

Constrained basin stability for studying transient phenomena in dynamical systemsAdrian van Kan,^{1,2,*} Jannes Jegminat,^{1,†} Jonathan F. Donges,^{3,4} and Jürgen Kurths^{3,5,6,7}¹*Department of Physics and Astronomy, University of Heidelberg, Im Neuenheimer Feld 226, D-69120 Heidelberg, Germany*²*Department of Physics, Imperial College London, Prince Consort Rd, London SW7 2BB, United Kingdom*³*Potsdam Institute for Climate Impact Research, P.O. Box 601203, D-14412 Potsdam, Germany*⁴*Stockholm Resilience Centre, Stockholm University, Kräftriket 2B, 114 19 Stockholm, Sweden*⁵*Department of Physics, Humboldt University Berlin, Newtonstr. 15, D-12489 Berlin, Germany*⁶*Institute for Complex Systems and Mathematical Biology, University of Aberdeen, Aberdeen AB24 3FX, United Kingdom*⁷*Department of Control Theory, Nizhny Novgorod State University, Gagarin Avenue 23, 606950 Nizhny Novgorod, Russia*

(Received 13 August 2015; revised manuscript received 9 January 2016; published 13 April 2016)

Transient dynamics are of large interest in many areas of science. Here, a generalization of basin stability (BS) is presented: constrained basin stability (CBS) that is sensitive to various different types of transients arising from finite size perturbations. CBS is applied to the paradigmatic Lorenz system for uncovering nonlinear precursory phenomena of a boundary crisis bifurcation. Further, CBS is used in a model of the Earth's carbon cycle as a return time-dependent stability measure of the system's global attractor. Both case studies illustrate how CBS's sensitivity to transients complements BS in its function as an early warning signal and as a stability measure. CBS is broadly applicable in systems where transients matter, from physics and engineering to sustainability science. Thus CBS complements stability analysis with BS as well as classical linear stability analysis and will be a useful tool for many applications.

DOI: [10.1103/PhysRevE.93.042205](https://doi.org/10.1103/PhysRevE.93.042205)**I. INTRODUCTION**

Many fields of science analyze dissipative dynamical systems in terms of their attractors. Thus it is an important challenge to quantify the stability of attractors with respect to a given perturbation. The most popular method is linear stability analysis, which considers infinitesimal perturbations. Menck *et al.* [1] suggest to complement this linear measure with basin stability (BS), which accounts for finite and even large perturbations. The application of BS to power grids has yielded novel mitigation strategies against superoutages [2,3]. BS is computed by estimating the volume of an attractor's basin. Therefore, it is not sensitive to different forms of transient dynamics. However, transient phenomena in complex systems are of large interest in many areas of science, such as climatic and, more generally, global change in Earth system science [4], epileptic seizures in neuroscience [5], ecosystem transitions in ecology [6], as well as in the previously mentioned study of super outages in power grids [2,3]. For example, in the case of climate change [4] and the great acceleration [7] as transient phenomena in the global social-environmental system [8,9], major efforts are invested into studying the maximum global mean temperature and its timing along the trajectory due to anthropogenic greenhouse gas emissions. Moreover, the model- and data-driven analysis of transient global change trajectories underlies many recently proposed frameworks for sustainable development such as tolerable environment and development windows [10], planetary boundaries [11,12], and the safe and just operating space for humanity [13].

Making BS sensitive to transients, we generalize it to a family of stability measures termed constrained basin stabilities (CBSs). As opposed to BS, CBS is not computed

from the entire basin of an attractor but only from a subset of the basin. The subset is defined by a generic constraint imposed on the transients. Thus CBS is sensitive to transients while maintaining the intuitiveness and simplicity of BS. To illustrate how CBS complements BS, we choose two specific constraints on transients, one based on the confinement of transient trajectories to certain regions in phase space and one based on transient duration, and apply them to the Lorenz system and a global carbon cycle model, respectively. In the former example, CBS anticipates a boundary crisis bifurcation. In the latter, we show that CBS represents a more intuitive measure for stability than BS because CBS reflects not only that perturbation-induced transients return to the attractor but also that they do so within a desirable time interval.

This paper is structured as follows. In Sec. II we introduce CBS and discuss some of its properties used in the further analysis. In Sec. III we present two examples of CBS analysis in dynamical systems: the paradigmatic Lorenz [14] model and a global carbon cycle model proposed by Anderies *et al.* [15]. Then, in Sec. IV we discuss the relevance of CBS and how it differs from established stability concepts. The paper concludes with closing remarks.

II. METHODS

Let the (not necessarily analytic) vector valued function f represent an autonomous potentially multistable dynamical system $\dot{\mathbf{x}} = f(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^n$, and let $\phi^t(\mathbf{x}_0)$ be the system state at time t on a trajectory starting at \mathbf{x}_0 at $t = 0$. The basin stability $S_B(A)$ of an attractor A of this system quantifies the probability that after a finite size perturbation trajectories return to A [1]. Perturbations within the attractor's basin $\mathcal{B}(A)$ return but the remaining ones fall into a different attractor. The probability distribution of perturbing a trajectory on the attractor to the state \mathbf{x} is given by $\rho(\mathbf{x})$. Thereby, BS is formally

*van_kan@stud.uni-heidelberg.de

†jegminat@iup.uni-heidelberg.de

defined as

$$S_B(A) = \int_{\Gamma} dx^n \rho(\mathbf{x}) \mathbf{1}_{B(A)}(\mathbf{x}), \quad (1)$$

where Γ denotes the state or phase space of the dynamical system and the indicator function is

$$\mathbf{1}_{B(A)}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in B(A), \\ 0 & \text{else.} \end{cases} \quad (2)$$

BS can be computed quickly once an attractor's basin is given: it equals the mass of ρ that is supported by the basin. However, in practice, the basin of attraction is usually not known and needs to be determined first. For this purpose, initial conditions are sampled according to the perturbation density ρ and then integrated until they reach an attractor. Thus, in computing BS only the long-term limit of the trajectories is used to determine if an initial condition lies in the basin of attraction. Therefore, by construction, BS does not depend on transient motion.

To generalize BS, here we propose instead to use the properties of transients to define a class of stability measures that we term *constrained basin stabilities* (CBSs). Calculating CBS requires a computational effort similar to that needed for BS, but CBS reveals additional information about the system that is encoded in the transient trajectories. We define a transient as the set of points belonging to the part of the trajectory between the initial condition $\mathbf{x}(0) = \mathbf{x}_0$ and reaching the attractor A ,

$$T(\mathbf{x}_0) = \{\phi^t(\mathbf{x}_0) \in \Gamma \setminus A \mid t \geq 0\}. \quad (3)$$

The fact that we define the attractor not to be part of the transient makes a difference for example in the case of trajectories induced by nonsmooth flows where an attractor may be reached within finite time. The idea of CBS is that a region in phase space is identified by some constraint on the transients starting from a subset of phase space

$$C = \{\mathbf{x} \in \Gamma \setminus A \mid \text{the transient from } \mathbf{x} \text{ satisfies a constraint}\}. \quad (4)$$

In other words, the transients starting from this *conditioned set* C satisfy the given constraint. For instance, we can choose C to be the set of states \mathbf{x} the transients starting from which exhibit monotonicity in the x_1 component. This is equivalent to demanding that the projection of a transient's velocity onto the basis vector \mathbf{e}_1 in x_1 direction is nonvanishing. Thus the conditioned set is $C_{mon} = \{\mathbf{x} \in \Gamma \setminus A \mid f(\phi^t(\mathbf{x})) \cdot \mathbf{e}_1 \neq 0 \forall t > 0\}$. If x_1 represents a population, the set C_{mon} is the set of initial conditions which do not lead to a population overshoot [16].

Incorporating an arbitrary constraint (not necessarily monotonicity) as an additional factor in Eq. (1), we formally define CBS as

$$S_B^C(A) = \int_{\Gamma} dx^n \rho(\mathbf{x}) \mathbf{1}_C(\mathbf{x}) \mathbf{1}_{B(A)}(\mathbf{x}). \quad (5)$$

The product of the two indicator functions checks whether a perturbation is inside of the attractor's basin and at the same time inside the region transients originating from which satisfy the prescribed constraint. Figure 1 illustrates the regions in a schematic two-dimensional phase space that are relevant for computing BS and CBS for a fixed point. Three useful

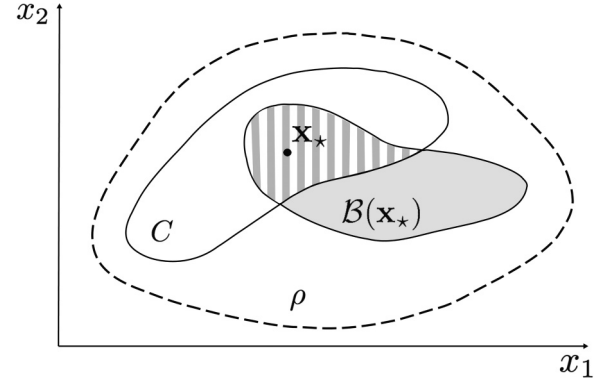


FIG. 1. Subspaces of a schematic two-dimensional phase space containing a fixed point \mathbf{x}_* : perturbations are sampled from the domain of the perturbation probability density ρ (area within dashed line). Their transients return to the fixed point when sampled from the basin of attraction $B(\mathbf{x}_*)$ (gray area). The set C (white within solid line) is the set of initial conditions which lead to transients that satisfy a given constraint, e.g., on the x_1 component. While BS is computed as the fraction of perturbations within the basin of attraction, CBS is the fraction of perturbations within the intersection $C \cap B(\mathbf{x}_*)$ (striped area). Thus CBS reflects the stability with respect to perturbations whose transients fulfill a given constraint.

properties of CBS follow directly from its definition. First, since $\mathbf{1}_C(\mathbf{x}) \leq 1$,

$$S_B(A) \geq S_B^C(A). \quad (6)$$

Secondly, let $\{C_i\}_{i \in I}$ be a partition of Γ , then

$$\sum_{i \in I} S_B^{C_i}(A) = S_B(A). \quad (7)$$

Thirdly, if $C_1 \subset C_2$ then

$$S_B^{C_1}(A) \leq S_B^{C_2}(A). \quad (8)$$

The novelty of CBSs is that they integrate information about the transients into the asymptotic framework of BS. This information is encoded in a set C of states the transients originating from which satisfy a given requirement, such as monotonicity in the previous example. We suggest classifying these requirements as static, dynamic, and integrated, depending on how much information is necessary to find out which desirable region of phase space corresponds to them as follows. (i) Static requirements *a priori* define a phase space region $\Gamma' \subset \Gamma$ that must not be entered by the transient. No further knowledge of the system or its dynamics is required. Examples are planetary boundaries in Earth system dynamics [11,12], minimum or maximum operating temperatures of a device or the evaluation of external functions (not f) of the system, e.g., the performance of a second system that depends on the system state \mathbf{x} . (ii) Dynamic requirements depend on velocity, thus more knowledge about the system is required: the dynamics, i.e., f , must be known. Using this knowledge, a region similar to Γ' is defined. Consider, for example, a roller coaster that must not exceed a certain velocity or acceleration or the requirement of monotonicity in economic output to exclude the burst of market bubbles. (iii) Integrated conditions depend not only on the current state of the transient but also on its past,

i.e., they operate on an infinite dimensional space and memory effects are possible. Despite this complexity, testing integrated conditions is often easy in practice as can be seen from the following examples: imposing a limited number of opinion changes of a political party, thresholding the time needed to reach an attractor, integrated damage in a climate model or imposing a minimum average power output of a wind farm per time interval. Note that each of the constraints implies a binary decision: CBS identifies a qualitative property in a transient.

In order to implement CBS [Eq. (5)] numerically, we need to discretize it. For simplicity of presentation, we choose the attractor to be a fixed point, $A = \{\mathbf{x}_*\}$, and consider a uniform distribution ρ of N initial conditions drawn from some subset of phase space approximated by a set of sampling points \mathbf{x}_i , $i \in \{1, \dots, N\}$ drawn at random from the phase space volume in question. This results in

$$S_B^C(\mathbf{x}_*, \varepsilon) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_C(\mathbf{x}_i) \Theta(\varepsilon - d_{\min}), \quad (9)$$

where d_{\min} is the minimal state-space distance (within the finite simulation time) between the fixed point \mathbf{x}_* and the transient $T(\mathbf{x}_i)$ and $\Theta(x)$ is the Heaviside function.

If a trajectory reaches a distance smaller than the threshold ε from the attractor within finite simulation time, we regard it to have reached the attractor. Furthermore, the uniformity of ρ implies $\rho(\mathbf{x}_i) = N^{-1}$. Operationally, for any attractor A (not necessarily a fixed point), we proceed as follows: (1) sample an initial condition \mathbf{x}_i according to ρ ; (2) integrate \mathbf{x}_i in time until it has reached an attractor; (3) if the reached attractor is A , count \mathbf{x}_i towards $S_B(A)$; (4) check if the transient originating from \mathbf{x}_i satisfies the constraint, if so count \mathbf{x}_i towards $S_B^C(A)$; (5) increase $i \rightarrow i + 1$; repeat until $i = N$.

The computational procedure outlined above by which we determine BS and CBS allows us to estimate the uncertainty of our estimates of BS and CBS. Since we consider a uniform perturbation ρ , we are effectively drawing initial conditions at random from the subset R of phase space where ρ is nonzero. The fraction p of the volume the basin $B(A)$ occupied by R is the true BS, i.e., the probability that we draw an initial condition from $B(A)$ at random. This implies that, effectively, our estimate of BS after drawing N initial conditions comes from a binomial distribution with expectation value $p = S_B$, which leads to the standard deviation

$$\sigma_{S_B(A)} = \frac{1}{N} \sqrt{S_B(1 - S_B)N} = \frac{1}{\sqrt{N}} \sqrt{S_B(1 - S_B)}. \quad (10)$$

Equation (10) also holds when $S_B(A)$ is replaced by $S_B^C(A)$, which follows from an argument analogous to the one above. It is important at this point to note that nonuniform distributions ρ are also admissible and make sense in certain applications when some perturbations need to be weighted more than others. However, an error estimate as in Eq. (10) is less straightforward to obtain for nonuniform ρ .

III. APPLICATION

To illustrate the versatility of CBS, we give examples of specific constraints in the paradigmatic Lorenz system [14] and in a global carbon cycle model by Anderies *et al.* [15].

In the Lorenz63 (L63) model we show how CBS can reveal precursory phenomena before the onset of a boundary crisis bifurcation. In the Anderies model, we argue that CBS reflects our intuition of stability of a desired state against perturbations better than standard BS. We illustrate in both examples how CBS can generate important new insights into the dynamics of complex systems, while being simple enough to be amenable to a quick interpretation.

A. Anticipating a boundary crisis bifurcation

The L63 system [14]

$$\dot{x} = \sigma(y - x), \quad (11)$$

$$\dot{y} = rx - y - xz, \quad (12)$$

$$\dot{z} = xy - bz \quad (13)$$

is a conceptual model of Rayleigh-Bénard convection. It is famous for exhibiting chaotic dynamics along with a rich dynamical behavior. Setting $\sigma = 10$ and $b = 8/3$, we begin by summarizing the bifurcation structure as the parameter $r \in [9, 26]$ increases. At first, two stable fixed points exist at $\mathbf{x}_*^{(\pm)} = (\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$, corresponding to left and right turning convection rolls, respectively. At $r_1 = 13.926$ a chaotic saddle appears. At $r_2 = 24.06$ this chaotic saddle undergoes a boundary crisis and becomes attractive. The fixed points lose their stability at $r_3 = 24.74$. Our goal is to anticipate this boundary crisis [17]. To this end, we choose a specific condition sensitive to long (chaotic) transients, since these are precursors of the crisis. With the following static constraint, we discriminate between transients that stay close to one of the fixed points \mathbf{x}_+ or \mathbf{x}_- (i.e., one sense of convective overturning) and chaotic transients that flip between them:

$$C^\pm = \{\mathbf{x} \in \Gamma \mid \phi^t(\mathbf{x}) \cdot \mathbf{n} \neq 0 \forall t > 0\}, \quad (14)$$

where the difference vector $\mathbf{n} = |\mathbf{x}_*^+ - \mathbf{x}_*^-|^{-1}(\mathbf{x}_*^+ - \mathbf{x}_*^-)$ is the normal of a plane H containing the origin that separates the phase space into two symmetric halves. Figure 2 shows two-dimensional cross sections of the three-dimensional basins of attraction $B(\mathbf{x}_*^\pm)$ and their intersections $C^\pm \cap B(\mathbf{x}_*^\pm)$ with the sets C^\pm defined above for two different values of the parameter r .

BS and CBS are computed for 15 625 initial conditions sampled from the uniform perturbation distribution $\rho(\mathbf{x}) = \frac{1}{40}\Theta(20 - |x|)\frac{1}{40}\Theta(20 - |y|)\frac{1}{30}\Theta(15 - |z|)$, which describes a box that roughly covers the attractor. To compute BS, we evolve each initial condition in time until it reaches either one of the fixed points or the chaotic attractor. The termination condition in the former case is that the trajectory enters an ε -ball around \mathbf{x}_*^\pm ; here $\varepsilon = 10^{-4}$. In the latter case, we iterate until the trajectory has crossed the plane H a large (but computationally feasible) number m of times; here $m = 400$. Figure 3 shows the resulting BS and CBS. The BS and CBS curves are identical for both fixed points due to their symmetry. Thus only the values for the positive fixed point are shown. For any conditioned set C and its complement $\bar{C} = \Gamma \setminus C$, Eq. (7) reduces to $S_B(A) = S_B^C(A) + S_{\bar{C}}^C(A)$. This implies that the difference between the two stability measures reflects the fraction of the basin from where (long) chaotic transients

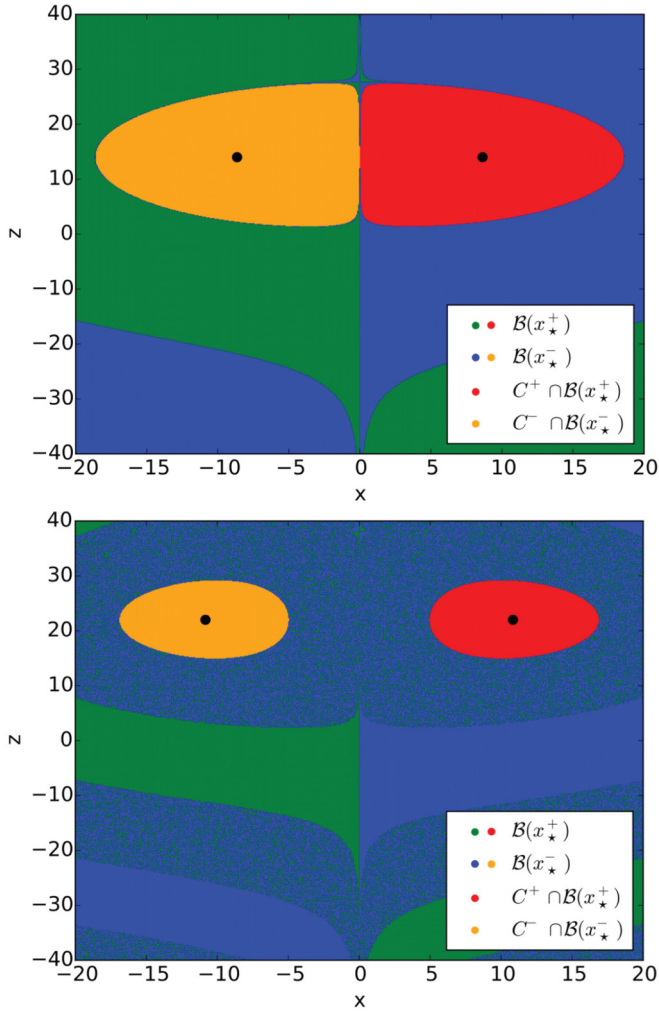


FIG. 2. Illustrations of the phase space structure of the L63 system defined by Eqs. (11)–(13). Both panels show cross sections, obtained by cutting along the plane containing the origin with normal $(1, 1, 0)$, of the basins of attraction $\mathcal{B}(x_\star^\pm)$ and their intersections $C^\pm \cap \mathcal{B}$ with the sets C^\pm for $r = 15$ (upper panel) and $r = 23$ (lower panel). In both panels, the blob-shaped region around each fixed point (black dots) is $C^\pm \cap \mathcal{B}$. Top panel: the total basin of a fixed point is composed of successive layers: it is given by $C^\pm \cap \mathcal{B}$ combined both with the region in the respective other half-space ($x > 0$ or $x < 0$) directly surrounding $C^\pm \cap \mathcal{B}$ and with the next layer of the same color in the fixed point's half plane. In the lower panel (coloring identical), the fractal structure of the basins, visible as intermingled green and blue sets, is apparent; it is associated with transient chaos. One observes that the fraction of the window covered by the sets C^\pm shrinks as r increases. This illustrates the general behavior observed in the L63 system and quantified by CBS, namely that the volume fraction of the three-dimensional sampling region occupied by C^\pm decreases continuously as r approaches r_3 .

originate. The magnitude of CBS reflects the opposite, i.e., the part of the basin from where trajectories fall into the fixed point without crossing H . For $9 \leq r \leq 14$, the two stability measures are constant but differ in their value. For $14 \leq r \leq 23$ the fraction of chaotic transients increases continuously (in agreement with Fig. 2), while the basin volume, i.e., BS, does not change. Between $r \approx 23$ and the bifurcation point at r_2 , the

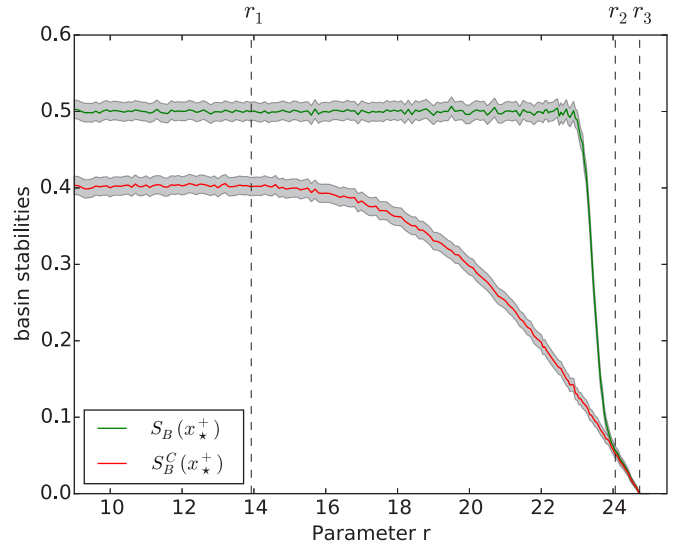


FIG. 3. BS [green (upper) line] and CBS [red (lower) line] of the positive fixed points in the L63 system as the parameter r is varied. From left to right, the vertical lines indicate the appearance of the chaotic set (r_1), the boundary crisis (r_2), and the stability loss of the fixed point (r_3). The difference between the two curves represents the fraction of the basin from where chaotic transients evolve. Note that BS exhibits a jump at r_2 which is blurred by the finite simulation time: the simulation stopped before all (increasingly long) transients had reached the fixed point. Because of the system's x - y symmetry the negative fixed point has the same BS and CBS. The gray envelopes represent $\pm 3\sigma$ according to Eq. (10).

basins of the fixed points collapse as the fraction of trajectories exceeding m crossings grows rapidly. In the limit $m \rightarrow \infty$, the basin size changes discontinuously at r_2 : the fraction of the basin that corresponds to the chaotic transients at $r < r_2$ suddenly feeds the newly born attractor when $r > r_2$. In Fig. 3 the drop is not discontinuous due to finite simulation time: long transients do not reach the fixed points before the simulation is ended. From r_2 to r_3 , both stability measures coincide because chaotic transients are absent due to the chaotic saddle now being attractive. For $r > r_3$, both BS and CBS vanish. In short, with BS alone we cannot anticipate the crisis at r_2 . However, combining it with CBS, the emergence of the chaotic set can be observed by the increasing fraction of chaotic transients within the fixed points' basins. Although CBS does not predict the bifurcation at r_2 , it does indicate the approaching crisis, while BS does not and thus it substantially complements the original BS by revealing additional information on the basin structure.

B. Anderies carbon cycle model

Anderies *et al.* [15] present a conceptual model of global carbon cycle dynamics in the Earth system formulated as a mass balance between three carbon stocks $\mathbf{x} = (c_a, c_t, c_m)$ which are nondimensional atmospheric c_a , terrestrial c_t , and marine stocks c_m , respectively. Formally, the model for the preindustrial case is given by

$$\dot{c}_t = P_{EN}(c_a, c_t) - H(c_t), \quad (15)$$

$$\dot{c}_m = D(c_a, c_m), \quad (16)$$

where total carbon in the system is conserved such that $c_a + c_t + c_m = 1$ and $c_a, c_t, c_m \geq 0$. The expressions describing the derivatives \dot{c}_t, \dot{c}_m are defined as follows: a harvesting term $H(c_t) = \alpha c_t$, where α determines the human offtake of terrestrial carbon stocks, a diffusion term between atmosphere and ocean $D(c_a, c_m) = 0.05(c_a - c_m)$, and net ecosystem productivity $P_{EN}(c_a, c_t) = 2.5c_t(1 - c_t/0.7)\{1.5c_a^{0.3}220T(c_a)^3 \exp[-7T(c_a)] - 110T(c_a)^4 \exp[-5T(c_a)]\}$ with $T(c_a) = 0.8c_a + 0.2$.

It is found in [15] that any initial condition converges to one of two fixed points of interest: either a desirable state $\mathbf{x}_*^d = ((c_a)_*^d, (c_t)_*^d, (c_m)_*^d)$ with vegetation or an undesirable global desert state \mathbf{x}_*^{ud} . At low values of $\alpha \in [0, 0.6]$, the desirable state is attractive, while the undesirable state is repulsive. At $\alpha_{\text{crit}} \approx 0.4$ a transcritical bifurcation occurs and the fixed points reverse their stability. Anderies *et al.* [15] study their model in the context of planetary boundaries interacting with each other. In order to define a safe operating space, they suggest to classify trajectories by whether they return to a certain small ε ball around the desirable fixed point \mathbf{x}_*^d by a certain critical time t_{crit} . Translating this into our framework, we obtain the condition

$$C = \{\mathbf{x} \in \Gamma \mid \exists t < t_{\text{crit}} \text{ such that } |\phi^t(\mathbf{x}) - \mathbf{x}_*^d| < \varepsilon\}, \quad (17)$$

where we choose $\varepsilon = 10^{-4}$. The constraint formulated in Eq. (17) is an integrated constraint since it depends on time. The motivation for this choice of constraint is that, although all trajectories converge to \mathbf{x}_*^d as $t \rightarrow \infty$ for $\alpha < \alpha_{\text{crit}}$ (since then \mathbf{x}_*^d is globally attractive), some trajectories pass very closely and slowly by \mathbf{x}_*^{ud} . These trajectories would entail catastrophic consequences for life on the planet. Therefore, they are identified by whether they exceed a certain return time threshold. Thus we compute CBS of the desirable fixed point based on Eq. (17). We consider a perturbation density ρ describing a depletion of the terrestrial carbon stock c_t (e.g., by immense wildfires). The released carbon is fed into the atmospheric carbon stock c_a . We implement this scenario using a uniform perturbation density on a line in phase space: the terrestrial carbon stock is depleted to a value $c_t \in [0, (c_t)_*^d)$, while the marine carbon stays constant, $c_m = (c_m)_*^d$, and the atmospheric carbon increases according to the carbon conservation law $c_a = 1 - c_m - c_t$. We draw $N = 500$ initial conditions. Figure 4 shows the set C in two-dimensional phase space for two different values of α . The more detailed dependence of BS and CBS on α is shown in Fig. 5: BS is discontinuous at α_{crit} (within numerical accuracy), whereas CBS exhibits a smooth monotonic decay from 1 to 0 on the interval $\alpha \in [0.02, 0.31]$. This reflects the fact that perturbations result in undesirable trajectories much more frequently as human carbon offtake increases until, at $\alpha \approx 0.31$, the return time for the considered perturbations always exceeds t_{crit} . Even though $\alpha < \alpha_{\text{crit}}$, none of the perturbed trajectories can avoid passing through a long quasiglobal desert state. On this parameter interval the desired state is unstable with respect to CBS but stable with respect to BS. We suggest that the former measure provides a more meaningful notion of Earth system resilience from an anthropocentric point of view: it measures the probability that perturbations decay within a predefined acceptable time horizon, while the

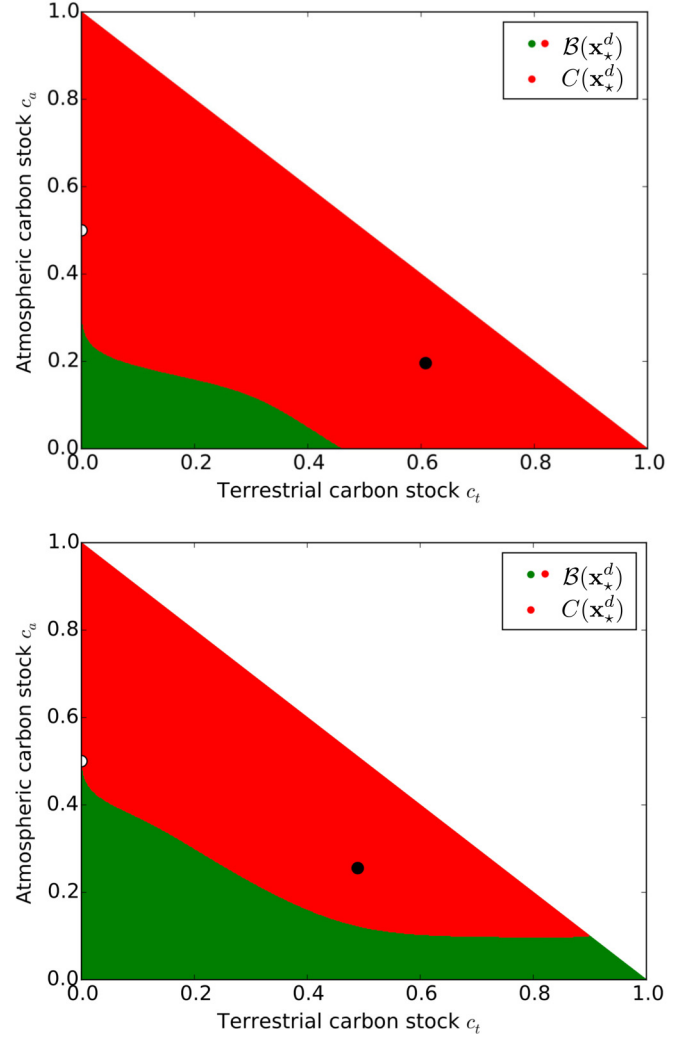


FIG. 4. Illustrations of the phase space structure of the Anderies system, Eqs. (15) and (16). Both panels show the (globally attracting) desirable fixed point (black dot) and the set C (red, upper part of triangle) and its complement B/C (green, lower part of triangle) for $\alpha = 0.15$ (upper panel) and $\alpha = 0.35$ (lower panel). The white dot at $(0, 0.5)$ is the desert state fixed point. One observes that the fraction of the two-dimensional finite phase space covered by the set C shrinks as α increases, in agreement with Fig. 5.

latter only measures the probability of returning within any (possibly infinite) time horizon. Further, CBS captures the change in transient structure and therefore reveals a signal of the transition in the Anderies model already at values of α significantly smaller than α_{crit} . In contrast, BS is discontinuous (within numerical accuracy) at α_{crit} and does not exhibit any precursory phenomena. Figure 4 shows the set C defined above in Eq. (17) and B , the basin of attraction of the desirable fixed point for two different values of α .

IV. DISCUSSION

We have defined CBS as a generalization of BS, thereby combining an asymptotic stability measure with information retrieved from transient behavior into a compact and intuitive measure. While BS is computed from an attractor's basin,

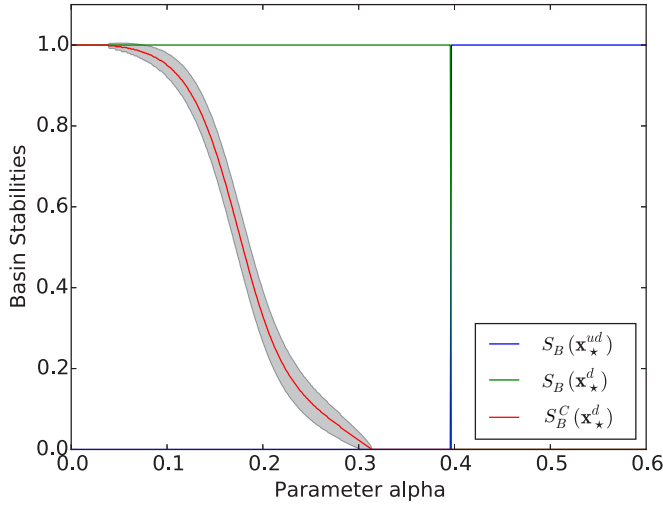


FIG. 5. BS (straight lines) and CBS (curved red line) of \mathbf{x}_*^d and \mathbf{x}_*^{ud} vs the human carbon offtake rate $\alpha \in [0, 0.6]$ for $t_{\text{crit}} = 90$. BS of \mathbf{x}_*^d is represented by the green straight line ($\alpha \leq \alpha_{\text{crit}} \approx 0.4$) and BS of \mathbf{x}_*^{ud} by the blue straight line ($\alpha \leq \alpha_{\text{crit}}$). At low values of α , the desirable state is stable against any strength of perturbation while the desert state is unstable. For $\alpha > 0.03$, an increasing fraction of perturbation-induced trajectories takes longer than t_{crit} to return to the desirable fixed point until, at $\alpha \approx 0.32$, CBS vanishes. By contrast, BS of both fixed points exhibits a jump at α_{crit} and thus no precursory phenomena can be observed there. The gray envelope represents $\pm 3\sigma$ according to Eq. (10).

CBS is computed from a subset of the attractor’s basin. The subset is defined by the transient behavior of trajectories originating from this subset. Thus CBS represents potentially very complicated transient behavior as an easy to interpret scalar quantity.

To underpin that a compact representation of transient behavior is highly relevant in applications, we have presented two examples using specific constraints on the transients. In the case of Rayleigh-Bénard dynamics in the scope of the Lorenz63 model we used the static constraint that the sense of rotation of convection rolls does not change. Here, CBS uncovers nonlinear precursory phenomena of a boundary crisis bifurcation. In the global carbon cycle model by Anderies *et al.* [15], we have studied the stability of the desirable state for a specific perturbation scenario under the premise that it can be restored within an acceptable time horizon. CBS reflects the fact that long return times to the attractor after a perturbation are not desirable. More generally, these applications demonstrate the following three main advantages of CBS over BS. (i) CBS provides useful information in the case of global attractors, while BS cannot be meaningfully applied (it is always equal to 1). (ii) Sudden changes in basin size are often preceded by a change in transient behavior. Extending linear notions of early warning signals for incipient bifurcations [6], CBS uncovers these nonlinear precursory phenomena in the case of the Lorenz63 model and helps anticipating the boundary crisis. (iii) CBS reflects the fact that certain perturbation-induced transients are often undesirable, e.g., long return times, thus allowing one to define highly relevant stability measures for a specific application.

The importance of BS lies in its applicability to a wide range of dynamical systems in various fields. The concept of CBS is even more general as it encompasses BS as a special case. However, to apply CBS, we must choose a specific constraint, such as a limit on the return time. This choice strongly depends on a specific application, revealing highly relevant information there but potentially not being as useful in other applications. By providing two examples of useful constraints and by defining CBS precisely, formally, and in close analogy to BS, we hope to facilitate the transfer of ideas between different applications and different generalizations of BS. For example, BS has been employed successfully to study power grid stability [2]. CBS could be used to develop more specific notions of stability, e.g., to impose that certain units recover quickly from megaoutages or to constrain the total energy loss on the way of recovery. Another example is ecology where BS has proven to be a useful concept and transients are important [18]. CBS could be used to quantify questions of how fast ecosystems recover or investigate potential early warning signals based on minimal abundances of certain species after transients. More generally, BS has successfully applied in resilience research [19,20] and we expect interesting results from further investigating the notion of constrained resilience based on constraints on transients. We expect that future work on CBS will yield a set of transient constraints that prove valuable across a wide range of different applications.

CBS can be used in both passive and active experimental settings. In the former, we have only limited or no control of the system, e.g., the Earth system. We start with some normative notion of undesirable transients as the time threshold in the Anderies example. Heitzig and Kittel [21] discuss desirability in relation to phase space topology. From a given notion of desirability, a constraint is derived. Then, CBS addresses the question of how stable the system is with respect to perturbation-induced transients given that only some of them are desirable. If a system parameter varies over time, CBS is capable of revealing a stability trend which can justify an action to reverse the parameter change. It remains an open problem how CBS can be inferred experimentally or from observational data if a satisfactory model of the system is not available. In principle, if long time series of some environmental parameter (e.g., forest cover on the Earth’s surface) can be derived from measurements and if many natural perturbations can be observed in the data, such as volcanic eruptions, then these can be exploited to estimate CBS. In the active setting, we use CBS to foster our understanding of the system without needing a normative proposition. The constraint and the perturbation are chosen such that new information about the structure of the basin of attraction is revealed. This situation is analogous to the Lorenz63 example: restricting the number of flips between the two halves of phase space (i.e., the two convection senses), the basin of attraction can be subdivided according to the number of flips. Thus CBS helps to characterize a system that is subject to perturbations. In this active setting, it is easier to measure CBS: the system parameters can be chosen freely and the number of perturbations is not restricted by historic events.

The specific condition (14) is reminiscent partly of the concept of “viability” [22,23], although there are significant differences: in particular, in our case, there are no man-

agement options and, more importantly, we are considering deterministic dynamics while viability theory incorporates stochastic and more generally nondeterministic processes. Furthermore, Eq. (14) can easily be generalized by allowing for trajectories to “pierce through” H once or multiple times—these generalizations are not related to viability theory. Another concept that has certain features in common with the condition (14) is “survivability” [24]. There, too, a desirable region of phase space is designated as in the case of our choice of C^\pm . However, survivability does not incorporate the asymptotic nature of BS: it depends on the fraction of trajectories starting in a designated region of phase space spending a certain time exclusively in that region. In particular, it does not depend on which attractor trajectories converge to in the long-time limit. For these reasons, CBS is different from both viability theory and survivability and presents a broadly applicable concept for quantifying stability of an attractor with respect to a given not only small perturbation, uniting both the asymptotic features of BS and the transient features of survivability. In conclusion, CBS represents a general framework to quantify the stability of attractors with broad applicability

in various fields with an interest in complex dynamical systems, ranging from physics and technology to sustainability science.

ACKNOWLEDGMENTS

The research reported in this article was developed within the working group “Complex systems perspectives on anthropogenic climate change” at the Natur- and Ingenieurwissenschaftliches Kolleg sponsored by the German National Academic Foundation (Studienstiftung des deutschen Volkes) and in the scope of the COPAN project on coevolutionary pathways at the Potsdam Institute for Climate Impact Research. J.F.D. thanks the Stordalen Foundation (via the Planetary Boundary Research Network PB.net) and the Earth League’s EarthDoc program for financial support. We thank the anonymous referees for their useful comments which helped to improve the readability of this paper and we are grateful for insightful discussions with Jobst Heitzig and Boyan Beronov.

A.v.K. and J.J. contributed equally to the reported research.

-
- [1] P. J. Menck, J. Heitzig, N. Marwan, and J. Kurths, *Nat. Phys.* **9**, 89 (2013).
 - [2] P. J. Menck, J. Heitzig, J. Kurths, and H.-J. Schellnhuber, *Nat. Commun.* **5**, 3969 (2014).
 - [3] P. Schultz, J. Heitzig, and J. Kurths, *New J. Phys.* **16**, 125001 (2014).
 - [4] T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press, Cambridge, UK, 2013).
 - [5] M. I. Rabinovich, P. Varona, A. I. Selverston, and H. D. Abarbanel, *Rev. Mod. Phys.* **78**, 1213 (2006).
 - [6] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, *Nature (London)* **461**, 53 (2009).
 - [7] W. Steffen, W. Broadgate, L. Deutsch, O. Gaffney, and C. Ludwig, *Anthropocene Rev.* **2**, 81 (2015).
 - [8] H.-J. Schellnhuber, in *Earth System Analysis: Integrating Science for Sustainability*, edited by H.-J. Schellnhuber and V. Wenzel (Springer, Berlin, 1998), pp. 3–195.
 - [9] H. J. Schellnhuber, *Nature (London)* **402**, C19 (1999).
 - [10] G. Petschel-Held, H.-J. Schellnhuber, T. Bruckner, F. L. Toth, and K. Hasselmann, *Clim. Change* **41**, 303 (1999).
 - [11] J. Rockström, W. Steffen, K. Noone, A. Persson, F. S. Chapin III, E. F. Lambin, T. M. Lenton, M. Scheffer, C. Folke, H. J. Schellnhuber, B. Nykvist, C. A. de Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sorlin, P. K. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. W. Corell, V. J. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J. A. Foley, *Nature (London)* **461**, 472 (2009).
 - [12] W. Steffen, K. Richardson, J. Rockström, S. E. Cornell, I. Fetzer, E. M. Bennett, R. Biggs, S. R. Carpenter, W. de Vries, C. A. de Wit *et al.*, *Science* **347**, 1259855 (2015).
 - [13] K. Raworth, Oxfam Policy Practice: *Clim. Change Resil.* **8**, 1 (2012), <http://policy-practice.oxfam.org.uk/publications/a-safe-and-just-space-for-humanity-can-we-live-within-the-doughnut-210490>.
 - [14] E. N. Lorenz, *J. Atmos. Sci.* **20**, 130 (1963).
 - [15] J. M. Anderies, S. Carpenter, W. Steffen, and J. Rockström, *Environ. Res. Lett.* **8**, 044048 (2013).
 - [16] J. A. Brander and M. S. Taylor, *Am. Econ. Rev.* **88**, 119 (1998).
 - [17] S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (Westview Press, Boulder, 2014).
 - [18] L. Dai, K. S. Korolev, and J. Gore, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 10056 (2015).
 - [19] C. S. Holling, *Annu. Rev. Ecol. Syst.* **4**, 1 (1973).
 - [20] C. Folke, S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C. Holling, *Annu. Rev. Ecol., Evol., Syst.* **35**, 557 (2004).
 - [21] J. Heitzig, T. Kittel, J. F. Donges, and N. Molkenhain, *Earth Syst. Dynam.* **7**, 21 (2016).
 - [22] J.-P. Aubin and P. Saint-Pierre, *Decis. Making Risk Manag. Sustain. Sci.* **2007**, 43 (2007).
 - [23] J.-P. Aubin, A. M. Bayen, and P. Saint-Pierre, *Viability Theory: New Directions* (Springer Science & Business Media, New York, 2011).
 - [24] F. Hellmann, P. Schultz, C. Grabow, J. Heitzig, and J. Kurths, [arXiv:1506.01257](https://arxiv.org/abs/1506.01257).