Supplementary Information

Formal Definition of Criticality

In the following, we propose more formal definitions of key terms described in Box 1, building on those provided in Winkelmann et al. (2022)Winkelmann et al. (2022), that can be useful when operationalising the proposed concepts in modelling or data analysis studies.

Definition: The *criticality* $C = C(x(t); \mathbf{p}, t, \xi)$ of a complex adaptive social system in a state $x(t)$ with parameter values described by a vector $\mathbf{p}$ under a perturbation regime ξ is the probability rate at which the system undergoes a path bifurcation when the social system is currently in $(x, \mathbf{p})$ and is perturbed by ξ.

Definition: A social system undergoes a *path bifurcation* when being perturbed from a state x and parameter values $\mathbf{p}$ to a state x' and parameter values $\mathbf{p}'$, and the trajectories starting at $(x, \mathbf{p})$ and $(x', \mathbf{p}')$ are qualitatively different.

Definition: The (deterministic or stochastic) trajectories starting at $(x, \mathbf{p})$ and $(x', \mathbf{p}')$ are said to be *qualitatively different* if there is no curve in state-parameter space from $(x, \mathbf{p})$ to $(x', \mathbf{p}')$ such that the trajectory starting at $(x'', \mathbf{p}'')$ changes continuously as $(x'', \mathbf{p}'')$ moves along that curve. To assess continuity, one might use the topology of pointwise convergence on the space of deterministic trajectories, or, for stochastic systems, the total variation distance metric on the space of stochastic trajectories.

Detecting criticality

In many systems, increased criticality is associated with characteristic quantitative changes in the system's functioning, detectable as Early Warning Signals (EWS) for impending critical transitions. A well-known example is the Critical Slowing Down (CSD) phenomenon: a system in a state of degrading resilience (increasing criticality) returns to equilibrium more sluggishly after a disturbance Scheffer (2009); Lenton (2020). Measures such as increasing time-series autocorrelation, variance or flickering can be used to identify this phenomenon and indicate an approaching tipping threshold. EWS have been found in the context of natural, environmental and ecological regime changes such as lake disturbances, forest transitions, and climate subsystems Scheffer (2009), and can go beyond CSD indicators Biggs et al. (2009). Directly transferring EWS analytical tools to the study of social systems can face difficulties, such as in quantifying uncertainties Boettiger and Hastings (2012); Dakos et al. (2015), and unreliability when applied to non-equilibrium processes Dablander et al. (2022). The highly complex, adaptive, multi-layered and agency-rich nature of social processes (Winkelmann et al., 2022) can magnify these drawbacks. Furthermore, many common forms of historical social data (e.g., surveys, voting behaviours) lack the granularity necessary for common EWS methods.

When more specific mechanisms for societal tipping are known or suspected, criticality can be assessed by simulation models of the processes in question. For example, agent-based models (ABMs) and adaptive network models can be used to analyse levels of criticality depending on varying properties of actors and social structures. ABMs allow for the simultaneous modelling of both contagion processes and network re-organizations, as well as agent heterogeneity and temporal dynamics Macy and Willer (2002). Calibrating models to real-world conditions remains an important research challenge, which requires interdisciplinary cooperation to validate model inferences Schlüter et al. (2019). Another research challenge is the development of network-analytical tools to investigate the criticality of real-world networks and nodes. Interdisciplinary knowledge sharing and dialogue will continue to be critical to increase the contribution of simulation modelling to the "crisis discipline" Bak-Coleman et al. (2021) of sustainability transformation science.

Furthermore, understanding how criticality can be shaped, and how likely societal systems are to undergo tipping processes can also be understood from identifying historical case studies of previous, potential tipping processes Hodbod et al. (2024). Societal tipping literature has often conceptually relied upon contagion- or diffusion-based mechanisms towards rapidly decarbonizing economies or increasing technological adoption Olsson and Moore (2024a). Yet, such mechanisms do not always reflect the more nuanced, non-linear approaches of transformation literatures – explicitly accounting for the complex nature of social systems which rarely exhibit deterministic or linear change Norström et al. (2022). For example, such complexity has been explored via identifying the historical and institutional factors leading to peace building and transformative justice, conditions influencing the success of feed-in tariffs and the 'Energiewende' in Germany Lipp (2007). There are also a great number of case-studies and empirical work in resilience and sustainability science, ranging from local, regional, to global governance systems, that have unravelled the social dynamics behind tipping points and phase shifts (often referred to as regime shifts). Here, the interplay of four interacting features of complex adaptive social-ecological systems – critical agency, social networks, bridging organisations, and institutions – has been identified as essential for active transformations towards biosphere stewardship Folke et al. (2005); Herrfahrdt-Pähle et al. (2020).