524 changing its sign. Thermodynamically, this corresponds to **sampling new, valid processes** from the  
525 continuum of accessible states within the hull. Conversely, an **exclusion attack** removes all points  
526 of  $\mathcal{D}$  except those near the vertices (or a vertex and a corresponding point on the base chord) that define  
527 a segment of opposing slope, as guaranteed by Lemma A.2 or Lemma A.3. Fitting the regression to this  
528 sparse subset yields a slope opposite to the original, which thermodynamically corresponds  
529 to **restricting the analysis to a coherent sub-ensemble** whose inherent trend opposes the original  
530 aggregation.

531 Consequently, the “attack vector” furnished by Lemma A.1 is not a mathematical artifact but  
532 a **consequence of the convex geometry inherent to thermodynamic state-space**.

533 In the context of the Multi-Level Model (Problem P1), the “points” are individual enthalpy-entropy pairs.  
534 Recalling the structured EEC dataset  $\Gamma = \{(\Delta_{\psi_i, n} S, \Delta_{\psi_i, n} H, T_{\psi_i, n})\}$ , each pair originates from a specific  
535 transformation process  $\psi_i$  at a specific temperature  $n$ . The vertices of the convex hull correspond  
536 to **individual measurements** that lie on the boundary of the pooled data region. **When each**  
537 **subpopulation is represented by a single measurement at a common temperature**—as in an  
538 isothermal subset  $\Omega(T^*)$  discussed in the case study—the removal or addition of a vertex **does** amount  
539 to excluding or including an entire thermodynamic subpopulation—as it is the case of bulk EEC reports  
540 in biological sciences.

541 This demonstrates that the apparent law-like correlation emerges **not from a hidden physical**  
542 **mechanism**, but from the geometric arrangement of thermodynamically distinct subpopulations in  
543 the  $(\Delta S, \Delta H)$  plane. The Exclusion-Inclusion Intervention (EII) exposes this dependency by showing the  
544 correlation hinges **entirely** on which subpopulations are included/excluded—the definitive hallmark of  
545 an aggregation artifact.

546 Applying this mathematical result to the enthalpy-entropy compensation problem yields the following  
547 conclusion:

#### 548 **Corollary 1 (Characterization of Commonly Reported EEC).**

549 *The commonly reported enthalpy-entropy compensation effect emerges from pooling distinct*  
550 *thermodynamic subpopulations in the Multi-Level Model (Problem P1). Consequently, the global*  
551 *correlations  $O_1$  and  $O_2$  inherit the **Geometric Vulnerability** of Theorem 1: they are unstable under*

552 *exclusion–inclusion interventions and must be recognized as **statistical phantoms**—artifacts of data*  
553 *aggregation, not universal physical relationships. This contextual intervention verdict identifies the EEC*  
554 *as a **composition-dependent artifact**, failing the most basic test of causal legitimacy.*

555 A final note pertains to the non-degeneracy condition in Theorem 1. This condition is satisfied for  
556 virtually all meaningful EEC datasets under an extremely mild thermodynamic assumption: that the  
557 considered transformations  $\psi_k$  induce strictly different heat capacity changes  $\Delta_{\psi_k} C_p$ . This follows  
558 directly from the relationship between arc lengths (or enthalpy ranges) and  $\Delta_{\psi_k} C_p$  established in  
559 Theorem 1.1 and Observation O2 of the PTF geometric analysis—a geometrical consequence  
560 of  $dh_k/ds_k = \tau^{16}$ . The only exception is the trivial case of a singleton dataset reported at a common  
561 temperature with identical Gibbs energy changes—a scenario already characterized as a direct  
562 manifestation of the Gibbs identity, not an independent global correlation<sup>16</sup>. **Therefore, geometric**  
563 **vulnerability is the generic condition for pooled enthalpy-entropy data.**

564 The abovementioned mathematical necessitates a careful **thermodynamic interpretation**, which is  
565 now examined next.

## 566 Thermodynamic Interpretation of the EII & Triage of Legitimate Physics

567 The mathematical truth established above points to a deep thermodynamic mis-specification inherent  
568 in the search for a universal EEC law—a mis-specification foreshadowed in the PTF analysis of discrete  
569 perturbations and the projection of the abstract hybrid thermodynamic manifold<sup>16</sup>. The conceptual error  
570 lies in **pooling data across distinct transformations** and projecting them onto the  $(\Delta S, \Delta H)$  plane. This  
571 operation discards the categorical structure of initial and final states of each specific transformation  
572 (e.g., molecular identity, phase composition) and the operational identity of the  
573 transformation  $\psi_k$  (represented by the difference operator  $\Delta_{\psi_k}$ ), treating points from different  
574 state-space trajectories as interchangeable. The representational reduction of thermodynamic  
575 hyperspace complexity to a two-dimensional scatter plot generates arrangements—specifically, the  
576 convex configurations identified in Theorem 1—that are **inherently prone to spuriously linear global**  
577 **patterns**. The resulting correlation is therefore a **statistical phantom**: an illusion of aggregation arising  
578 from the geometric vulnerability of pooled data, not an expression of a unified physical law.

579 This verdict on the global EEC **does not negate all notions of enthalpy-entropy linkage**. Instead,  
580 it **relocates legitimate behavior to its proper domain**—a distinction enabled by the integrated results  
581 of the PTF and the causal analysis above.

- 582 • **Within a single subpopulation  $S_k$** , the relationship is governed by the structural law  $dh_k/ds_k =$   
583  $\tau$ . The trajectory is generally nonlinear but, under the conditions specified in Observation O1 in  
584 the PTF<sup>16</sup>, can be validly approximated as **locally linear**. This is a legitimate, deterministic feature  
585 of the thermodynamic state-space.
- 586 • **Across subpopulations**, Corollary 1 identifies observed linearity as a **statistical phantom**—a  
587 spurious artifact of data aggregation with no causal basis.

588 In summary, the framework performs a necessary **thermodynamic-causal triage**. It identifies the  
589 reported global EEC as a **level-confounding aggregation artifact** while precisely defining the space for  
590 legitimate physics: the functional dependence of the structural law **within** a well-defined  
591 thermodynamic population. In this correct **intra-population** context, the relationship is a direct  
592 manifestation of the system's thermodynamics, categorically distinct from the phantom realm of  
593 aggregated correlations.

594 The verdict above, however, invites a final, more profound interrogation: **if a correlation escapes**  
595 **combinatorial fragility, does it thereby earn the status of a physical law autonomous from  $F$ ?**

596 To answer this, the second structural intervention of the causal verification framework—the **Global**  
597 **Confounder Intervention (GCI)** is now deployed.

## 598 Analytical Execution of GCI: Structural Non-Autonomy

599 The GCI's purpose is to quantify **structural independence**: it tests whether an observed pattern  
600 possesses the autonomous, law-like behavior of a genuine physical relationship, or is merely a passive,  
601 mathematical shadow of the underlying thermodynamics encoded in  $F$ . The decisive question is  
602 therefore: **Can a correlation that survives the EII demonstrate autonomy from  $F$ ?**

### 603 **Proposition 1 (Structural Non-Autonomy Under GCI).**

604 Let  $\Psi$  be a family of transformations and  $\Gamma$  be an associated EEC dataset as defined in Appendix B.  
605 Under the Global Confounder Intervention (GCI) that fixes the intensive variable  $\tau$  (an input to  $F$ ) to a  
606 constant value  $\tau'$ , any global correlation derived from the aggregated outputs  $O_1, O_2$  is structurally non-  
607 autonomous—and the correlation is wholly determined by the causal skeleton  $F$ .

#### 608 **Proof:**

609 By the thermodynamic identities for  $h_k(\tau)$  and  $s_k(\tau)$ ,  $(s_k, h_k)$  are outputs of  $F$  parameterized by  $\tau$ .  
610 Since  $F$  is a set of structural laws that are algebraic functions of  $\tau$ ,  $do(\tau = \tau')$  causes an algebraic  
611 collapse of the enthalpy and entropy data pairs into singletons for each  
612 subpopulation,  $(s_k(\tau'), h_k(\tau'))$ —a thermodynamic necessity. Consequently, the within-group  
613 covariance vanishes,  $\text{Cov}\left((s_k(\tau'), h_k(\tau')) \mid do(\tau = \tau')\right) = 0$ , and the parametrization of any global  
614 correlation derived from these outputs reduces to a standard Pearson correlation computed over points  
615 that are themselves direct outputs of  $F$ . Moreover, as a series of such interventions  $do(\tau = \tau'_1), do(\tau =$   
616  $\tau'_2), \dots, do(\tau = \tau'_n)$  is performed, the evolution of  $O(\tau)$  is fully locked to the propagation of  $\tau$  through  $F$ .  
617 The group-specific constants  $h_{k,0}, s_{k,0}$  are part of  $F$ ; they provide no independent source of variation.  
618 Therefore,  $O(\tau)$  tracks  $F$  deterministically across the entire temperature window—it is non-  
619 autonomous. ■

620 Thus, Proposition 1 ascertains  $O_j$  failure of the GCI's test for causal legitimacy. It is **not autonomous**,  
621 as its behavior is wholly dependent on the perturbation of  $\tau$ . It is **functionally dependent**, existing only  
622 for a specific instantiation of  $F$ 's inputs—the particular window of  $\tau$ , the set of functions  $\{c_k(\tau)\}$ , and the

parameters  $\{s_{k,0}, h_{k,0}\}$ . Consequently, it is an **epiphenomenon**: a side-effect of the specific configuration of the system ( $F$  conditioned by the exogenous variables), not an autonomous cause. Therefore,  $O_j$  is a **structurally redundant epiphenomenon**—a mathematical function of  $\tau$  that contains no additional information beyond what is described by  $F$ . The GCI delivers the final—perhaps the strongest—causal verdict: any apparent global correlation is a dependent epiphenomenon of  $F$ , lacking the modular autonomy that defines a physical law.

Hence, Proposition 1 complements Corollary 1 to form a **complete causal refutation**. The execution of (EII) proves aggregated correlations are fragile **context-dependent phantoms**, while Proposition 1 (GCI) proves that even a compositionally stable correlation is a **redundant mathematical shadow** of the known law  $F$ . Together, they demonstrate that the observed global EEC cannot occupy any causal middle ground: it is either **fragile** (an aggregation artifact) or **redundant** (not a novel law). In both cases, it fails the tests for genuine causal legitimacy. This final intervention closes the last conceptual loophole, confirming that the search for a universal enthalpy-entropy compensation "law" was chasing a **causal illusion**—an epiphenomenon generated by the geometry of thermodynamic state-space and mistaken for a foundational principle.

## Mechanisms of a Common Artifact

The conclusion that the EEC is a statistical phantom prompts a profound question: if such correlations are geometric artifacts, **why are they observed so frequently in practice?** The answer lies in two fundamental mechanisms that transform standard experimental and analytical practices into reliable generators of this specific illusion.

The first mechanism stems from the standard design of comparative studies, where the experimenter varies a single structural parameter—such as alkyl chain length or a substituent—while holding the core physicochemical transformation  $\psi_i$  constant. This constitutes a **tightly constrained perturbation**, modulating remote molecular degrees of freedom while leaving the essential **process coordinate** intact. When different transformations share a similar process coordinate, their thermodynamic response to such a discrete perturbation is **systematically correlated**: the relative change of  $\Delta H$  with respect to  $\Delta S$  is preserved. Consequently, the means  $(\overline{\Delta S_{\psi_k}}, \overline{\Delta H_{\psi_k}})$ —representative of horizontal and vertical anchoring—for each process do not scatter randomly but form a **sequence of steps that are co-linearly arranged**—manifesting as either a positive (EEC) or negative (AEEC) slope—in the  $(\Delta S, \Delta H)$  plane. This pre-existing **co-linear skeleton**, when superimposed with the small local curvature induced by temperature variation around each mean, produces an aggregated data cloud geometrically destined to yield a high- $R^2$  ordinary least-squares fit. The regression line does not discover a new law; it **rediscovers the alignment pre-engineered into the dataset** by the selection of transformations with a similar process coordinate and preserved respective changes of  $\Delta S$  and  $\Delta H$ —a condition frequently, though not exclusively, met by studying homologous molecular series.

A second, complementary mechanism operates even for non-homologous processes via a fundamental thermodynamic constraint. Theorem 4 of the PTF<sup>16</sup> establishes that all data measured at a common temperature  $T^*$  are **confined within a parallel band** of slope  $T^*$  in the  $(\Delta S, \Delta H)$  plane,

661 bounded by the minimum and maximum  $\Delta G$  values. Aggregating data from multiple temperatures—a  
 662 routine practice in meta-analyses—superimposes several such distinct bands. Under common  
 663 experimental conditions (e.g., a narrow biologically relevant temperature range, similar solvent  
 664 environments), these bands can overlap substantially, creating a **seemingly coherent, global linear**  
 665 **trend** from the union of thermodynamically independent sub-populations.

666 In practice, these mechanisms operate synergistically. A series of similar transformations is typically  
 667 investigated over a limited temperature interval. The resulting dataset is thus **doubly primed** for artifact  
 668 generation: its points are pre-aligned by mechanistic similarity *and* collectively bounded by  
 669 temperature-specific confinement bands. The researcher, adhering to sound comparative practice,  
 670 unintentionally constructs the ideal geometric conditions prescribed by Theorem 1. The robust,  
 671 reproducible linear correlation that emerges is consequently mistaken for a physical compensation law,  
 672 when it is in fact the **inevitable geometric outcome of projecting constrained thermodynamic data**  
 673 **onto a two-dimensional subspace**.

674 Finally, the **arbitrary realizability** established in the PTF<sup>16</sup> demonstrates that such correlations are not  
 675 merely common unintentional byproducts—they can be **intentionally constructed at will**. Any desired  
 676 linear trend (compensation, anti-compensation, or null) can be engineered by appropriate selection or  
 677 design of the transformation processes  $\psi_k$ . Consequently, the observation of a specific linear pattern  
 678 in any given dataset carries no necessary implication of a unique physical cause; it may equally reflect  
 679 intentional curation, unconscious selection bias, or the geometric inevitabilities already described. This  
 680 ultimate flexibility underscores that the EEC is fundamentally a **property of the representation**, not a  
 681 constraint imposed by physics.

682 **EEC Illustration: Protein Denaturation**

|                                                                            |                    |        | $\Delta\psi_i C_p$<br>$\left(\frac{J}{mol\ K}\right)$ | $\Delta\psi_i H$<br>$\left(\frac{kJ}{mol}\right)$ | $\Delta\psi_i S$<br>$\left(\frac{J}{mol\ K}\right)$ | $\Delta\psi_i H$<br>$\left(\frac{kJ}{mol}\right)$ | $\Delta\psi_i S$<br>$\left(\frac{J}{mol\ K}\right)$ |
|----------------------------------------------------------------------------|--------------------|--------|-------------------------------------------------------|---------------------------------------------------|-----------------------------------------------------|---------------------------------------------------|-----------------------------------------------------|
| <b>Ribonuclease A</b><br><b>(<i>RiA</i>)</b>                               | $\psi_{RiA}$       | 13,600 | 43.5                                                  | 2.37                                              | 6.70                                                | 6.06                                              | 17.8                                                |
| <b>Parvalbumin</b><br><b>(<i>Par</i>)</b>                                  | $\psi_{Par}$       | 11,500 | 46.0                                                  | 1.36                                              | 2.80                                                | 6.12                                              | 16.8                                                |
| <b>Egg-white</b><br><b>lysozyme (<i>Ewl</i>)</b>                           | $\psi_{Ewl}$       | 14,300 | 51.7                                                  | 2.02                                              | 5.52                                                | 6.24                                              | 17.6                                                |
| <b><math>\beta</math>-Trypsin (<math>\beta Tr</math>)</b>                  | $\psi_{\beta Tr}$  | 23,800 | 57.7                                                  | 1.33                                              | 3.43                                                | 6.23                                              | 17.9                                                |
| <b><math>\alpha</math>-Chymotrypsin</b><br><b>(<math>\alpha Ch</math>)</b> | $\psi_{\alpha Ch}$ | 25,200 | 57.7                                                  | 1.26                                              | 3.57                                                | 6.17                                              | 18.0                                                |
| <b>Papain (<i>Pap</i>)</b>                                                 | $\psi_{Pap}$       | 23,400 | 60.1                                                  | 0.93                                              | 1.60                                                | 6.22                                              | 17.0                                                |

683 Table 5 Selected Proteins Denaturation Process Dataset. Reported  $(\Delta_{\psi_k} S, \Delta_{\psi_k} H, T_k)$  and  $\Delta_{\psi_k} C_p$   
684 (Privalov & Gill<sup>21</sup>)

685 The purpose of this case study is to make tangible the analytical and ontological arguments made  
686 throughout the development of this work—specifically with respect to the EEC. The protein denaturation  
687 case study serves several objectives: (i) to facilitate visualizing the outcome of interventions; (ii) to  
688 demonstrate the strength of structural verdicts with a minimal dataset; (iii) to demonstrate the strength  
689 of structural verdicts with respect to a statistically robust correlation; and (iv) to provide contrast  
690 between statistical methods and the CVF.

## 691 Process Definition & Data Generation

692 A family of protein denaturation transformations are defined according to Appendix B:

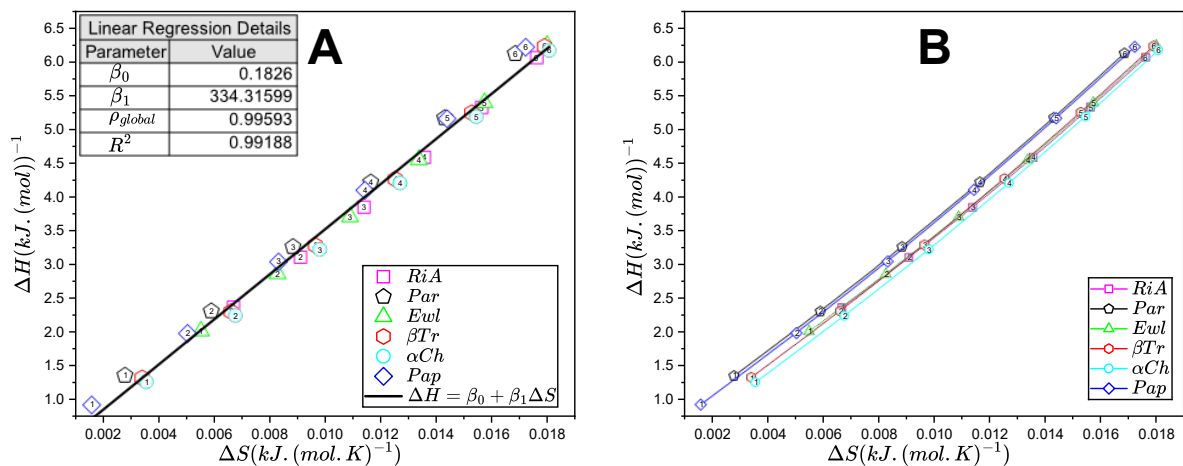
$$700 \quad \Psi = \left\{ \psi_{pr}: \text{Folded } pr \rightarrow \text{Unfolded } pr \mid pr \in \{\text{RiA, Par, Ewl, } \beta\text{Tr, } \alpha\text{Ch, Pap}\} \right\}$$

693 Original data for the selected proteins at  $T = \{298, 383\}K$  are given in Table 5, with  $\Delta_{\psi_k} C_p$  reported  
694 at 323K. The approximate relation,  $\overline{\Delta_{\psi_k} C_p} \approx \Delta H_{\text{range}} / T_{\text{range}} \approx \Delta_{\psi_k} C_{p,323}$  indicates a nearly  
695 constant  $\Delta_{\psi_k} C_p$  for each protein  $i$ . Therefore,  $\overline{\Delta_{\psi_k} C_p}$  was used to generate four additional equispaced  
696 data points within the temperature interval, yielding six temperatures per protein:  $T \in$   
697  $\{298, 315, 332, 349, 366, 383\}K$ . Verification of using  $\overline{\Delta_{\psi_k} C_p}$  to generate additional intermediate points is  
698 conducted by comparing the end points  $\Delta_{\psi_{k,383}} S$  with respect to the originally reported values by Privalov  
699 & Gill<sup>21</sup>. A maximum deviation of less than  $\sim 2\%$  was observed confirming the validity of the assumption.

701 The complete dataset  $\Gamma$  was then constructed according Appendix B. It comprises six proteins, each  
702 measured at the same six temperatures, with each subset  $\Gamma_i$  containing six triplets  $(\Delta_{\psi_{k,n}} S, \Delta_{\psi_{k,n}} H, T_n)$ ,  
703 resulting in a total of 36 data points. This structure directly instantiates the Multi-Level Model  
704 (Problem P1): each protein defines a distinct subpopulation  $S_k$ , temperature acts as the global  
705 continuous confounder  $\tau$ , and the measured variables are  $s_k(\tau) = \Delta_{\psi_k} S(T)$  and  $h_k(\tau) = \Delta_{\psi_k} H(T)$ .

## 706 The Illusion: A Compelling Global Correlation

707 Figure 2A presents the pooled dataset  $\Gamma$  with an ordinary least-squares (OLS) regression applied  
708 indiscriminately; the regression parameters and  $\rho_{\text{global}}$  are shown within the plot. The result is a  
709 seemingly robust linear law with high correlation. This high correlation coefficient represents the  
710 canonical evidence for enthalpy-entropy compensation in protein unfolding.



711

712 **Figure 2 (EEC Manifestation in the Protein Denaturation Dataset).** (A) Grand OLS regression of all  
 713 points in  $\Gamma$ , demonstrating a strong global correlation. (B) Real skeletal representations of the  
 714 individual subsets  $\Gamma_k$ , displaying the generic parallel-curve (spline-connected) behavior with  
 715 near-perfect local OLS fits (not shown). Data labels {1,2,3,4,5,6} correspond to temperatures  
 716 {298,315,332,349,366,383}K.

717 Figure 2B displays the same data but with spline-connected trajectories for each individual protein  $\Gamma_i$ .  
 718 This reveals the **local reality**: each protein obeys its own, distinct nonlinear trajectory governed by the  
 719 universal structural law  $F$  ( $d\Delta H/d\Delta S = T$ ). The near-perfect linear regression (not shown) for each  
 720 subset  $\Gamma_i$  produces a set of approximately parallel lines<sup>16</sup>. Consequently, the apparent global linearity in  
 721 Panel A is not a unifying physical feature but a **geometric composite** of these distinct, parallel, curved  
 722 state-space paths, whose differing arc lengths arise from their distinct  $\Delta\psi_k C_p$  values.

723 This visual separation provides a preliminary indication of failure under an **Exclusion-Inclusion**  
 724 **Intervention (EII)**—the purported “law” is not invariant to the level of aggregation. It also offers visual  
 725 insight into the **non-degeneracy** of the convex hull formed by the data at different common  
 726 temperatures.

## 727 Statistical Analysis of the Dataset: A Robust Global Correlation

728 To demonstrate the contrast between statistical diagnosis and causal verification, the protein  
 729 denaturation EEC dataset was subjected to a comprehensive set of standard statistical methods, all  
 730 operating at Pearl's Rung 1 (Association). The full details of these analyses are provided in  
 731 Supplementary Information Section S2; a summary is presented here.

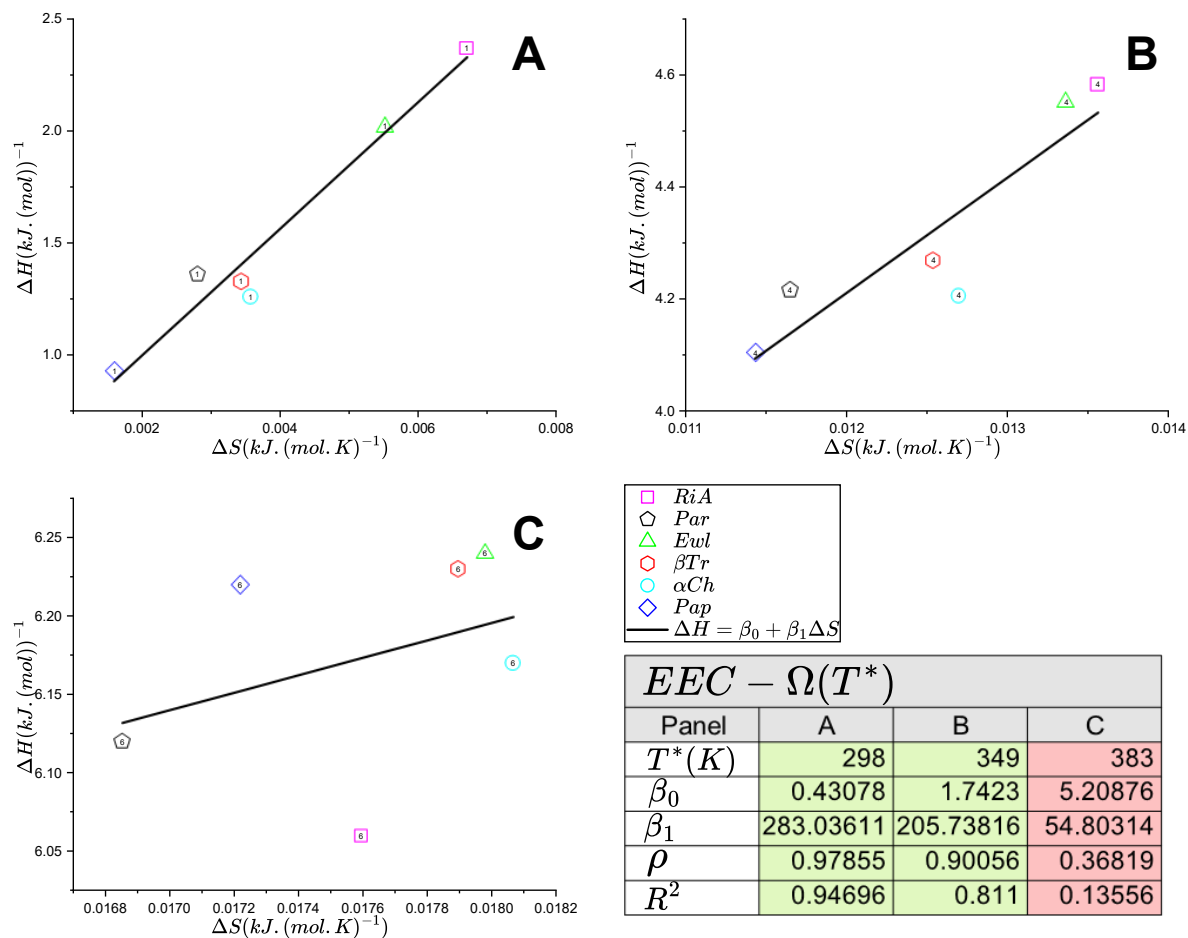
732 Five statistical models were constructed hierarchically. Model M1 is a simple OLS regression ( $\Delta H \sim \Delta S$ ),  
 733 establishing the baseline global correlation. Model M2 adds temperature as a covariate ( $\Delta H \sim \Delta S + T$ ).  
 734 Model M3 further adds protein identity as a fixed effect ( $\Delta H \sim \Delta S + T + \text{Protein}$ ). Mixed-effects models  
 735 M4 ( $\Delta H \sim \Delta S + (1 \mid \text{Protein})$ ) and M5 ( $\Delta H \sim \Delta S + T + (1 \mid \text{Protein})$ ) incorporate random intercepts by  
 736 protein. In all models, the  $\Delta S$  coefficient remained highly significant ( $p < 0.001$ ). Notably, adding

temperature to the mixed model (M5) worsened model fit (AIC increased from -33.49 to -31.88), and temperature itself was not a significant predictor ( $p = 0.532$ ).

Leave-one-protein-out analysis confirmed that the global correlation is stable under sampling variation: the OLS slope remained within  $338.8 \pm 3.4$  and  $R^2$  remained consistently high ( $> 0.98$ ) regardless of which protein was removed. Partial correlation analysis controlling for temperature yielded  $\rho = 0.9928$ , virtually identical to the raw Pearson correlation ( $r = 0.9926$ ), confirming that temperature does not explain the global correlation.

**All statistical analyses confirm that the EEC correlation is statistically robust.** Adding more parameters, controlling for covariates, removing individual proteins, or accounting for hierarchical structure does not eliminate the correlation. Yet, as analytically demonstrated previously, the CVF's Rung 2 interventions—EII and GCI—reveal that the correlation is structurally fragile: it is context-dependent (EII) and non-autonomous (GCI). The statistical methods detect robustness; the CVF detects the phantom. This contrast is the central epistemological distinction of this work. Next visualization of the GCI and composite GCI/EII are conducted demonstrating the true nature of the correlation.

## 752 The Global Confounder Intervention: Dissolving the Phantom



753

754 *Figure 3 (Investigating EEC Robustness at Different Common Temperatures). (A) Isothermal*  
755 *correlation at  $T = 298$  K, showing a strong linear fit. (B) Correlation at  $T = 349$  K, exhibiting significant*  
756 *weakening. (C) Correlation at  $T = 383$  K, where the linear pattern fails.*

757 A controlled Global Confounder Intervention (GCI) is performed by fixing the confounder  $\tau$  (temperature)  
758 and examining the correlations within each isothermal subset  $\Omega(T^*)$ . This procedure directly mirrors the  
759 most common experimental scenario for reporting the EEC, where data are collected at a single  
760 temperature and aggregated across processes. Figure 3 presents the results for three representative  
761 temperatures.

762 At  $T = 298$  K (Panel A), the correlation remains strong ( $R^2 = 0.958$ ), closely mimicking the slope and  
763 visual appearance of the grand illusion from the pooled OLS fit of  $\Gamma$ . This reflects the specific, incidental  
764 arrangement of each process path's projection at this specific temperature.

765 At  $T = 349$  K (Panel B), the correlation weakens significantly ( $R^2 = 0.811$ ), demonstrating the  
766 non-coincidental evolution of the different transformation trajectories due to their  
767 distinct  $\Delta_{\psi_i} C_p$  values—or, equivalently, that the linear pattern is not temperature-invariant.

768 Finally, at  $T = 383$  K (Panel C), the correlation catastrophically fails ( $R^2 = 0.136$ ). The purported  
769 “compensation law” effectively vanishes when examined across a wider temperature range, rendering  
770 the physical information assumed to be encoded in the slope  $\beta_1$  vacuous.

771 This progression is an experimental illustration of the EEC being an **epiphenomenon**: a side-effect of  
772 the specific configuration of the system. A genuine causal law would exhibit modular autonomy—its  
773 core relationship would persist under perturbation of  $\tau$ . Here, the correlation is wholly dependent on a  
774 specific window of the system's inputs; change  $T$ , and the pattern disintegrates.

## 775 Geometric Vulnerability, Arbitrary Realizability, and Composite Interventions

776 Figure 4 elucidates the mathematical mechanism underlying the GCI results by displaying the convex  
777 hulls for the isothermal subsets at **298 K**, **349 K**, and **383 K**.

778 At **T = 298 K** (Panel A), the hull comprises two increasing monotone chains, which facilitates the  
779 high- $R^2$  linear fit shown in Figure 2A for the full set  $\Omega(298)$ . A closer inspection reveals that **both a**  
780 **strong EEC** ( $\rho \approx 1$ ) and a **moderate AEEC** ( $\rho \approx -0.84$ ) coexist within  $\Omega(298)$ , corresponding  
781 respectively to the subsets  $\{\text{Pap}, \text{Ewl}, \text{RiA}\}$  and  $\{\text{Par}, \beta\text{Tr}, \alpha\text{Ch}\}$ . The AEEC arises from consistent  
782 negative correlations between the protein pairs  $\text{Par}-\beta\text{Tr}$ ,  $\text{Par}-\alpha\text{Ch}$ , and  $\beta\text{Tr}-\alpha\text{Ch}$ . **This selective**  
783 **manifestation of opposing correlations is a direct geometric consequence of an Exclusion–**  
784 **Inclusion Intervention (EII)**: by including or excluding specific protein subpopulations, the sign of the  
785 global slope can be arbitrarily chosen.

786 Crucially, both the observed EEC and AEEC **weaken at T = 349 K** (Panel B) and **completely disintegrate**  
787 **at T = 383 K** (Panel C). At this higher temperature, two of the previously negative pairwise correlations  
788 ( $\text{Par}-\beta\text{Tr}$ ,  $\text{Par}-\text{Ch}$ ) reverse sign, and other pairwise correlations—both positive and negative—likewise  
789 invert their 298 K behavior. The convex hull at 383 K evolves into **four distinct monotone chains** (two

790 increasing, two decreasing), creating the ideal geometric condition for vertex-based slope manipulation  
791 described in **Lemma A.2**. This geometric reorganization—a direct consequence of the **Global**  
792 **Confounder Intervention (GCI)**—explains the **catastrophic failure of the EEC** that was constructible  
793 from the full set  $\Omega(298)$ , as shown in Figure 2C.

794 Together, these observations illustrate a **composite intervention**: the EII reveals the correlation’s  
795 dependence on compositional selection, while the GCI exposes its fragility under parametric  
796 perturbation. This dual vulnerability visually confirms **Theorem 1 (Geometric Vulnerability)** and  
797 **Proposition 1 (Structural Non-Autonomy Under GCI)**. In every case, the OLS slope is **not a stable**  
798 **property of the underlying thermodynamics** but a fragile function of which data points constitute the  
799 convex hull. It also confirms the principle of **arbitrary realizability**: both compensation and  
800 anti-compensation are geometrically constructible from the same thermodynamic substrate,  
801 depending solely on selective inclusion or exclusion of processes.

802 The protein denaturation case study has achieved its objectives. It visualized the outcome of EII and GCI  
803 interventions on a real thermodynamic dataset: the global EEC correlation is context-dependent (EII)  
804 and non-autonomous (GCI). It demonstrated the strength of structural verdicts with a minimal  
805 dataset—only six proteins and six temperatures per protein were sufficient to reveal the phantom,  
806 because the verdict follows from the analytical structure of  $F$ , not from sample size. It demonstrated  
807 that a statistically robust correlation—one that survives hierarchical regression, mixed-effects models,  
808 leave-one-out analysis, and partial correlation (SI Section S2)—can still be structurally fragile. Finally, it  
809 provided a clear contrast: Rung 1 statistical methods confirm robustness; Rung 2 interventions reveal  
810 the phantom. The case study thus makes tangible the analytical and ontological arguments of this work,  
811 rendering the abstract CVF logic visible and concrete.

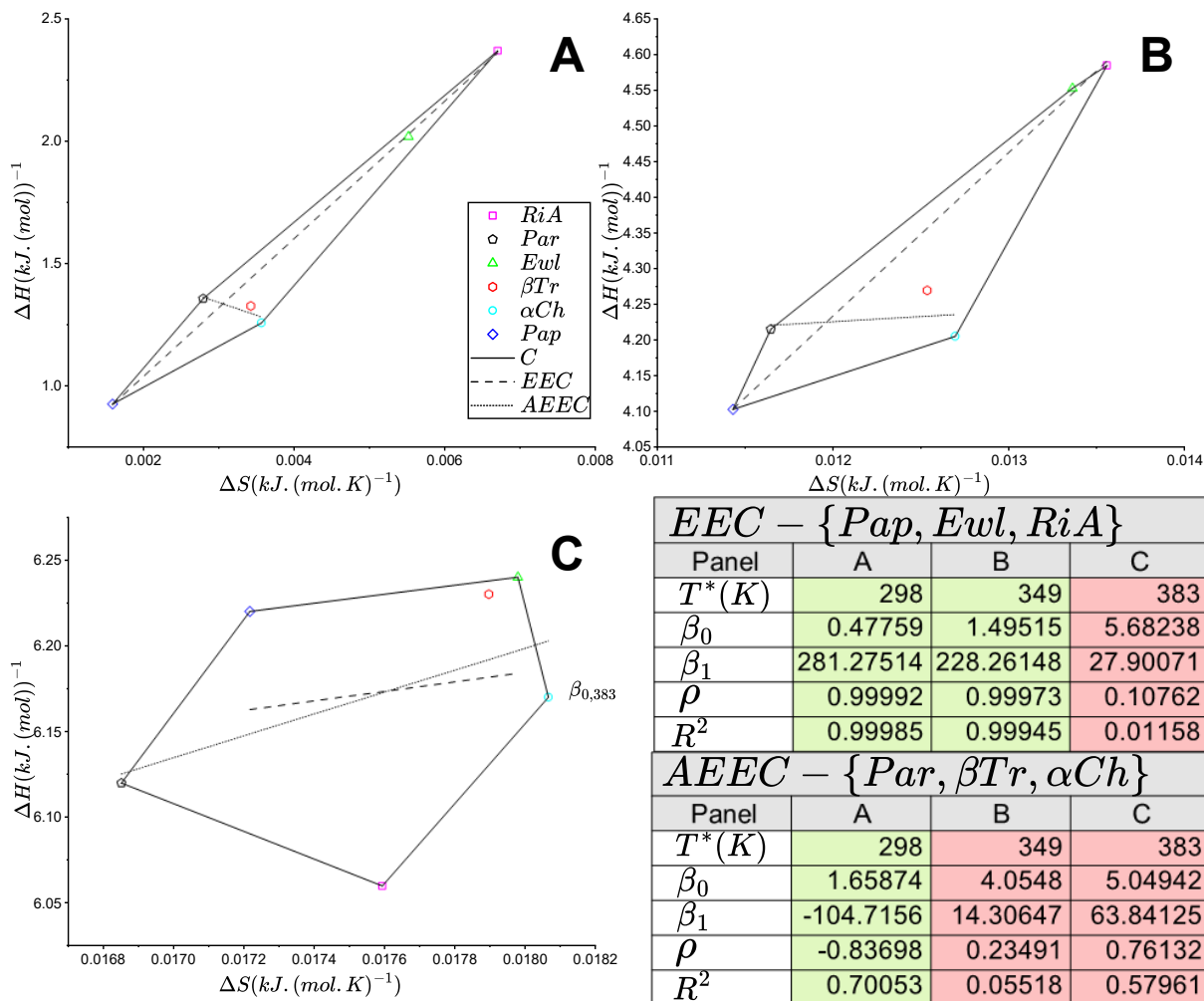


**Figure 4. (Investigating EEC under Selective EII & Demonstrating the OLS Vulnerability Theorem 1).** Convex Hull  $C$  shown for all panels demonstrating existence of different correlations between different proteins and their evolution with temperature. The evolution of select observed EEC and AEEC under EII: (A) Demonstration of both a strong EEC and a moderate AEEC among different proteins at  $T = 298$  K. (B) At  $T = 349$  K The selected EEC holds while the AEEC disintegrates. (C) At  $T = 383$  K both EEC and AEEC correlations disintegrate.

## Concluding Remarks

This work introduced the causal verification paradigm and the associated Causal Verification Framework (CVF)—a natural and complementary expansion of the causal inference landscape. The CVF addresses a foundational problem: given a known causal skeleton  $F$ , can we determine whether an observed statistical summary  $O$  is a genuine autonomous law or merely a derivative phantom?

The philosophy is straightforward: leverage established knowledge—the causal skeleton  $F$ —to adjudicate the status of a phenomenological observation. Physical laws supersede observations; they are the fixed reference against which all claims of causal legitimacy are tested. The CVF operationalizes this philosophy through the HFCM architecture, which separates known structural laws from observed patterns, and through the Structural Compliance Principle (which qualifies the generic principle of autonomy within the verification paradigm) and the Stability (Invariance) Principle (which defines the

operational interventions to test compliance). Together, these principles define interventions that test whether  $O$  is modular relative to  $F$  at two levels: contextual and structural.

The CVF is not a competitor to existing causal identification methods. It is their complement. Identification methods ask: "What causes  $Y$ ?"—a problem of discovery. The CVF asks a different, foundational question: "Is this observed correlation a phantom or a law?"—a problem of verification where the causal skeleton  $F$  is known. Together, identification and verification complete the causal inference landscape. One discovers the unknown; the other verifies the observed against existing knowledge.

Technically, the CVF operates at Pearl's Rung 2 (Intervention). It does not ask whether a statistical summary  $O_j$  is associated with  $V$  (Rung 1), nor does it merely adjust for confounding. It asks whether  $O_j$  survives interventions. This is the strongest form of causal proof: if  $O_j$  fails to remain invariant under structural or contextual interventions, it is not autonomous from  $F$ —it is a statistical phantom. If it survives both, it is a stable genuine law that captures a causal mechanism beyond what is embodied within  $F$ .

The EEC investigation serves as a warning to trusting high- $R^2$  observational correlations. The CVF's dual interventions—EII and GCI—exposed the global correlation as context-dependent (Corollary 1) and non-autonomous (Proposition 1), delivering a structural verdict of "statistical phantom." The verdict follows from the structure of  $F$ , not from empirical coverage. Because of analytical execution, the verdict is universal. Even a minimal dataset was sufficient to observe the phantom, as illustrated throughout the protein denaturation case study (Figures 3 and 4).

By ignoring thermodynamic necessities, EEC reports turn a blind eye to existing knowledge. The structural law  $dH_k/dS_k = \tau$  is always present and universal<sup>16</sup>, whether explicitly considered or not. Temperature is often either not explicitly considered or buried in the  $T\Delta S$  term. The CVF, specifically through the GCI, reveals its effect on the stability of the EEC correlation. Thus, the EEC investigation serves as proof of concept demonstrating both the need and effectiveness of the CVF.

## Data Availability

Data used were extracted from cited literature and will be made available on request.

## Supplementary Material

Supplementary Information includes historical perspectives on the EEC controversy (Table S1, Section S1) and statistical analysis details (Section S2).

## Declaration of Competing Interest

The author declares that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

862

863

# Declaration of Generative Ai and Ai-Assisted Technologies in the Writing Process

864

865

866

867

During the preparation of this work, the author used DeepSeek-V4 to improve the language and clarity of the manuscript and format equations into LaTeX and Microsoft Word equation structures. After using this tool/service, the author reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

868

## Funding

869

870

The author received no specific funding for this work. The Article Processing Charge (APC) could be covered upon acceptance and will be acknowledged accordingly.

871

## Acknowledgement

872

TBA

873

## Nomenclature

874

### General

| $K$                    | Number of distinct thermodynamic subpopulations (process families)                                       |
|------------------------|----------------------------------------------------------------------------------------------------------|
| $S_k$                  | The $k$ -th subpopulation, corresponding to a well-defined transformation process $\psi_k$               |
| $s_{k,n}, h_{k,n}$     | Observed entropy change and enthalpy change for the $n$ -th measurement in subpopulation $S_k$           |
| $\tau, \tau_i$         | Global continuous confounder (temperature); $\tau_i$ is a specific measurement value                     |
| $c_k(\tau)$            | Subpopulation-specific structural function, analogous to $\Delta_{\psi_k} C_p(\tau)$                     |
| $\mathcal{D}$          | Pooled global dataset: $\mathcal{D} = \bigcup_{k=1}^K \{(s_{k,i}, h_{k,i}, \tau_i)\}$                    |
| $\bar{s}, \bar{h}$     | Grand means of $s$ and $h$ over $\mathcal{D}$                                                            |
| $\rho_{\text{global}}$ | Global Pearson correlation coefficient calculated from $\mathcal{D}$                                     |
| $\beta_0, \beta_1$     | Intercept and slope of the global Linear Regression Model (LRM): $\Delta H = \beta_0 + \beta_1 \Delta S$ |

875

### Causal Verification Framework (HFCM)

| <b>HFCM</b>   | Hierarchical Functional Causal Model: $\langle U, V, F, O \rangle$                                                     |
|---------------|------------------------------------------------------------------------------------------------------------------------|
| $U$           | Set of <b>exogenous (background) variables</b> determined outside the model                                            |
| $V$           | Set of <b>endogenous variables</b> determined within the model: $V = \{V_1, \dots, V_n\}$                              |
| $F$           | Set of <b>structural laws</b> : $F = \{f_1, \dots, f_n\}$ , where $f_i: \text{PA}_i \cup U_i \rightarrow V_i$          |
| $O$           | Set of <b>observed correlations</b> : $O = \{O_1, \dots, O_m\}$ , each a statistical summary of realized values in $V$ |
| $\text{PA}_i$ | Parents of variable $V_i$ within $V$                                                                                   |
| $O_1$         | Global linear correlation $\rho_{\text{global}}$ (or the fitted LRM parameters $\beta_0, \beta_1$ )                    |

|       |                                                                                   |
|-------|-----------------------------------------------------------------------------------|
| $O_2$ | Ecological correlation: $\text{Corr}(\bar{s}_k, \bar{h}_k)$ across subpopulations |
|-------|-----------------------------------------------------------------------------------|

## Interventions & Verification Tests

|                 |                                                                                        |
|-----------------|----------------------------------------------------------------------------------------|
| $\hat{\beta}_1$ | Ordinary least squares (OLS) slope estimate from a dataset                             |
| $\mathcal{C}$   | Convex hull of a planar dataset $\mathcal{D}$ : $\mathcal{C} = \text{Ch}(\mathcal{D})$ |
| $V_L, V_R$      | Leftmost and rightmost vertices of $\mathcal{C}$ (by $x$ -coordinate)                  |
| $\mathcal{B}$   | Base chord of $\mathcal{C}$ , the segment $V_L V_R$                                    |
| $m_{ab}$        | Slope of the segment between points $a$ and $b$ : $m_{ab} = (y_b - y_a)/(x_b - x_a)$   |

## Acronyms & Categories

|             |                                                                     |
|-------------|---------------------------------------------------------------------|
| <b>EEC</b>  | Enthalpy–Entropy Compensation                                       |
| <b>AEEC</b> | Anti-Enthalpy–Entropy Compensation (negative slope correlation)     |
| <b>LRM</b>  | Linear Regression Model ( $\Delta H = \beta_0 + \beta_1 \Delta S$ ) |
| <b>MLM</b>  | Multi-Level Model (formalization of Problem P1)                     |
| <b>HFCM</b> | Hierarchical Functional Causal Model                                |
| <b>EII</b>  | Exclusion–Inclusion Intervention                                    |
| <b>GCI</b>  | Global Confounder Intervention                                      |
| <b>OLS</b>  | Ordinary Least Squares                                              |
| <b>SP</b>   | Simpson’s Paradox                                                   |

## References

- (1) Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the computational design of solid catalysts. *Nature Chemistry* **2009**, 1 (1), 37-46. DOI: <https://doi.org/10.1038/nchem.121>.
- (2) Ceder, G. Opportunities and challenges for first-principles materials design and applications to Li battery materials. *MRS Bulletin* **2010**, 35 (9), 693-701. DOI: <https://doi.org/10.1557/mrs2010.681>.
- (3) Krug, R. R.; Hunter, W. G.; Grieger, R. A. Enthalpy-entropy compensation. 2. Separation of the chemical from the statistical effect. *The Journal of Physical Chemistry* **1976**, 80 (21), 2341-2351. DOI: <https://doi.org/10.1021/j100562a007>.
- (4) Krug, R. R.; Hunter, W. G.; Grieger, R. A. Statistical interpretation of enthalpy–entropy compensation. *Nature* **1976**, 261 (5561), 566-567. DOI: [10.1038/261566a0](https://doi.org/10.1038/261566a0).
- (5) Exner, O. On the enthalpy-entropy-relationship. *Collect. Czech. Chem. Commun.* **1964**, 29, 1094-1113. DOI: <https://doi.org/10.1135/cccc19641094>.
- (6) Agrawal, R. K. On the compensation effect. *Journal of thermal analysis* **1986**, 31 (1), 73-86. DOI: [10.1007/BF01913888](https://doi.org/10.1007/BF01913888).
- (7) Barrie, P. J. The mathematical origins of the kinetic compensation effect: 1. the effect of random experimental errors. *Physical Chemistry Chemical Physics* **2012**, 14 (1), 318-326, [10.1039/C1CP22666E](https://doi.org/10.1039/C1CP22666E). DOI: [10.1039/C1CP22666E](https://doi.org/10.1039/C1CP22666E).
- (8) Barrie, P. J. The mathematical origins of the kinetic compensation effect: 2. the effect of systematic errors. *Physical Chemistry Chemical Physics* **2012**, 14 (1), 327-336, [10.1039/C1CP22667C](https://doi.org/10.1039/C1CP22667C). DOI: [10.1039/C1CP22667C](https://doi.org/10.1039/C1CP22667C).
- (9) Mianowski, A.; Radko, T.; Siudyga, T. Kinetic compensation effect of isoconversional methods. *Reaction Kinetics, Mechanisms and Catalysis* **2021**, 132 (1), 37-58. DOI: [10.1007/s11144-020-01898-2](https://doi.org/10.1007/s11144-020-01898-2).

(10) Cooper, A.; Johnson, C. M.; Lakey, J. H.; Nöllmann, M. Heat does not come in different colours: entropy–enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions. *Biophysical Chemistry* **2001**, 93 (2), 215-230. DOI: [https://doi.org/10.1016/S0301-4622\(01\)00222-8](https://doi.org/10.1016/S0301-4622(01)00222-8).

(11) Chodera, J. D.; Mobley, D. L. Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design. *Annu Rev Biophys* **2013**, 42, 121-142. DOI: 10.1146/annurev-biophys-083012-130318 From NLM.

(12) Wells, P. R. Linear Free Energy Relationships. *Chemical Reviews* **1963**, 63 (2), 171-219. DOI: 10.1021/cr60222a005.

(13) Himanen, L.; Geurts, A.; Foster, A. S.; Rinke, P. Data-Driven Materials Science: Status, Challenges, and Perspectives. *Advanced Science* **2019**, 6 (21), 1900808. DOI: <https://doi.org/10.1002/advs.201900808>.

(14) Efstathiou, G. H0 revisited. *Monthly Notices of the Royal Astronomical Society* **2014**, 440 (2), 1138-1152. DOI: 10.1093/mnras/stu278 (accessed 1/7/2026).

(15) Camps-Valls, G.; Gerhardus, A.; Ninad, U.; Varando, G.; Martius, G.; Balaguer-Ballester, E.; Vinuesa, R.; Diaz, E.; Zanna, L.; Runge, J. Discovering causal relations and equations from data. *Physics Reports* **2023**, 1044, 1-68. DOI: 10.1016/j.physrep.2023.10.005 (accessed 2026/07/27).

(16) Justanieah, A. M. Predictive Thermodynamic Framework: Skeletal Constructs and Geometric Constraints for Data Validation with Application to Enthalpy–Entropy Correlations. *Results in Engineering* **2026**, 112598. DOI: <https://doi.org/10.1016/j.rineng.2026.112598>.

(17) PEARL, J. *Causality*; Cambridge University Press, 2009. DOI: <https://doi.org/10.1017/CBO9780511803161>.

(18) Bareinboim, E.; Correa, J. D.; Ibeling, D.; Icard, T. On Pearl’s Hierarchy and the Foundations of Causal Inference. In *Probabilistic and Causal Inference: The Works of Judea Pearl*, Vol. 36; Association for Computing Machinery, 2022; pp 507–556.

(19) Peters, J.; Bühlmann, P.; Meinshausen, N. Causal Inference by using Invariant Prediction: Identification and Confidence Intervals. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **2016**, 78 (5), 947-1012. DOI: <https://doi.org/10.1111/rssb.12167> (accessed 7/7/2026).

(20) Martin Arjovsky, L. B., Ishaan Gulrajani, David Lopez-Paz. Invariant Risk Minimization. *Preprint at* **2020**, *arXiv:1907.02893v3*. DOI: <https://doi.org/10.48550/arXiv.1907.02893>.

(21) Privalov, P. L.; Gill, S. J. Stability of Protein Structure and Hydrophobic Interaction. In *Advances in Protein Chemistry*, Anfinsen, C. B., Edsall, J. T., Richards, F. M., Eisenberg, D. S. Eds.; Vol. 39; Academic Press, 1988; pp 191-234.

## 934 Appendix A: Geometric Conditions for OLS Slope Reversal

935 Let  $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^n \subset \mathbb{R}^2$  with  $n \geq 3$ , and let  $\hat{\beta}_1 \neq 0$  be the ordinary least-squares (OLS) slope from  
 936 regressing  $y$  on  $x$  over  $\mathcal{D}$ . Denote the convex hull of the dataset by  $C = \text{conv}(\mathcal{D})$ ;  
 937 assume  $C$  is **non-degenerate** (i.e., not a line segment).

938 Let  $V_L$  and  $V_R$  be the vertices of  $C$  with the minimal and maximal  $x$ -coordinates, respectively. The  
 939 segment  $\mathcal{B} = V_L V_R$  is the **base chord** of  $C$ .

940 A **monotone chain** is a sequence of consecutive vertices along the boundary  $\partial C$  for which  $x$  increases  
 941 and  $y$  is monotonic (either non-decreasing or non-increasing).

942 **Lemma A.1 (Interior-segment existence)**

943 There exist points  $P, Q \in C$  such that the slope

946 
$$m_{PQ} = \frac{y_Q - y_P}{x_Q - x_P}$$

944 has sign opposite to  $\hat{\beta}_1$ . Moreover,  $P$  and  $Q$  can be chosen either in the interior of  $C$ , on the boundary  
945 of  $C$ , or one of each.

947 **Proof:**

948 Pick any interior point  $O \in \text{int}(C)$ . Since  $\text{int}(C)$  is non-empty, there exists  $\epsilon > 0$  such that a disk of  
949 radius  $\epsilon$  centered at  $O$  lies entirely within  $\text{int}(C)$ . Within this disk, choose a diameter with slope  $m =$   
950  $-\text{sgn}(\hat{\beta}_1)$ . Because  $C$  is convex and the disk is contained in  $\text{int}(C)$ , the line through this diameter  
951 intersects  $C$  in a closed segment  $[P, Q]$  with  $P \neq Q$ . The slope of  $PQ$  equals  $m$ , which has sign opposite  
952 to  $\hat{\beta}_1$ .

953 Quantitatively, the OLS slope of the augmented dataset is a weighted average of the original slope  $\hat{\beta}_1$  and  
954 the segment slope  $m_{PQ}$ :

956 
$$\hat{\beta}'_1 = \frac{n \cdot \hat{\beta}_1 + m \cdot m_{PQ}}{n + m}$$

955

957 where  $n$  is the number of original points and  $m$  is the number of new points along segment  $PQ$ . Since  
958 the denominator is positive, the sign of  $\hat{\beta}'_1$  is determined by the numerator. As  $m \rightarrow \infty$ :

960 
$$\text{sgn}(\hat{\beta}'_1) \rightarrow \text{sgn}(m_{PQ})$$

959

961 which has sign opposite to  $\hat{\beta}_1$ . Thus, sufficient new points along  $PQ$  will force the global OLS slope  
962 toward the opposite sign. ■

963 **Note.** Lemma A.1 formalizes the intuitive fact that inside any non-degenerate convex hull one can  
964 always find a segment whose slope opposes the global OLS trend. This guarantees the existence of an  
965 “attack vector” for manipulating the regression slope.

966 **Lemma A.2 (Vertex-vertex reversal)**

967 If  $C$  contains vertices from both an **increasing** monotone chain and a **decreasing** monotone chain,  
968 then:

- 969 • If  $\hat{\beta}_1 > 0$ , any two vertices taken from a decreasing chain yield a slope  $< 0$ .
- 970 • If  $\hat{\beta}_1 < 0$ , any two vertices taken from an increasing chain yield a slope  $> 0$ .

971 In either case the resulting slope has sign opposite to  $\hat{\beta}_1$ .

972 **Proof:**  
 973 On a decreasing chain,  $y$  decreases as  $x$  increases; therefore the slope between any two vertices on  
 974 that chain is negative. On an increasing chain,  $y$  increases with  $x$ , so the slope between any two vertices  
 975 is positive. Selecting vertices from the chain whose slope sign opposes  $\hat{\beta}_1$  gives the required opposite  
 976 slope. ■

### 977 **Lemma A.3 (Chord-vertex reversal)**

978 If every vertex of  $C$  lies on monotone chains of the **same** direction (all increasing or all decreasing), then  
 979 no pair of vertices yields a slope opposite to  $\hat{\beta}_1$ . In this situation, for any vertex  $V_k$  there exists a  
 980 point  $Q$  on the **base chord**  $\mathcal{B} = V_L V_R$  such that the slope

$$982 \quad m_{V_k Q} = \frac{y_Q - y_k}{x_Q - x_k}$$

981 has sign opposite to  $\hat{\beta}_1$ .

### 983 **Proof:**

984 Because all vertices belong to monotone chains of the same direction, the slope between any two  
 985 vertices shares the sign of  $\hat{\beta}_1$ . Let  $m_c$  be the slope of the base chord  $\mathcal{B}$ . Choose  $Q \in \mathcal{B}$  as follows:

- 986 • If  $\hat{\beta}_1 > 0$ , pick  $Q$  with  $x_Q > x_k$  and with  $y_Q$  below the line through  $V_k$  having slope  $m_c$ ;  
 987 then  $m_{V_k Q} < 0$ .
- 988 • If  $\hat{\beta}_1 < 0$ , pick  $Q$  with  $x_Q < x_k$  and with  $y_Q$  above that line; then  $m_{V_k Q} > 0$ .

989 In both cases  $m_{V_k Q}$  has sign opposite to  $\hat{\beta}_1$ . The convexity of  $C$  guarantees that such a point  $Q$  exists on  
 990 the base chord. ■

### 991 **Connection to the Main Text**

992 These lemmas provide the geometric underpinning for **Theorem 1 (Geometric Vulnerability of OLS)** in  
 993 the main text. They guarantee that, for any non-degenerate planar dataset, one can always construct a  
 994 segment—either between two vertices (Lemma A.2) or between a vertex and a point on a chord  
 995 (Lemma A.3)—whose slope opposes the original OLS trend. Adding data points along such a segment  
 996 systematically shifts the global OLS slope, enabling the deliberate construction (or destruction) of a  
 997 spurious linear correlation—the essence of the **Exclusion-Inclusion Intervention (EII)**.

## 998 **Appendix B: Physicochemical Transformations and EEC Dataset**

### 999 **B.1 Family of Transformations**

1000 Let there exist a family of  $m$  physicochemical transformations:

$$1001 \quad \Psi = \{\psi_1, \psi_2, \dots, \psi_m\}$$

where each  $\psi_k$  is a well-defined transformation between fixed thermodynamic states with stoichiometries, conditions, and composition fully specified. Each  $\psi_k$  belongs to a broader class (e.g., a homologous series of adsorbates, ligands, or surface reactions). The index  $k$  denotes the identity of the subpopulation or process.

## B.2 Structured Dataset for EEC

Each transformation  $\psi_k$  is associated with a dataset  $\Gamma_k$  containing  $n_k$  measurement triplets:

$$\Gamma_k = \{(\Delta_{\psi_{k,n}} S, \Delta_{\psi_{k,n}} H, T_{\psi_{k,n}})\}_{n=1}^{n_k}$$

The union of all process-level data is:

$$\Gamma = \bigcup_{k=1}^m \Gamma_k$$

Under  $\psi_k$  transformation, the associated standard-state (fixed composition by definition) change in enthalpy and entropy are determined by:

$$\Delta_{\psi_k} S(T) = \Delta_{\psi_{k,0}} S + \int_{T_{\psi_{k,0}}}^{T'} \frac{\Delta_{\psi_k} C_p(T)}{T} dT,$$

$$\Delta_{\psi_k} H(T) = \Delta_{\psi_{k,0}} H + \int_{T_{\psi_{k,0}}}^{T'} \Delta_{\psi_k} C_p(T) dT,$$

$$\frac{d\Delta_{\psi_k} H(T)}{\Delta_{\psi_k} S(T)} = T$$

where  $\Delta_{\psi_{k,0}} S$  and  $\Delta_{\psi_{k,0}} H$  are the entropy and enthalpy changes computed at the reference temperature  $T_{\psi_{k,0}}$ ; and  $T' \in \{T_{\psi_{k,n}}\}_{n=1}^{n_k}$ .

Thus, these equations define the underlying causal skeleton  $F$  (equivalently, the thermodynamic skeleton) that governs the behavior of all EEC datasets  $\Gamma_k$  individually:

$$(\Delta_{\psi_k} S, \Delta_{\psi_k} H) = F(\psi_k, T) \forall k$$

## B.3 Data Subsets and Subpopulations

A subpopulation  $\mathcal{K}$  is defined as a subset of the family  $\Psi$ :

$$\mathcal{K} \subseteq \Psi$$

For a fixed intensive variable value  $T^*$ , define the subset of data:

$$\Omega(T^*) = \{(\Delta_{\psi_{k,n}} S, \Delta_{\psi_{k,n}} H, T_{\psi_{k,n}}) \in \Gamma \mid T_{\psi_{k,n}} = T^*\}$$

1028 i.e., a subset that contains all data from all subpopulations measured at the same temperature.

1029 **Connection to the Main Text**

1030 Appendix B summarizes the definition of transformations, the data structuring, and thermodynamic

1031 identities associated with an EEC dataset mentioned throughout the main text of the article.

1032 Comprehensive treatment is provided in the PTF formalism<sup>16</sup>.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [CVFSIScRev4.2.docx](#)