



From transient shocks to unexpected outcomes: Disruptive drivers in scenario pathways

Andrew G. Ross 

Forschungszentrum Jülich GmbH, Institute of Climate and Energy Systems, Jülich Systems Analysis, Wilhelm-Johnen-Strasse, 52425 Jülich, Germany

ARTICLE INFO

Dataset link: <https://doi.org/10.5281/zenodo.19364824>

JEL classification:

Q40
C63
D81
Q55

Keywords:

Disequilibrium
Scenario extremes
Unexpected outcomes
Resilience
Energy transition
Complexity

ABSTRACT

Scenario pathways (e.g. for the energy transition) often use a single trajectory or a band. That is not sufficient when one needs to understand not just that outcomes differ but what drives those differences, whether a transient shock, a different storyline, or uncertainty about how factors interact. Doing so requires tracking disequilibrium along pathways, comparing runs across “worlds” or storylines, and surfacing outcomes that are unlikely under a central view but plausible when how factors interact is uncertain. Cross-Impact Balance (CIB) is a well-established method for generating pathways. This paper extends CIB to formalise and implement disequilibrium, extremes, and unexpected outcomes in pathway runs, and defines four run types that respectively emphasise one-off shocks, extremes under alternative regimes, influence-structure uncertainty that widens over time, and exogenous shocks as a baseline for comparison. The approach is applied to a socio-technical decarbonisation pathway for illustration. Together, these extensions support stress-testing, comparison across storyline or regime assumptions, and exploration of rare or surprising futures, and help analysts distinguish results that are stable across those assumptions from those that depend on structural uncertainty about the influence table.

1. Introduction

Scenario and pathway analysis is widely used to explore how the future might unfold under different assumptions. Frameworks such as the Shared Socioeconomic Pathways provide qualitative narratives that describe world futures and are combined with quantitative pathways for integrated assessment (Riahi et al., 2017; O'Neill et al., 2017). Similar approaches are used in energy and climate context scenarios, where storylines are built from consistent factor combinations and then linked to various assessment models (Nakicenovic and Swart, 2000; O'Neill et al., 2014; Guivarch et al., 2017).

Nevertheless, several dimensions remain underexplored in practice. These are classified following McCollum et al. (2020). First, the quantification of short-lived departures from equilibrium along pathways is methodologically underdeveloped. That is, policy analysis would benefit from knowing how far a pathway is from a stable, self-consistent state at each step. Second, the exploration of pathways under alternative structural assumptions such as disruptive drivers or extremes is often limited. Comparing outcomes under distinct “worlds” or storylines is not routinely separated from other sources of variation. Third, the exposure of rare or overlooked outcomes when the influence structure is uncertain receives insufficient attention, including tail events

and what has been termed grey or black swans (Walker et al., 2003; Lin and Emanuel, 2016; Taleb, 2007).

When scenario work informs policy, a single central trajectory or a single band of outcomes is often insufficient (Weitzman, 2009; Rosen, 2025). Decision-makers need to understand both how much outcomes can vary and what kind of cause lies behind that variation. A min-max band added to a central trajectory shows the range of outcomes when inputs are varied, but it does not separate the effect of a one-off shock from the effect of a different storyline or from the effect of uncertainty about how the system works. The band remains an uninterpretable range. To the best of the author's knowledge, existing scenario pathway and storyline work (including overviews of quantitative scenario techniques (Guivarch et al., 2017)) does not provide a single framework that separates these influences so that the kind of shock or storyline or uncertainty behind variation can be identified and communicated (Lempert et al., 2003; McCollum et al., 2020). The pathway framework set out in this paper supports such attribution when scenarios are built as distinct runs, each activating a different combination of shock, regime, uncertain influence structure, and exogenous disturbances. Where an illustration bundles several of these, interpretation follows that joint setup.

E-mail address: a.ross@fz-juelich.de.

<https://doi.org/10.1016/j.techfore.2026.124893>

Received 1 April 2026; Received in revised form 29 July 2026; Accepted 19 September 2026

Available online 29 September 2026

0040-1625/© 2026 The Author. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Open access funding enabled and organized by Projekt DEAL.

Cross-Impact Balance (CIB) provides a framework in which these gaps can be addressed. In non-technical terms, CIB represents a system as a small set of factors (descriptors), each taking one of a few ordered qualitative states, together with a table of expert judgements on how each factor's state pushes the others towards or away from their own states. This table is the cross-impact matrix (CIM). From a starting combination of states, factors are repeatedly updated towards the states best supported by the rest, moving the combination towards a self-consistent one, and over a sequence of periods this yields a scenario pathway (Weimer-Jehle, 2006). This multi-period use is common in CIB applications (Schweizer and Kriegler, 2012; Weimer-Jehle et al., 2020; Schweizer, 2020).

This paper extends CIB along three dimensions, namely disequilibrium along pathways (short-lived departures from a stable state), pathways under distinct regimes or “worlds” (extremes), and unexpected outcomes when the influence table (the CIM) is uncertain. The aim is to provide a single framework in which pathway outcomes can be attributed to the mechanism that produced them, so that transient disturbances, regime differences, and influence-structure uncertainty are each separately visible. Throughout, a storyline is the qualitative account of how a future unfolds, a “world” is the alternative context that the storyline describes, and a regime is its formal representation as a particular cross-impact matrix. A shock is an exogenous disturbance applied to a pathway, whereas uncertainty refers to incomplete knowledge of that matrix, which is then sampled. These extensions operate within a known representational constraint, namely that CIB's natural unit is a self-consistent scenario on a workshop-sized descriptor set, not a complete dynamical-systems model of the system being analysed. The extensions are partial remedies for the epistemic risks that arise from this, with the full treatment deferred to Section 6.

In this paper, a systematic approach is set out that extends CIB in this way. Four run types are developed and illustrated, three of which build on the new extensions and one on existing stochastic exogenous shocks retained for comparison. Each type addresses a different cause of variation (transient shock, a joint regime-and-shock specification, system uncertainty, or random disturbances alone), so that pathways are explored in four distinct ways and both the extent and the kinds of causes of variation in outcomes are visible. The distinction is qualitative. Each run type is a distinct combination of mechanisms, so numerical differences between types are not interpretable as the quantitative effect of a single factor. The quantification of disequilibrium along pathways is formalised here and is not provided by other CIB or scenario pathway methods in this form. Formal disequilibrium metrics (consistency margin, burden, partial convergence) and pathway-level sampling from an uncertain CIM are not offered in existing CIB pathway or scenario overviews (Guivarch et al., 2017).

The four run types comprise runs under a one-off stronger disturbance in a single period (stress-testing and resilience), runs under a joint regime-and-shock specification (comparison across such joint specifications), runs under uncertain influence structure (rare or surprising endpoints), and runs with exogenous stochastic disturbances only (baseline of pure noise). Including this fourth type as a distinct run provides a baseline against which the added effect of shock, the joint regime-and-shock specification, and structural uncertainty can be compared. Together they support stress-testing, comparison across joint regime specifications, and the identification of rare or surprising futures.

The three extensions to CIB (disequilibrium, extremes, and unexpected outcomes) are the main contribution of this paper, which is, to the author's knowledge, the first to combine these dimensions within a single CIB pathway framework. The approach is general and applicable to any CIB application. For illustration, it is applied here to a socio-technical pathway based on the data of Ross and Ross (2026), which uses AI-simulated expert panels for its generation. Each run type can involve tail-like or rare outcomes in a different sense (realisation tails, storyline or stress tails, or structural-uncertainty tails).

The four types together allow one to assess whether results are robust across joint regime specifications, depend on the chosen joint extremes specification, or only appear when structural uncertainty is included (Lempert et al., 2003; Taleb, 2007). Thematically, realisation tails can also surface rare or disruptive pathway outcomes in the spirit of grey swans, which are plausible but often neglected (Lin and Emanuel, 2016). The terms “grey-swan-type” and “black-swan-type” are reserved here for the third run type (structural-uncertainty tails) so that the cause of rarity remains clearly attributed.

This paper is organised as follows. Section 2 provides further non-technical detail on the scenario literature, the evidence for the limitation of single-band presentation, and the four run types and their interpretation. Section 3 outlines the CIB method and the novel extensions in formal terms. Section 4 describes the simulation scenarios and points to the Appendix A and the Supplementary Material, which specify the simulation set-up and notation. Section 5 reports simulation outcomes for the four run types. Section 6 discusses and concludes.

2. Conceptual background

The following develops the conceptual basis outlined in the introduction. It adds the scenario and pathway literature, the limitation of single-band presentation, and the four run types and their interpretation. Section 3 then provides the technical detail. As described in Section 1, CIB represents a system through descriptors that take ordered qualitative states, with pairwise expert judgements on how each descriptor's state bears on the others collected in the cross-impact matrix (the influence table). A succession rule then updates each descriptor towards the state best supported by the others in their current configuration.

The scenario pathways and storylines literature stresses trajectories over time and narrative assumptions about alternative “worlds” to explore how the future might unfold. Reviews of scenario techniques for energy and environmental research emphasise the need to broaden the capacity to deal with complexity and uncertainty and to combine qualitative and quantitative elements (Guivarch et al., 2017; Kwakkel and Pruyt, 2013; Bai et al., 2016). That literature supports comparison across narrative assumptions and stress-testing of strategies under different futures (Lempert et al., 2003).

The gaps outlined in the introduction are reflected in this literature. McCollum et al. (2020) and Weitzman (2009) suggest that the future is shaped as much by extremes as by central trends. Work on resilient climate mitigation, for example, argues for expanded futures analysis rather than relying on conventional modelling alone (Gambhir and Lempert, 2023). Broader exploration of the scenario space is supported by evidence that many scenarios are needed for robust energy system insights and that single or narrow scenario sets can miss important outcomes (Frey et al., 2025). Representativeness of scenario sets is argued to require attention to “unknown unknowns” and to structural uncertainty in how factors interact (Eriksson et al., 2022; Walker et al., 2003). The third run type specifically addresses this gap by exposing endpoints that a central matrix would underweight.

The limitation of presenting outcomes as a single band has been widely discussed (Lempert et al., 2003; Rosen, 2025). A min-max band does not separate the effect of a one-off shock from that of a different storyline or from uncertainty about how the system works. In risk assessment, nonprobabilistic cross-impact approaches have been criticised for not guaranteeing conservative risk estimates and for consistency criteria that can admit very few or no scenarios (Salo et al., 2022). Such simplified approaches are widespread in scenario pathway practice. McCollum et al. (2020) note that they are (too) common in energy scenario modelling.

CIB is commonly used for the generation of multi-period scenario pathways (Schweizer and Kriegler, 2012; Kemp-Benedict et al., 2019; Schweizer, 2020). Pathways evolve from a cross-impact matrix over a sequence of periods so that each period yields a scenario. Linked

together, these scenarios form a pathway. The approach is used across many domains, including socio-technical scenarios for net-zero and other transformation pathways, and supports scenarios that combine qualitative storylines with quantitative models in plausible societal contexts for model assessment (Weimer-Jehle et al., 2020; Guivarch et al., 2017). Existing CIB pathway work has focused on consistency and multi-scale linking (Guivarch et al., 2017; Kemp-Benedict et al., 2019).¹

Standard CIB pathway work generates scenario sets by identifying consistent configurations (combinations of factor states in which each factor is in the state best supported by the others). Each configuration represents a distinct “world”. This paper takes a complementary approach, running ensembles under each of the four run types and reporting pathway distributions. The extremes run type is the closest analogue to a conventional scenario set, producing three regime-based alternatives, but the output under each regime is a pathway distribution, not a single consistent configuration. The extensions developed below add dimensions that this literature does not yet address. Within this CIB pathway framework, the four run types developed in this paper serve distinct analytical purposes. The following discusses each in turn.

The first of the four run types addresses what happens when the system is subjected to a one-off stronger disturbance in a single period. Its purpose is stress-testing, namely to make visible the response to a short-lived disruption and to quantify how far the pathway is from a stable, self-consistent CIB state at each step. The insights are the resilience or vulnerability of pathways to an amplified event and the extent of transient disequilibrium along the path. Outcomes can include rare realisations when shock sequences are extreme (realisation tails). Decision-makers can use this to identify where buffering or contingency planning is most needed.

The second type addresses what happens when the analyst assumes a different underlying storyline (for example a constrained “world”) and holds it fixed for the whole pathway. Each storyline is represented by a distinct influence structure (regime). The purpose is to compare outcomes across storylines and to assess sensitivity to the choice of storyline. The insights are which outcomes are robust across alternative regimes or storylines and which depend on the assumed world, which is useful in contested policy contexts.² This type explores storyline or stress tails (outcomes under deliberately adverse or “tail” storylines).

The third type addresses what happens when the influence table is not treated as known but is instead sampled from an uncertain specification, with uncertainty optionally widening over time and disturbances applied. The purpose is to expose pathway endpoints that are rare under a single central matrix but plausible under structural uncertainty, including tail-risk and grey-swan outcomes (Walker et al., 2003; Lin and Emanuel, 2016; Taleb, 2007). The insight is the share and identity of such rare but plausible states (endpoints a central matrix may underweight) when the influence structure is uncertain (structural-uncertainty tails). The aim is to broaden the set of futures considered and to avoid overconfidence in a single view.

¹ Other cross-impact approaches derive scenario probabilities and integrate with risk or decision analysis (Salo et al., 2022; Roponen and Salo, 2024), but typically do not provide pathway-level dynamics, disequilibrium metrics, or the separation of shock, storyline, and structural uncertainty in a single run framework (Guivarch et al., 2017). That overview and the CIB developments it discusses (e.g. linked CIB for multi-scale consistency Kemp-Benedict et al., 2019) address different dimensions rather than pathway-level disequilibrium or uncertain-CIM sampling.

² Robustness means similar outcome distributions under different assumed regimes or storylines. Material dependence means those distributions differ across assumptions (Guivarch et al., 2022; Lempert et al., 2006). Structural shock intensity and other disturbance settings may be specified per regime. Where they co-vary with the regime matrix, pathway contrasts reflect that joint specification. Attribution to storyline or influence structure alone requires that such settings be held fixed across regimes or varied in a separate factorial design.

The fourth type uses a single nominal backbone CIM and applies only exogenous structural and dynamic stochastic disturbances, without uncertain-CIM sampling or regime change (the realised matrix still varies period-to-period when structural shocks are drawn). Its purpose is to isolate the exogenous shock channels from matrix-sampling uncertainty and storyline regimes, and thereby to provide a reference run type for qualitative contrast with the other three. It can yield realisation tails when shock sequences are extreme. This run type is outlined in Ross and Ross (2026) and is therefore not developed further in this paper. The formal treatment of the four types and their implementation is given in Section 3.

3. Methods

The following subsections outline the method in sequence. First, the standard CIB framework and multi-period pathways are presented, and then the three extensions (disequilibrium, extremes, and unexpected outcomes), each defined and described in terms of the implementation within PyCIB (an open-source CIB analysis package) (Ross, 2026). The purpose and interpretation of each run type, and the distinction between realisation, storyline, and structural-uncertainty tails, are as set out in Section 2. The following subsections give the formal definitions and implementation.

3.1. Standard CIB and multi-period pathways

At the core of the CIB method is the cross-impact matrix C , the table of pairwise influence scores between descriptor states. Descriptors are factors, each with a set of discrete ordinal states. Following Weimer-Jehle (2006), a scenario z is one state chosen per descriptor (i.e. one state per factor), and z_j denotes the state of descriptor j in scenario z . For such a scenario, the impact score of target descriptor j in state l is

$$\theta_{j,l}(z) = \sum_{i \neq j} C_{i \rightarrow j}(z_i, l), \quad (1)$$

where i runs over descriptors, z_i is the state of source descriptor i in scenario z , l denotes a possible state of descriptor j , and $C_{i \rightarrow j}(z_i, l)$ is the impact of source descriptor i in state z_i on target descriptor j in state l .

In regime-aware runs the matrix used in this formula is the active matrix $M_{r(t)}$ for period t (Section 3.3). Otherwise, it is the single CIM C . The impact balance for descriptor j is the vector of these scores over all its states. A scenario is consistent if, for every descriptor j , the chosen state z_j attains the maximum (ties allowed). The implementation treats a descriptor as inconsistent only if some other state has a strictly greater score. Equality or near-equality within a small numerical tolerance is treated as a tie (consistent). That is, for each descriptor j and for every state l of that descriptor,

$$\theta_{j,z_j}(z) \geq \theta_{j,l}(z) \quad \text{for all } l. \quad (2)$$

In this application, “consistency” means the CIB rule (Eq. (2), under which the chosen state attains the maximum impact score), not consistency in the sense of a joint probability distribution over scenarios satisfying elicited constraints, as in probabilistic cross-impact methods. When representativeness or scenario-set consistency is discussed in the broader literature (for example by Eriksson et al. (2022)), that may refer to transparency, plausibility, or persistence of scenarios. Here the term is used only in the CIB sense. State shares or proportions reported in the results (for example in timeline figures) are frequencies over pathway runs, not scenario probabilities from a single joint distribution. The present approach does not provide scenario probabilities or conservative risk bounds in the sense of Salo et al. (2022).

The succession operator produces from a scenario z a successor scenario z' by setting each descriptor to a state that attains the maximum impact score given z . That is, for each descriptor j , $z'_j \in \arg \max_l \theta_{j,l}(z)$ (with a tie-breaking rule when the maximum is not unique).

Starting from an initial scenario, succession is applied repeatedly. If after some number of steps the scenario equals its successor, a fixed-point attractor has been reached. If the successor was already encountered at an earlier step, a cycle attractor is identified, namely a finite repeating sequence of scenarios. Formally, a fixed point satisfies $z' = z$. A cycle is a list of scenarios $(z^{(1)}, \dots, z^{(m)})$ such that the successor of $z^{(m)}$ is $z^{(1)}$. The succession dynamics described here are related to the Markov chain formalisation of [Schweizer et al. \(2023\)](#), which characterises pathway transitions between socio-economic states with probabilistic diagnostics and information-theoretic measures. The extensions developed in this paper address transient disequilibrium, regime-based extremes, and uncertain-CIM sampling rather than the stationary properties of the transition process.

The implementation may cap within-period iteration at a maximum of $K_{\max}$ steps. When that cap is applied and neither a fixed point nor a cycle has been reached within $K_{\max}$ steps, the last scenario $z^{(K_{\max})}$ is taken as the period outcome (partial convergence). That scenario need not satisfy Eq. (2). In this analysis, the succession operator is local in that one descriptor is updated per step (the one with the largest gap between current and maximum impact score), so each step counts as a single-descriptor update.

In “dynamic” runs,³ a time grid T of periods is defined (e.g. $T = \{t_1, \dots, t_N\}$). An element $t \in T$ is referred to as a period ([Ross and Ross, 2026](#)). A pathway is the sequence of scenarios $(z(t))_{t \in T}$ realised at each period. Within each period, the model applies succession from the current state (or from an initial scenario in the first period), optionally after cyclic transitions for selected descriptors and threshold-based modifications to the active matrix.

The threshold-triggered CIM changes and cyclic descriptors (state in the current period dependent on the previous period) are implemented in PyCIB in line with the non-linear transformation pathway approach of [Vögele et al. \(2019\)](#), among others. In that approach, a tree of consistent scenarios is built per sub-period. Here an ensemble of single-path simulations is run, with disequilibrium, regimes, and uncertain CIM added. Structural and dynamic shocks (random perturbations to the CIM and to impact scores during succession, respectively) are described in Section 3.5 and are used alongside these extensions in this analysis.

3.2. Disequilibrium

This extension formalises the quantification of short-lived departures from equilibrium along pathways (Section 2). The following metrics define the implementation. Disequilibrium is defined here as the situation in which the realised scenario at a period is not fully characterised by a single equilibrium condition under the active matrix. That is, the state actually produced may be transiently inconsistent, in that at least one descriptor is not in the state that maximises impact balance for that period’s matrix. For a scenario z and descriptor j , the consistency margin (the margin by which the chosen state leads or lags the best alternative) is

$$\mu_j(z) = \theta_{j,j}(z) - \max_{l \neq j} \theta_{j,l}(z), \quad (3)$$

³ Dynamic CIB is not dynamics in the sense economists validate against data. Used honestly, it is structured expert judgement stepped through time on a thin cross-impact scaffold for exploration and stress-testing. It is not a state-space model with identified propagation, micro-founded adjustment, or projections that assume a validated mapping from states to observations. Mistaking elicitation for identification, or treating these trajectories as policy-grade outputs, overstates what the scaffold can bear. Parts of the CIB literature have also pushed the method towards structural policy appraisal and similar uses where the elicited layer cannot carry the burden. Where the same artefacts are presented as quantitative dynamic analysis of real adjustment, that use is inadmissible by ordinary standards for structural dynamics and policy appraisal.

where $\theta_{j,j}(z)$ is the impact score defined in Eq. (1). The global consistency margin is

$$\mu(z) = \min_j \mu_j(z). \quad (4)$$

If descriptor j has only one state, the set $\{l \neq j\}$ is empty. In that case $\mu_j(z)$ is taken as zero (and such descriptors are included in the minimum in Eq. (4)). The scenario is consistent under the active matrix if and only if $\mu(z) \geq 0$. When within-period succession is capped (partial convergence), the period outcome may have $\mu(z) < 0$.

The implementation quantifies disequilibrium further using distances and a time-to-consistent measure. The [Hamming \(1950\)](#) distance between two scenarios z and z' (same descriptor set) is the number of descriptors in which they differ (hence d_H is integer-valued, and the implementation may report it as a float):

$$d_H(z, z') = \sum_j \mathbf{1}[z_j \neq z'_j]. \quad (5)$$

For a scenario z and the active matrix, the distance to the nearest attractor is the minimum of $d_H(z, z^*)$ over all scenarios z^* that are fixed-point attractors or belong to a cycle attractor. When exact enumeration of attractors is infeasible, the implementation reports the Hamming distance from z to the attractor reached by succession from z (rather than the minimum over all attractors). The distance to the consistent set is the minimum of $d_H(z, z^c)$ over all scenarios z^c that satisfy Eq. (2). When exact enumeration of consistent scenarios is infeasible (e.g. in this application, 3^{15} scenarios), the implementation reports the Hamming distance from z to the first consistent scenario reached by succession from z .

The time to the consistent set is the smallest number of succession steps $k \geq 0$ such that the scenario reached after k steps is consistent. If no such k exists within the iteration cap, the quantity is undefined. The implementation reports this quantity as “time to equilibrium” in the pathway diagnostics. The same diagnostics may label the Hamming distance to the attractor reached by succession as “distance to equilibrium” (distinct from the distance to the consistent set). The cumulative disequilibrium burden along a pathway is

$$B((z(t))_{t \in T}) = \sum_{t \in T} \max(0, -\mu(z(t))). \quad (6)$$

These metrics (margins, distances, time to consistent set, burden) are computed per period from the realised scenario and the active matrix and reported alongside the realised pathway.

The distance to the nearest attractor, tracked here as a period-by-period diagnostic, is related to tipping-distance ideas in the complexity-science literature. [Campfens et al. \(2025\)](#) apply related attractor-basin reasoning in the energy transition context, using feedback loop analysis with succession to identify intervention points from which cascades towards desired attractors can be initiated.

3.3. Extremes

As outlined in Section 2, pathways can be run under different assumed regimes. The extension is implemented as regime-aware dynamics. The model admits multiple cross-impact matrices in this setting. The user defines a set of named regimes, each associated with its own CIM, so there are as many distinct CIMs as regimes.

A regime is a named structural mode. Each regime r has a base matrix $M_{\text{base}}^{(r)}$ (or uses a shared fallback) and an optional ordered list of modifiers. The active matrix M_r for regime r is obtained by applying those modifiers in sequence to the base matrix.

$$M_r = (\varphi_L \circ \dots \circ \varphi_1)(M_{\text{base}}^{(r)}), \quad L \geq 1; \quad M_r = M_{\text{base}}^{(r)} \quad \text{when } L = 0. \quad (7)$$

Here $\circ$ denotes functional composition, so that $(\varphi_k \circ \varphi_{k-1})(M)$ is the matrix obtained by applying φ_{k-1} to M and then φ_k to the result, with modifiers applied in order φ_1 , then φ_2 , ..., then φ_L . For each period $t \in T$, the model uses the active matrix of the current regime for

impact scoring and succession, and all consistency and disequilibrium measures for that period refer to $M_{r(t)}$. In the simplest configuration, the run is assigned an initial regime r_0 and the regime is held fixed for the whole pathway, with $r(t) = r_0$ for all $t \in T$, so the pathway answers the question of what happens when the world follows that regime's influence structure.⁴

Defining extremes as regime-based storylines (alternative CIMs) rather than as events or as diversity bounds under a single CIM supports comparison of robustness versus storyline-dependence and makes stress-testing under adverse influence structures a distinct, interpretable run type in principle. As Ross and Ross (2026) note, in much existing CIB work that refers to extremes, the term denotes either extreme events (such as droughts, floods, or other rare high-impact occurrences) or the maximally diverse scenarios that bound the scenario space under one matrix. Regime-based extremes as a run type are not modelled in that way. The approach taken here is therefore clearly distinct from that literature. In the general formalism, regime (fixed CIM), structural shocks, and uncertain-CIM sampling are separate optional layers (Sections 3.4 and 3.5).

3.4. Unexpected outcomes

The influence table can be treated as uncertain (Section 2), so that many plausible matrices are tried across runs and the full range of pathway endpoints, including rare or surprising ones, is visible. Unexpected outcomes are pathway endpoints or trajectories that arise when the matrix is sampled from an uncertain specification (Walker et al., 2003; Kwakkel and Pruyt, 2013).

Structural uncertainty is implemented by sampling the CIM and running pathways, in contrast to approaches that resolve expert inconsistency via optimisation to a single consistent distribution (Salo et al., 2022; Roponen and Salo, 2024).⁵ When the matrix supports sampling (e.g. from judgement distributions with an explicit uncertainty or confidence encoding), each matrix cell is drawn independently from a distribution around a point estimate.

A cell is indexed by the source descriptor, source state, target descriptor, and target state. The point estimate for that cell is $\bar{c}$, and the scale parameter for its confidence code c (e.g. derived from expert confidence or from a prior) is $\sigma(c)$. An optional multiplicative scale λ , which may depend on run or period (e.g. $\lambda(t)$), scales the spread of that distribution. In the implementation used here, the period schedule is linear in the period index, $\lambda(t_i) = 1 + i$ for $i = 0, \dots, N - 1$ (periods indexed from zero). The sampled value for the cell is then

$$\bar{c} = \text{clip}(\mathcal{N}(\bar{c}, (\sigma(c) \cdot \lambda)^2), c_{\min}, c_{\max}), \quad (8)$$

where $\mathcal{N}(\bar{c}, (\sigma(c) \cdot \lambda)^2)$ denotes the normal distribution with mean $\bar{c}$ and variance $(\sigma(c) \cdot \lambda)^2$ (standard deviation $\sigma(c) \cdot \lambda$), and $\text{clip}(\cdot, c_{\min}, c_{\max})$ denotes clipping to the admissible CIM score range $[c_{\min}, c_{\max}]$ (when the score range is unbounded, clipping may be omitted). The full sampled matrix $\bar{C}$ is built by drawing every cell in this way (with a specified random seed or stream for reproducibility). A fresh matrix can be drawn once per run (so the whole pathway uses one sampled CIM) or once per period (so the influence structure may change from period to period). In the unexpected-outcomes run type, the latter is used, re-sampling each period with period-specific $\lambda(t)$. With the implemented schedule $\lambda(t_i) = 1 + i$, uncertainty widens over time.

⁴ Alternatively, a regime transition rule can be supplied. At each period the rule may depend on the current scenario and path history to select $r(t)$, so the pathway may switch between regimes and hence between CIMs over time. The implementation records the active regime and matrix provenance per period so that results can be attributed to the chosen regime.

⁵ In those approaches, a “consistent distribution” is a single joint probability distribution over all scenarios that satisfies the elicited constraints. Here, uncertainty is represented by sampling many CIMs and running pathways, with no single joint distribution over scenarios.

Different runs therefore realise different influence structures. Combined with succession and optional regime or shock mechanisms, this produces an ensemble of pathways whose distribution may include states that would be rare under the central matrix. Here the source of tail risk is uncertainty in the influence structure (Section 2), so the ensemble can expose grey- and black-swan-type outcomes (Lin and Emanuel, 2016; Taleb, 2007). Reporting is at pathway level. The analyst inspects the distribution of endpoint states or of full trajectories across runs to quantify the share of runs that realise each outcome. Structural and dynamic shocks (additive perturbations to the matrix and to impact scores during succession) can be combined with sampled matrices for stress-testing.

3.5. Exogenous stochastic shocks

This analysis uses two kinds of exogenous stochastic shocks that are implemented and outlined in Ross and Ross (2026). Structural shocks are random perturbations applied to the cross-impact matrix. The influence table is additively (or otherwise) disturbed, then used for scoring and succession in that period or run. Structural shocks thus represent unexpected changes to the system's influence structure. Dynamic shocks, by contrast, perturb the balance for each descriptor when computing the successor scenario (e.g. by a draw from an autoregressive process over periods), so that the step-by-step dynamics are disturbed rather than the matrix itself.

The perturbation at each descriptor-state and period follows the AR(1) process $\eta_t = \rho\eta_{t-1} + u_t$ with $u_t \sim \mathcal{N}(0, (1 - \rho^2)\tau^2)$, so that the long-run standard deviation is τ . Both shock types serve to stress-test pathways under random disturbances. In this analysis, structural and dynamic shocks are used where appropriate alongside the three extensions described in the preceding subsections, and the extensions add complementary diagnostics and scenario dimensions that the shocks do not themselves provide.

4. Simulation scenarios

The general method (Section 3) is independent of the choices below and would apply to any cross-impact matrix. The illustrative example here focuses on socio-technical decarbonisation pathways and uses data sourced from Ross and Ross (2026). Germany's socio-technical energy transition to net zero by 2050 is the subject of that storyline. The 15 descriptors comprise policy stringency, EU policy alignment, renewables deployment, grid development, permitting pace, technology costs, investment availability, fossil fuel prices, energy security, public acceptance, land-use availability, electrification pace, hydrogen's role, and demand-side factors, with Decarbonisation Outcome as the timeline descriptor. The baseline pathway is initialised in 2025 at a state of medium policy ambition and a lagging grid. That is, the starting configuration reflects moderate policy commitment and underdeveloped grid infrastructure. Thereafter, a period of demand adjustment and fossil-price pressure is traversed, whilst the decarbonisation outcome remains at a medium level. A net-zero-aligned state is subsequently reached, supported by strong renewables deployment and expanded grid capacity. The consolidation phase is characterised by declining technology costs, strong policy stringency, and ample investment. The pathway concludes in 2050 at a net-zero-aligned endpoint with some residual grid and public acceptance constraints. The full and detailed endpoint storyline, of which the foregoing is a condensed qualitative account, is provided in Ross and Ross (2026).

The technical simulation specification adopted in this paper is described next. Impact scores are defined on a bounded scale and, in this application, that scale is the interval $[-3, +3]$ (as per the underlying data). Any sampled or perturbed cell value is clipped to this range. The uncertain-CIM variant uses confidence-coded expert judgements (Ross and Ross, 2026). Each cell has a point estimate $\bar{c}$ and a confidence code c that maps to a standard-deviation scale $\sigma(c)$, so that

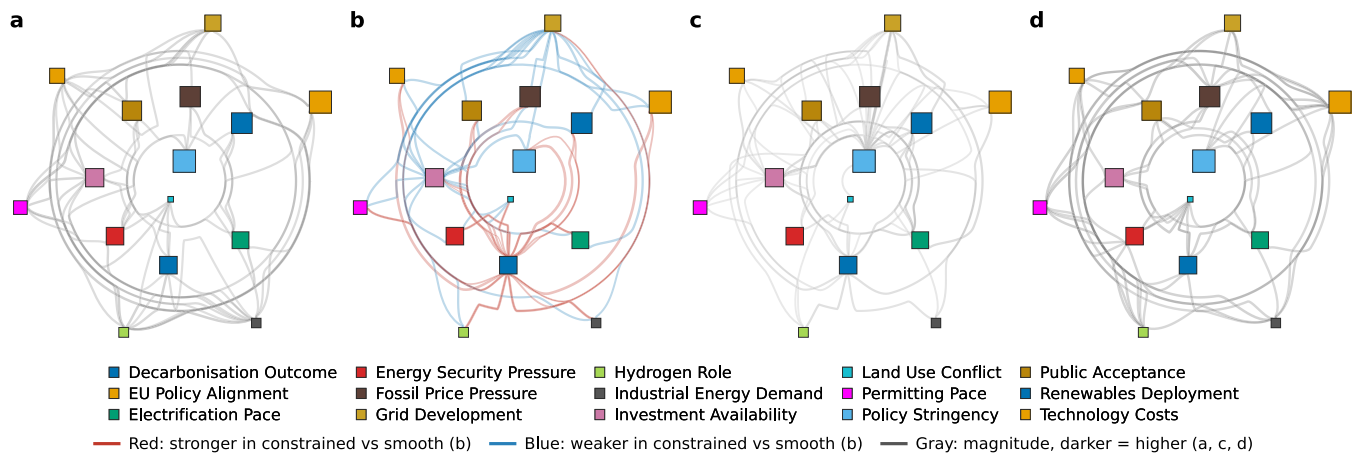


Fig. 1. Structural view of the cross-impact matrix under four contrasts (same layout and node-size scale, out-strength from the base CIM; edges in the upper quartile by weight, i.e. at or above the 75th percentile, are shown). a, Structural-shock illustration: mean absolute confidence-calibrated structural shock per directed edge (over state pairs), one draw; edge width and darkness show magnitude. b, Extremes: absolute impact difference between constrained-transition and smooth-transition regimes, $|\text{constrained-transition} - \text{smooth-transition}|$. c, Regime spread: range of mean-absolute impact across smooth-transition, turbulent-transition, and constrained-transition CIMs. d, Unexpected: standard deviation of impact strength across samples from the uncertain matrix (same edge width and darkness convention as a).

the sampled value for the cell is drawn from $\mathcal{N}(\bar{c}, (\sigma(c) \cdot \lambda)^2)$ and then clipped to $[-3, +3]$. Three regimes are used in the extremes run type, referred to throughout as the smooth-transition, turbulent-transition, and constrained-transition regimes. The smooth-transition regime uses the unmodified nominal CIM paired with the lowest structural shock scaling. The turbulent-transition regime uses a pro-transition amplified CIM with an intermediate shock scaling. The constrained-transition regime uses a constraint-scaled CIM with the highest shock scaling, so convergence is slowest under this regime. Each regime is paired with a regime-specific structural shock scaling α (see Appendix A).⁶

The simulation runs four workflows. The full set of parameter and run-type choices, including the data source, time grid, initial scenario, run count, partial convergence and succession settings, structural and dynamic shock parameters, regime definitions, and the unexpected-outcome threshold, together with notation and assumed values, are given in Appendix A and in the Supplementary Material (S1). Specifically, Table S1 defines the notation used in the main text and in Appendix A and, where applicable, gives the simulation assumed values.

5. Results

Outcomes from four run types are reported in this section, each corresponding to a distinct combination of assumptions and disturbances. A selection of the results is presented in two figures. Fig. 1 provides a structural view of the cross-impact matrix. All descriptors and their pairwise influences are shown in a single layout, with four panels, each showing a different contrast derived from perturbations of the cross-impact matrix. The descriptors are not standalone, in that they interact through the cross-impact matrix, so that the state of each descriptor influences the impact balance of others, and succession updates the scenario step by step. The system is therefore changing over time along each pathway. Fig. 2 focuses on a single descriptor, the timeline descriptor (Decarbonisation Outcome), and shows how the ensemble share of each of its three states evolves over time across pathway runs. In each panel, one line is plotted per run family (with three lines for the extremes), so that pathway outcomes under the different assumptions can be compared across the time horizon.⁷

⁶ In the empirical pathways reported here, the regimes each use a fixed regime-specific CIM and a fixed regime-specific structural shock scaling α (Table S1 in the Supplementary Material). Differences among those three ensemble lines therefore jointly reflect storyline-style changes to the matrix and changes in shock intensity, and should be read with that design in view.

Outcomes differ between run types because each combines a different set of mechanisms. These concern whether the influence structure is sampled as uncertain, held to a fixed regime-specific backbone, or held to a single nominal backbone subject to structural and dynamic shocks, together with which shock types and amplification rules are applied. Each of the four run types can be put in the form of a simple question (e.g. “What if one year is stressed much more strongly than the others?” versus “What if alternative joint regime-and-shock specifications are contrasted, for example a smoother transition against stronger headwinds?” versus “What if the strength with which factors affect one another is only partly known?” versus “What if a single nominal influence picture is assumed, whilst structural and dynamic shocks keep perturbing pathways?”). In order, these questions map to the run types named in the legend of Fig. 2, namely Disequilibrium, Extremes, Unexpected, and Exogenous. This yields outcome distributions that are interpretable, in that differences between run types can be qualitatively attributed to the combination of mechanisms in play, and both the extent of outcome variation and the scenario logic that produced it are visible.

Fig. 1 shows where variation arises in the influence structure. In this visualisation, each descriptor is a coloured node and each directed cross-impact is an edge, where edge width and darkness encode the quantity described for each panel, and only edges in the upper quartile by weight are shown. Node size encodes out-strength from the base CIM (the sum of absolute impact strengths on outgoing edges from each descriptor). Panel (a) shows edges most affected by a calibrated structural-shock draw (an illustration that matches the disequilibrium run type in structure but not in parameter values). Panel (b) shows where the constrained-transition regime differs from the smooth-transition regime. Panel (c) shows where regime spread is largest across the three regime specifications (smooth-transition, turbulent-transition, and constrained-transition). Panel (d) shows where impact strengths are most sensitive to CIM sampling under uncertainty.

Fig. 2 compares pathway outcomes across run types. The three panels give the ensemble share of each state of the timeline descriptor

⁷ The Extremes (Smooth transition, Turbulent transition, and Constrained transition) legend labels identify the three extremes regimes by their influence structure and shock intensity. The three states of the Decarbonisation Outcome descriptor, High emissions, Medium emissions, and Net-zero aligned, are the outcomes tracked on the vertical axis of Fig. 2. The two sets of terms are distinct.

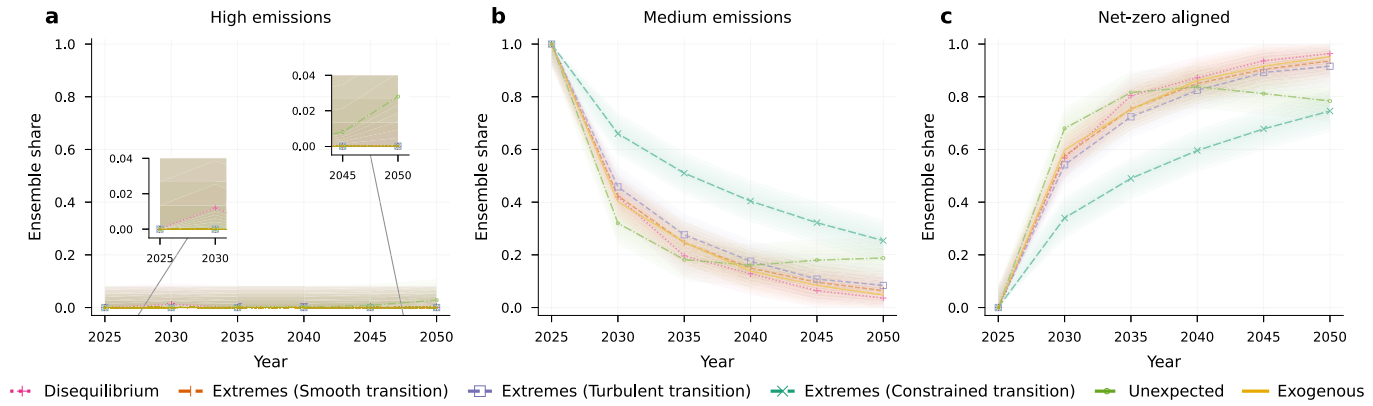


Fig. 2. Pathway ensemble shares for the timeline descriptor (Decarbonisation Outcome): one state per panel (a–c), all run types overlaid. a, High emissions; b, Medium emissions; c, Net-zero aligned. Lines: Disequilibrium (nominal backbone CIM, structural shocks, amplified event period); Extremes (three extremes regimes: fixed regime-specific CIM paired with regime-specific scaling α ; no sampling, no switching); Unexpected (uncertain CIM, per-period re-sampling with $\lambda(t)$, dynamic and structural shocks); Exogenous (nominal backbone CIM, structural and dynamic shocks; no regime or sampling uncertainty). Lines are ensemble proportions; shaded bands are derived from 99% Wilson score intervals at each period ($n = 500$ runs per type), with a fixed half-width of 0.08 applied for visual legibility, drawn as nested layers.

(Decarbonisation Outcome) over time for each run type, with 99% Wilson score intervals (confidence intervals for binomial proportions that remain well-behaved near zero or one) over 500 runs. Stress-testing is reflected in the disequilibrium and exogenous ensembles. The former includes one amplified event period (2030), so the transient response to a stronger perturbation is visible. The latter shows structural and dynamic shocks about a nominal backbone CIM without uncertain-CIM sampling or regime change.

The sharp 2030 increase in Net-zero aligned ensemble share visible across all run types reflects a common baseline dynamic. All share the same initial condition (middle ordinal state per descriptor), the same period grid and succession settings, and are anchored in a baseline CIM whose attractor structure is pro-transition (Appendix A). It follows that early-period movement is broadly aligned, whereas the workflow-specific mechanisms (event amplification, regime-specific CIM and α , uncertain-CIM sampling, and dynamic shocks) primarily determine the extent of spread, persistence, and tail outcomes thereafter. The same shared pull also helps explain why High emissions remains near zero in most workflows. Under the common initialisation and succession setup, trajectories are directed towards the dominant pro-transition basin. In this calibration, non-zero High-emissions mass is concentrated in the unexpected-outcomes workflow as a small late tail, whilst the others remain at or near zero.

The three extremes ensembles correspond to the smooth-transition, turbulent-transition, and constrained-transition regimes, as defined in Section 4. Each ensemble combines a regime-specific CIM with regime-specific structural scaling α . The CIMs are the unmodified nominal matrix, a pro-transition amplified copy, and a constraint-scaled copy. These ensembles do not isolate storyline effects alone. Divergence among the lines reflects joint change in influence structure and shock intensity. Where the three lines diverge, outcomes are sensitive to that joint extremes specification. Close alignment, by contrast, indicates that outcomes do not vary materially across those three regimes. The three ensembles do not correspond to a design in which structural scaling α is held fixed whilst only the storyline matrix changes.

The unexpected-outcomes scenario (uncertain CIM, re-sampled per period with period-widening $\lambda(t)$, plus dynamic and structural shocks) yields a distribution of final states across the ensemble. In the final simulation period, 2.8% of pathways end in High emissions, which is the only final state classified as unexpected in this ensemble (empirical final-state share below 0.15 and strictly positive). Here “unexpected” means rare in this Monte Carlo ensemble under that threshold, not an externally validated tail probability under the central matrix. That share is non-negligible and supports discussion of tail risk and of

endpoints that a single point-estimate matrix would underrepresent once sampling and shocks are active.

The three-way distinction (robust across the three extremes regimes vs. sensitive to that joint specification vs. uncertainty-driven) can thus be read from the figure and from the unexpected-outcomes classification, bearing in mind that sensitivity to the joint extremes specification mixes CIM variant and structural scaling α . Under partial convergence, the disequilibrium run exhibits transient disequilibrium in 43% of period outcomes (initial period included), with a mean cumulative disequilibrium burden of 16.5. The cumulative disequilibrium burden is the sum over periods of the negative consistency margin when inconsistent (Eq. (6)), so this value indicates cumulative stress along pathways and confirms that the within-period cap prevents full convergence in many periods.

The four run types represent distinct transition scenarios, each reflecting a different underlying system, which can be qualitatively described as follows. In the disequilibrium scenario, an amplified crisis hits in 2030, throwing multiple factors out of alignment simultaneously, with elevated technology costs, slower permitting, and mounting land-use pressures the main sources of friction, before the transition recovers and reaches net zero in subsequent periods. In the smooth-transition world, policy commitment delivers grid expansion and renewables deployment, technology costs fall steadily, and land-use conflict remains the principal unresolved tension by 2050. In the turbulent-transition world, the drive towards decarbonisation is stronger than in the smooth-transition world, but greater system turbulence partially offsets this advantage, leaving the transition fractionally less advanced by 2050, particularly in how large a part hydrogen comes to play. In the constrained-transition world, the forces working against the transition are strengthened, with grid limitations and investment shortfalls bearing more heavily on outcomes, policy direction present but unable to deliver at the pace required, and one in four pathways remaining in medium-emissions states by 2050. The unexpected-outcomes scenario describes a world in which how strongly each aspect influences the others is itself uncertain and grows over time, generating a small but noticeable share of high-emissions pathways that a more certain assessment would not reveal. In the exogenous scenario, the transition faces shocks of various kinds throughout the period rather than in a single acute event, and nearly all pathways nonetheless reach net zero.

The same initial scenario, time grid, run count, baseline structural σ_{struct} , and convergence settings are held constant across run types, whilst selected shock settings are run-type-specific by design, so that differences in the plotted outcome distributions reflect the run type

(mechanism) under a controlled setup. The distinction drawn in the results between the four run types is qualitative. The numerical difference between two run types, however, is not interpretable as the quantitative effect of a single factor. Each run type differs in several respects at once, so the difference is a compound of all of them. In particular, the three extremes lines differ in both CIM and α simultaneously. Furthermore, many parameter choices (e.g. structural α values, event period and multiplier, dynamic shock parameters, unexpected threshold, storyline scaling factors for the transitions and regime CIMs) are conventional or illustrative. Results and any comparison across run types are conditional on those choices and are not asserted to be robust to re-parameterisation. Full specification and brief notes on parameter sources and sensitivity are given in [Appendix A](#).

6. Discussion and conclusions

This paper has introduced three extensions new to CIB pathway analysis, namely formalising disequilibrium along pathways, comparing pathways across regimes represented by alternative cross-impact matrices (the extremes illustration uses a joint regime-and-shock setup), and surfacing rare endpoints by sampling from an uncertain matrix specification. These are organised together with exogenous shocks into four run types in one framework. The types line up with different notions of tail or rare outcomes in the sense discussed in the results. Used together, they make both the spread of outcomes and the storyline behind each run easier to discuss, but they do not by themselves deliver a quantitative attribution of differences to a single lever. Unknown unknowns are addressed here only within the current descriptor set and influence structure, so that endpoints that a central matrix may underweight can still be examined. Exploring missing factors or relations (e.g. omitted descriptors or interactions) is outside the scope of this paper.⁸

This combination is not available in one place in the existing scenario and CIB pathway literature. It comprises four run types in a single framework, pathway-level disequilibrium diagnostics, and sampling from an uncertain matrix to expose rare trajectories or endpoints, plus the three-way reading of tails used in the illustration (robust across the joint extremes specification, sensitive to it, or driven by structural uncertainty). The implementation can be applied and extended to other cross-impact matrices and CIB applications.

For analysts, the main benefit is transparency of design. Instead of one band that mixes shocks, competing storylines, and doubt about how factors interact, the types differ in the mechanisms included in each run, making those distinctions explicit in interpretation. Stress-testing is explicit. Comparison across the three extremes lines and the unexpected-outcomes classification supports discussion of robustness and of tails without relying on a single central run. The disequilibrium and exogenous run types show response to an amplified period and to ongoing shocks. Where partial convergence applies, pathway-level diagnostics such as cumulative disequilibrium burden and the fraction of periods inconsistent summarise transient stress alongside ensemble outcomes. In the illustration, regime matrices and structural scaling move together. Separating storyline from shock intensity would mean holding one fixed whilst varying the other in a further design.

Much of the methodological literature on dynamic CIB has avoided an explicit acknowledgement of a structural tension, namely that dynamics are layered onto a representation whose natural unit is a

self-consistent scenario on a small, workshop-sized factor list. In that sense, dynamic CIB can force trajectory-like behaviour onto a scaffold that was never intended to encode a complete state space or governing equations for the real system. Adding uncertainty aspects, shocks, or regimes does not remove that constraint. Nor does the period grid resolve the temporal dimension (a gap common to all dynamic CIB approaches). Outputs are therefore best read as coherence and propagation of expert-structured judgements over those coordinates.

In complexity-science terms, the attractors identified by CIB are structural coherence conditions rather than emergent properties. That is, they are descriptor combinations that are mutually reinforcing under the expert-elicited interaction structure and logically entailed by the cross-impact matrix. Strong emergence, where system-level properties are not reducible to the lower-level interactions that produced them [Anderson \(1972\)](#) and [Chalmers \(2006\)](#), is outside what the framework claims to model. The attractor and basin-of-attraction analysis developed here draws on dynamical systems concepts. The disequilibrium metrics track a pathway's distance from the nearest attractor at each step, a quantity that maps to tipping-distance ideas from that literature (Section 3.2). A formal step in this direction is taken in [Ross et al. \(2026\)](#), which applies Green-Kubo Linear Response Theory to the CIB drift matrix as a framework for deriving macroscopic properties from microscopic dynamics. Fuller integration with complexity science would require an adaptive matrix whose interaction strengths update as the pathway evolves. That extension is not pursued here. Developing it would be a substantive step towards complexity-science integration of CIB. The complexity-science vocabulary used here (attractors, basins of attraction, tipping distances) is accordingly best read as diagnostic language for characterising pathway outputs on the elicited scaffold, not as a claim that CIB models emergent system dynamics. The attractor structure is a property of the cross-impact matrix, not of the underlying socio-technical system.

Subject to that reading, the disequilibrium diagnostics and run-type contrasts described in Sections 3.2 and 5 are the appropriate objects for reporting. They support transparent discussion of mechanism and tails on the elicited scaffold, and they ground claims about transient stress in the formal state space and update rules rather than in a complete dynamical-systems representation. They do not substitute for expanding the descriptor set or for model linkage where important mechanisms are absent from the formal coordinates.

Declaration of competing interest

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

Acknowledgements

Special thanks go to Fiona for insightful discussions about probability concepts, and to Alan Ross for contributions to improving the methodology. This paper is funded by the Helmholtz Association under the programme “Energy System Design”. Open Access is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) (491111487). The author utilised LLM's such as ChatGPT, as language and code assistants during the preparation of this work. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Appendix A

This appendix specifies the parameter and run-type choices used in the simulation. Notation follows Section 3, with symbols and assumed values given in Table S1 in the Supplementary Material. As in Section 4, the admissible CIM score range $[c_{\min}, c_{\max}]$ (the clip interval in Eq. (8)) is $[-3, +3]$, and sampled or perturbed cell values are clipped to this

⁸ That dimension could be pursued by eliciting or discovering additional descriptors (e.g. via horizon scanning, stakeholder workshops, or scenario discovery on quantitative models), by testing alternative or sparse influence structures to see which relations materially change pathway outcomes, or by extending uncertain-CIM sampling so that candidate descriptors or links are included in some runs and excluded in others. Such extensions are left for future work. Nevertheless, given the contribution of [Ross and Ross \(2026\)](#), the cost of generating additional expert-elicited data has decreased, making such potential extensions feasible.

Table A.1
Simulation run types and main options.

Option	Disequilibrium	Extremes	Unexpected	Exogenous
CIM sampling	No	No	Yes ^a	No
Regimes	No	Yes ^b	No	No
Structural shocks	Yes ^c	Yes ^d	Yes ^c	Yes ^c
Dynamic shocks	No	No	Yes	Yes
Partial convergence	Yes	Yes	Yes	Yes

^a One uncertain specification per run, with the active CIM re-sampled each period using $\lambda(t)$.

^b Three extremes ensembles, no regime switching, each pairing a regime-specific CIM with regime-specific structural scaling.

^c Structural shocks use multiplicative-magnitude scaling with baseline $\alpha = 0.3$ (with one amplified event period in the disequilibrium run type).

^d For Extremes, structural scaling is regime-specific.

range. The exact values and code are available from the repository referenced in the Data statement. Where a value is derived from the CIM data (e.g. σ_{struct} from confidence codes) or from the CIM structure (e.g. K_{max} , initial scenario), this is noted in the relevant paragraph that follows, whilst other numerical choices are conventional or illustrative. The caveat in the Results (Section 5) applies. Results and comparisons are conditional on these choices and are not asserted to be robust to re-parameterisation.

Table A.1 summarises the four run types (CIM sampling, regimes, structural and dynamic shocks, and partial convergence). The rows should be read alongside the parameter descriptions in this appendix and with Table S1 (in the Supplementary Material) for symbols and numerical values.

The illustrative empirical application uses CIM data from Ross and Ross (2026). The matrix C (Section 3.1, Eq. (1)) has 15 descriptors with three ordered states per descriptor, covering policy, technology, infrastructure, and demand-side dimensions, and score and confidence columns (point estimate $\bar{c}$ and confidence code c mapping to scale $\sigma(c)$ in Eq. (8)) are used. The baseline structural shock strength σ_{struct} (the standard-deviation scale of the random perturbation applied to the CIM in structural shocks, Section 3.5) is derived from the CIM confidence codes using the median of the confidence-derived standard deviations. The timeline descriptor used for the ensemble share of each of its states over time across pathway runs and for defining unexpected final states of the pathway $(z(t))_{t \in T}$ (i.e. the descriptor whose state at the last period is used for unexpected-outcome classification) is Decarbonisation Outcome. Each cell supplies an impact score on the scale $[-3, +3]$ and a confidence code mapping to a scale $\sigma(c)$, so that both the baseline structural shock strength and the uncertain-CIM sampling (Eq. (8)) reflect the experts' stated uncertainty.

Pathways are defined over a discrete set of periods and require a starting scenario. The time grid T (Section 3.1) consists of six periods $t_1, \dots, t_N$ at five-year intervals (2025, 2030, 2035, 2040, 2045, 2050). The initial scenario (the starting scenario for the pathway $(z(t))_{t \in T}$, i.e. the scenario from which succession is applied in the first period) is defined as the middle ordinal state of each descriptor under the predefined descriptor-state ordering used in the implementation (Ross and Ross, 2026).

So that transient disequilibrium and path diversity can be observed along pathways, within-period succession is capped rather than run to full convergence (Section 3.1). The within-period iteration cap K_{max} (the maximum number of succession steps per period before the period outcome is taken as $z^{(K_{\text{max}})}$) is set to $\max(1, \lfloor n_{\text{desc}}/2 \rfloor)$, where $\lfloor x \rfloor$ denotes the floor of x (greatest integer not exceeding x). When equilibrium relaxation is applied, a separate iteration limit of 1000 steps is used (relaxation towards a scenario satisfying Eq. (2)). Local succession is applied, updating one descriptor per step within each period (the most inconsistent, i.e. the one with the largest gap between its current and maximum impact score).

Given this structure, under attractor analysis, the baseline CIM exhibits a single dominant fixed-point attractor (the pro-transition state, Net-zero aligned with Strong grid and High investment) with a large basin share (the fraction of scenarios in that attractor's basin of attraction, i.e. that converge to that attractor under succession). Under full within-period convergence, most periods would therefore end at that scenario and pathway diversity would be limited. The within-period cap therefore ensures that many periods stop before reaching an attractor.

Structural shocks (Section 3.5, random perturbations applied to the CIM) use multiplicative-magnitude scaling. The perturbation is scaled by a factor α . Under this mode, the effective shock magnitude is the draw multiplied by $1 + \alpha|\bar{c}|/c_{\text{max}}$, where $\bar{c}$ is the cell point estimate and $c_{\text{max}} = 3$ normalises to the CIM score range. The baseline structural scaling α (the multiplier applied to the shock magnitude, Table S1) is 0.3 for the disequilibrium, unexpected-outcomes, and exogenous-shocks run types. One designated period t (2030) is given an amplified structural σ_{struct} (the scale of the random perturbation, multiplier 4.5 on σ_{struct}) so that a transient event is visible in the disequilibrium ensemble. For the extremes run type, the structural scaling α takes the values 0.20 (smooth-transition regime), 0.50 (turbulent-transition regime), and 0.60 (constrained-transition regime), chosen so that the three regime ensembles differ clearly in pathway spread. No descriptor-specific or state-specific scaling is applied.

Dynamic shocks (Section 3.5, random perturbations applied to the impact scores $\theta_{j,i}(z)$ during succession) are used in the unexpected-outcomes and exogenous-shocks run types. They are applied and discussed in Ross and Ross (2026). In this calibration, dynamic-shock settings are run-type-specific, with $\tau = 0.36$ and $\rho = 0.65$ for unexpected outcomes, and $\tau = 0.50$ and $\rho = 0.20$ for exogenous shocks (Table S1). The innovation distribution is normal.

In the unexpected-outcomes run type (Section 3.4), one uncertain matrix (specification) is used per run. The active CIM is re-sampled each period from that specification with multiplicative scale $\lambda(t)$, using the implemented linear schedule $\lambda(t_i) = 1 + i$ (equivalently: base 1.0 and additive increment 1.0 per period), so that uncertainty widens over time. Each cell is drawn as in Eq. (8). An outcome is classified as unexpected when the empirical share of pathways ending in that final state (of the timeline descriptor) is below 0.15 and strictly positive (the final-state share threshold in Table S1), a conventional cut-off for "rare" in this application). A near-miss threshold $\epsilon = 0.25$, a conventional cut-off for reporting outcomes just above that range, is used where applicable.

For the extremes run type (Section 3.3), three regimes r define three separate pathway ensembles (Table S1). There is no regime switching, and each run stays in one regime for the whole pathway, so $r(t) = r_0$ for all $t \in T$. In this empirical illustration, the extremes runs use fixed regime-specific CIMs (no sampling) together with regime-specific structural shock scaling, which is consistent with common storyline-style implementations in the existing scenario literature. This configuration is illustrative rather than exhaustive. The technical framework, however, treats regime choice, structural-shock scaling, and uncertain-CIM sampling as separable optional layers, so implementations can instead hold one layer fixed whilst varying another (e.g. fixed scaling across regimes, fixed regime CIM across scaling levels), or allow regime switching over time. This layer-separable implementation extends common storyline-style implementations in the current literature, even though the specific empirical configuration shown here is intentionally consistent with them.

In the present setup, each regime's matrix is supplied as a fixed $M_{\text{base}}^{(r)}$ with no further runtime composition of modifiers ($L = 0$ in Eq. (7)). The turbulent-transition and constrained-transition matrices are deterministic multiplicative scalings of the baseline cells (Table S1), not new elicitations. The turbulent-transition matrix strengthens impacts involving policy stringency, electrification pace, and renewables deployment (as source or target). The constrained-transition matrix

strengthens involvement of the Decarbonisation Outcome descriptor, weakens impacts that target grid development and investment availability, and applies an additional down-weighting to impacts towards selected pro-transition target states so that the constrained-transition attractor can dominate for illustration. A separate robustness evaluation in the implementation operates on the uncertain matrix specification and is not the same object as these deterministic regime CIMs, and only the pathway figures in the main text use this deterministic extremes setup.

Parameters derived from the CIM or its structure include σ_{struct} (median of confidence-derived scales from the CIM), the score range $[-3, +3]$ and clip interval, the descriptor set and timeline descriptor, and the initial scenario (middle ordinal state per descriptor under the predefined ordering). The within-period cap $K_{\text{max}} = \max(1, \lfloor n_{\text{desc}}/2 \rfloor)$ is a design choice to allow transient disequilibrium, since larger K_{max} would allow more convergence and typically reduce pathway diversity. The remaining numerical choices (e.g. baseline α , event period and multiplier, run-type-specific τ and ρ , $\lambda(t)$ growth, unexpected-outcome threshold, run count) are conventional or illustrative.

No full sensitivity analysis is conducted. The focus of the paper is on pathway mechanisms (disequilibrium, regime-based extremes, uncertain influence structure, and exogenous disturbances) and on illustrating how each can be implemented and read from ensemble outputs, rather than on calibrating a single “best” parameter set or on quantifying robustness of numerical magnitudes across the full input space. Nevertheless, the qualitative implications of changes in the main illustrative parameters are straightforward. A larger structural α or a higher event multiplier would increase pathway spread and transient stress, a higher final-state share threshold would classify more states as unexpected, stronger $\lambda(t)$ growth would widen uncertainty over the horizon, and different τ or ρ would change the persistence and magnitude of dynamic shocks. As noted in Section 5, results and comparisons are conditional on these choices and are not asserted to be robust to re-parameterisation.

A fuller treatment of parameters could nonetheless connect selected numerical settings to evidence, when the analyst wishes to tighten plausibility or to conform to a given stress-test protocol. This might be done by using structured expert judgement for shock severity and for cut-offs that operationalise “rare” outcomes, by adopting stress-test templates from policy or risk practice, or by checking qualitative consistency with stylised facts or with reference projections from other models. None of this claims that the CIB pathway specification is identified from a likelihood. Structural shock intensity, dynamic shock persistence, and widening matrix uncertainty can trade off in producing similar ensemble dispersion, so reporting ranges or scenario suites can dominate the search for a single “best” parameter vector. Such extensions are left for subsequent work. This paper focuses on the mechanism taxonomy and on a transparent baseline specification.

A focused robustness protocol is recommended to accompany the illustrative baseline. First, convergence-policy sensitivity could be checked by re-running all run types under strict within-period convergence (no partial cap) and comparing endpoint shares, cumulative disequilibrium burden, and timeline trajectories against the capped setting, where materially different outcomes indicate convergence-policy dependence. Second, the unexpected-outcome classification could be stress-tested by a threshold sweep (e.g. 0.05, 0.10, 0.15, 0.20) and by Monte Carlo size checks (larger run counts) to ensure that rare-state labels are not artefacts of sampling noise. Third, to improve attribution, a factorial design could separate uncertainty channels by varying one layer at a time (regime matrix, structural scaling α , uncertain-CIM sampling, dynamic shocks). These checks do not replace the mechanism taxonomy proposed here, but make its empirical interpretation more robust and transparent.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.techfore.2026.124893>.

Data availability

Code and data supporting the simulations and figures in this paper are available at <https://doi.org/10.5281/zenodo.19364824>.

References

- Anderson, P.W., 1972. More is different. *Science* 177, 393–396. <http://dx.doi.org/10.1126/science.177.4047.393>.
- Bai, X., van der Leeuw, S., O'Brien, K., Berkhout, F., Biermann, F., Brondizio, E.S., Cudennec, C., Dearing, J., Duraipah, A., Glaser, M., Revkin, A., Steffen, W., Syvitski, J., 2016. Plausible and desirable futures in the anthropocene: A new research agenda. *Glob. Environ. Change* 39, 351–362. <http://dx.doi.org/10.1016/j.gloenvcha.2015.09.017>.
- Campfens, J.K., Duygan, M., Binder, C.R., 2025. Initiating social tipping dynamics in energy transitions: A novel analytical approach for exploring feedback loops and intervention points. *Environ. Innov. Soc. Transit.* 56, 100973. <http://dx.doi.org/10.1016/j.eist.2025.100973>.
- Chalmers, David J., 2006. Strong and weak emergence. In: Clayton, Philip, Davies, Paul (Eds.), *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*. Oxford University Press, <http://dx.doi.org/10.1093/acprof:oso/9780199544318.003.0011>.
- Eriksson, E.A., Hallding, K., Skånberg, K., 2022. Ensuring representativity of scenario sets: The importance of exploring unknown unknowns. *Futures* 139, 102939. <http://dx.doi.org/10.1016/j.futures.2022.102939>.
- Frey, U.J., Cao, K.K., Sasanpour, S., Buschmann, J., Breuer, T., 2025. The benefits of exploring a large scenario space for future energy systems. *Nat. Commun.* 17, 873. <http://dx.doi.org/10.1038/s41467-025-67593-9>.
- Gambhir, A., Lempert, R., 2023. From least cost to least risk: Producing climate change mitigation plans that are resilient to multiple risks. *Front. Clim.* 5, 1149309. <http://dx.doi.org/10.3389/fclim.2023.1149309>.
- Guivarch, C., Le Gallic, T., Bauer, N., Fragkos, P., Huppmann, D., Jaxa-Rozen, M., Keppo, I., Kriegler, E., Krisztin, T., Marangoni, G., Pye, S., Riahi, K., Schaeffer, R., Tavoni, M., Trutnevyte, E., van Vuuren, D., Wagner, F., 2022. Using large ensembles of climate change mitigation scenarios for robust insights. *Nat. Clim. Change* 12, 428–435. <http://dx.doi.org/10.1038/s41558-022-01349-x>.
- Guivarch, C., Lempert, R., Trutnevyte, E., 2017. Scenario techniques for energy and environmental research: An overview of recent developments to broaden the capacity to deal with complexity and uncertainty. *Environ. Model. Softw.* 97, 201–210. <http://dx.doi.org/10.1016/j.envsoft.2017.07.017>.
- Hamming, R.W., 1950. Error detecting and error correcting codes. *Bell Syst. Tech. J.* 29, 147–160. <http://dx.doi.org/10.1002/j.1538-7305.1950.tb00463.x>.
- Kemp-Benedict, E., Carlsen, H., Kartha, S., 2019. Large-scale scenarios as ‘boundary conditions’: A cross-impact balance simulated annealing (CIBSA) approach. *Technol. Forecast. Soc. Change* 143, 55–63. <http://dx.doi.org/10.1016/j.techfore.2019.03.006>.
- Kwakkel, J.H., Pruyt, E., 2013. Exploratory modeling and analysis, an approach for model-based foresight under deep uncertainty. *Technol. Forecast. Soc. Change* 80, 419–431. <http://dx.doi.org/10.1016/j.techfore.2012.10.005>, future-Oriented Technology Analysis.
- Lempert, R.J., Groves, D.G., Popper, S.W., Bankes, S.C., 2006. A general, analytic method for generating robust strategies and narrative scenarios. *Manag. Sci.* 52, 514–528. <http://dx.doi.org/10.1287/mnsc.1050.0472>.
- Lempert, Robert J., Popper, Steven W., Bankes, Steven C., 2003. *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis. Technical Report*, RAND Corporation.
- Lin, N., Emanuel, K., 2016. Grey swan tropical cyclones. *Nat. Clim. Change* 6, 106–111. <http://dx.doi.org/10.1038/nclimate2777>.
- McCollum, D.L., Gambhir, A., Rogelj, J., Wilson, C., 2020. Energy modellers should explore extremes more systematically in scenarios. *Nat. Energy* 5, 104–107. <http://dx.doi.org/10.1038/s41560-020-0555-3>.
- Nakicenovic, N., Swart, R. (Eds.), 2000. *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, URL <https://www.ipcc.ch/report/emissions-scenarios/>.
- O'Neill, B.C., Kriegler, E., Ebi, K.L., Kemp-Benedict, E., Riahi, K., Rothman, D.S., van Ruijven, B.J., van Vuuren, D.P., Birkmann, J., Kok, K., Levy, M., Solecki, W., 2017. The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Glob. Environ. Change* 42, 169–180. <http://dx.doi.org/10.1016/j.gloenvcha.2015.01.004>.
- O'Neill, B.C., Kriegler, E., Riahi, K., Ebi, K.L., Hallegatte, S., Carter, T.R., Mathur, R., van Vuuren, D.P., 2014. A new scenario framework for climate change research: the concept of shared socioeconomic pathways. *Clim. Change* 122, 387–400. <http://dx.doi.org/10.1007/s10584-013-0905-2>.

- Riahi, K., van Vuuren, D.P., Kriegler, E., Edmonds, J., O'Neill, B.C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuasmasma, J.C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humenöder, F., Da Silva, L.A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Krey, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J.C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., Tavoni, M., 2017. The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Glob. Environ. Change* 42, 153–168. <http://dx.doi.org/10.1016/j.gloenvcha.2016.05.009>.
- Roponen, J., Salo, A., 2024. A probabilistic cross-impact methodology for explorative scenario analysis. *Futures Foresight Sci.* 6, e165. <http://dx.doi.org/10.1002/ffo2.165>.
- Rosen, R.A., 2025. Problems with creating useful scientifically valid futures scenarios. *Futures* 167, 103566. <http://dx.doi.org/10.1016/j.futures.2025.103566>.
- Ross, A.G., 2026. *PyCIB Cross-Impact Balance (CIB) analysis package*.
- Ross, A.G., Gershenson, J., Kleefeld, A., 2026. Beyond consistent scenarios: Deriving indirect influence, transition resistance, and adjustment dynamics. <http://dx.doi.org/10.48550/arXiv.2607.12414>.
- Ross, A.G., Ross, A.M., 2026. AI-Simulated expert panels for Socio-Technical scenarios and decision guidance. <http://dx.doi.org/10.48550/arXiv.2603.29470>.
- Salo, A., Tosoni, E., Roponen, J., Bunn, D.W., 2022. Using cross-impact analysis for probabilistic risk assessment. *Futures Foresight Sci.* 4, e2103. <http://dx.doi.org/10.1002/ffo2.103>.
- Schweizer, V.J., 2020. Reflections on cross-impact balances, a systematic method constructing global socio-technical scenarios for climate change research. *Clim. Change* 162, 1705–1722. <http://dx.doi.org/10.1007/s10584-019-02615-2>.
- Schweizer, V.J., Jamieson-Lane, A.D., Cai, H., Lehner, S., Smerlak, M., 2023. Pathways for socio-economic system transitions expressed as a Markov chain. *PLoS One* 18, e0288928. <http://dx.doi.org/10.1371/journal.pone.0288928>.
- Schweizer, V.J., Kriegler, E., 2012. Improving environmental change research with systematic techniques for qualitative scenarios. *Environ. Res. Lett.* 7, 044011. <http://dx.doi.org/10.1088/1748-9326/7/4/044011>.
- Taleb, N.N., 2007. *The Black Swan: The Impact of the Highly Improbable*. Random House, New York.
- Vögele, S., Poganietz, W.R., Mayer, P., 2019. How to deal with non-linear pathways towards energy futures: Concept and application of the cross-impact balance analysis. *TATuP-Z. Tech. Theor. Prax.* 28, 20–26. <http://dx.doi.org/10.14512/tatup.28.3.20>.
- Walker, W., Harremoës, P., Rotmans, J., van der Sluijs, J., van Asselt, M., Janssen, P., von Krauss, M.K., 2003. Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integr. Assess.* 4, 5–17. <http://dx.doi.org/10.1016/j.techfore.2005.06.005>.
- Weimer-Jehle, W., 2006. Cross-impact balances: A system-theoretical approach to cross-impact analysis. *Technol. Forecast. Soc. Change* 73, 334–361. <http://dx.doi.org/10.1016/j.techfore.2005.06.005>.
- Weimer-Jehle, W., Vögele, S., Hauser, W., Kosow, H., Poganietz, W.R., Prehofer, S., 2020. Socio-technical energy scenarios: state-of-the-art and CIB-based approaches. *Clim. Change* 162, 1723–1741. <http://dx.doi.org/10.1007/s10584-020-02680-y>.
- Weitzman, M.L., 2009. On modeling and interpreting the economics of catastrophic climate change. *Rev. Econ. Stat.* 91, 1–19. <http://dx.doi.org/10.1162/rest.91.1.1>.

Andrew G. Ross is a senior scientist at the Forschungszentrum Jülich, specialising in (macro)economic assessments of various energy- and non-energy related policies, shocks, and scenarios. He holds a PhD in Economics from the University of Strathclyde.