Modeling and Analysis of Mood Dynamics in the Bipolar Spectrum

Hugo Gonzalez Villasanti⬤, *Member, IEEE*, and Kevin M. Passino, *Fellow, IEEE*

Abstract—This article introduces a nonlinear ordinary differential equation model of mood dynamics for disorders on the bipolar spectrum. Motivated by biopsychosocial findings, the model characterizes mood as a 2-D state corresponding to manic and depressive features, enabling the representation of most diagnoses of bipolar and depressive disorders. We perform a mathematical analysis of conditions for the mood to stabilize to euthymia and discuss its psychotherapeutic implications. Furthermore, a computational analysis applied to pharmacotherapy depicts a mechanism that results in a switch from depression to mania when the bipolar disorder was misdiagnosed as major depressive disorder, and an antidepressant is administered without a mood stabilizer. This work innovates by offering a concise representation of most features of mood disorders in existing mathematical models, providing a framework for studying dynamics in the bipolar spectrum.

Index Terms—Bipolar spectrum, dynamical systems, mood disorders, psychology, stability.

I. INTRODUCTION

BIPOLAR disorders are prevalent and disabling mental disorders that have received significant attention from clinical (e.g., psycho/pharmacotherapy), scientific (e.g., genetic or neurotransmitters), and technological [e.g., electroconvulsive therapy (ECT) or repetitive transcranial magnetic stimulation (rTMS)] perspectives [1], [2]. Foundational to this work is mathematics (e.g., statistics or electromagnetic theory for the axon) and engineering (e.g., electrical engineering for rTMS/ECT devices). Here, we use a mathematical (nonlinear) dynamical system model to integrate scientific findings of bipolar disorders, characterize time trajectories of mood via nonlinear and computational analyses, and connect these to psychotherapeutic practice. Psychodynamic processes have been analyzed from the dynamical systems perspective, particularly the mood "swings" in bipolar disorders or recurrent depression, as will be detailed in the following. Our approach is a qualitative synthesis of multiple existing experimental studies in a dynamical systems-level representation. We build on, then move past, the reductionistic paradigm.

The mathematical and computational modeling and analysis of mood have been studied using different approaches. In [3], the bipolarity of mood is characterized using a thermodynamics perspective, and fixed, periodic, and chaotic attractors

are discussed. While no mathematical models were used, such dynamics are partially supported by the self-report data from bipolar patients analyzed in [4], which shows that mood swings are not truly cyclic, but chaotic. Studies that employed specifically designed experiments to validate their models are mostly limited to drift-diffusion models on binary choice ("Flanker task"), studying the role of rumination, attention, and executive functions in mood disorders [5], [6]. However, these studies are constrained to only specific features of mood. On the other hand, nonlinear stochastic models have been constructed to represent several dynamic features of mood disorders, with different levels of connection to neurobiological and psychological determinants. The effects of noise on bifurcations and episode sensitization in mood disorders have been described with linear oscillator models [7], [8]. In [9] and [10], bipolarity is envisioned as arising from two possible types of regulation of a bistable system that results in mood oscillations characterized by two variables, depression and mania, with some connections to mood disorder determinants and therapy. Analysis in [11] considered a nonlinear limit cycle model of mood variations based on biochemical reaction equations. In [12], a model of the behavioral activation system (BAS), linked to bipolar disorder episodes, is constructed via a nonlinear stochastic model. The work studies monostability and bistability of mood oscillations by comparing simulations with empirical and observational data. Mood regulation is analyzed in [13] with an inverted pendulum model, and therapy interventions are represented as feedback controllers. The mixed state is accounted for in the oscillator model in [14]. While these studies focus on a particular subset of depressive disorder determinants using up to two state variables, integrative dynamical models accounting for biological and psychological determinants of depressive disorder were proposed in [15], with an emphasis on psychosocial states and, in [16], with an emphasis on neurobiological factors. Additional models considering depression are in [17] where a finite-state machine is used and in [18] with a stochastic model of random aspects of mood. Also, other general computational/mathematical models are used in psychiatry and psychology (e.g., see [5], [19], and [20]) relevant to the abovementioned models. However, all the abovementioned studies are limited in terms of experimental model validation and computational and mathematical analysis.

In this article, unlike the past research mentioned earlier, the mathematical model represents a broader range of features: 1) euthymia, mania, depression, the mixed state, anhedonia, hedonia, and flat or blunted affect, all on a continuum of

numeric values that represent the extent to which someone possesses a mood characteristic or its "severity level"; 2) mood dynamics for smooth variations between mood states, e.g., mania to euthymia to depression, and back, and severity levels for these as mood changes continuously; and 3) "attractors" that mood can fall in to and get stuck (e.g., in a severely depressed state), parameterized in terms of a person's diagnosis (e.g., BD-I or MDD), and other characteristics of a person (defined below); some justification for such attractors approach is given in [21] where such basins were found experimentally for a group of depressed of patients. A computational analysis of mood attractors is given for multiple cases on the bipolar spectrum (e.g., for BD-I, BD-II, cyclothymia, and MDD with attractors for euthymia, depression, mania, and the mixed state). This shows how the characterization of mood disorders here can be related to a "dimensional diagnosis," i.e., when mood stays in certain regions or visits a region for some length of time, a specific type of disorder is indicated. Also, unlike the abovementioned work, our mathematical model and analysis provide conditions for when a person will return to euthymia no matter how their mood is perturbed, i.e., when euthymia is a "global" mood attractor. It is explained how these conditions have clinical implications, in psycho and pharmacotherapy, to stabilize a person to euthymia. Also, our analysis provides a novel explanation of the mechanism underlying the mood stabilizer [e.g., lithium carbonate (Li_2CO_3)], which illustrates how it changes from being an antimanic to antidepressant agent, and also how it operates in the mixed state, serving simultaneously as an antimanic and antidepressant agent. These results resolve the statement where researchers identify the "…paradoxical effects of Lithium as both an antidepressant and antimanic agent" [1]. Next, via simulations, it is shown that, for some persons with bipolar disorder under a depression episode, antidepressants' administration could result in a mood trajectory that moves to the fully manic state.

To the limit scope of our work, we ignore: 1) details of emotion regulation and "fast" dynamics of emotions that occur on a time scale of less than 2–5 s [22] (but do consider the longer term influence of emotions on mood as in [23]); 2) effects of stress (e.g., see [24]); 3) mood-congruent attention (see [1], [23], and [25]); 4) positive/negative rumination and interepisode features (e.g., [26] and [27]); 5) BAS/behavioral inhibition system (BIS) sensitivity (e.g., [28] and [29]); 6) goal pursuit (e.g., [30]) and goal dysregulation [31]; and 7) "slow" dynamics that occur on a time scale of multiple years (e.g., seasonal influences [1], [32]). We broadly exclude the neural level, but the model nonlinearities are informed by it via referenced scientific studies.

This work aims to uncover principles of mood dynamics common across the spectrum. We include features via how they influence mood, not via an explicit model of their dynamics (e.g., a dynamical model of emotion regulation dysfunction or the coupling between mania and sleep). Our ability to consider cross-spectrum issues arises due to the use of a "dimensional approach," as opposed to the "categorical" one in [33], since a comparative analysis of these two approaches for modeling work is beyond the scope of this article.

II. METHODS

A. Mood States, Trajectories, and Regions

The mood is represented in two dimensions here; however, consider briefly the 1-D case where the mood state is simply a scalar. Let $D(t) \geq 0$ and $M(t) \geq 0$ represent the severity of depression (respectively, mania) at time $t \geq 0$ (these are "diagnosis-level" variables, not specific symptoms). Consider Fig. 1(a) (see caption first), but ignore everything except the red/blue vertical line. Place both $D(t)$ and $M(t)$ on this line (one dimension). Thinking of depression and mania as opposites, assume that $D(t)$ is measured on the blue vertical axis and $M(t)$ on the red vertical axis. "×" represents normal/euthymic. If $D(t)$ starts at normal and its severity increases, mood decreases in the downward direction along the blue line (toward a bottom "pole"). If $M(t)$ starts at normal and its severity increases, the mood point increases in the upward direction along the red line (toward the other pole). Representing mood in this manner allows for the colloquial manner of discussing "mood swings" going "up and down" that ignores the mixed state.

The "mood state" [1] is represented here via two independent variables at time $t \geq 0$, depressive mood $D(t)$, and manic mood $M(t)$. Numerical values and scales could result from standard instruments for measuring depression (e.g., BDI or HAM-D; see [34]), mania (e.g., YMRS; see [1]), or the mixed state (e.g., see [35] and [36]). Here, we assume that the numeric results from such instruments are aggregated or scaled so that they take on values between zero and one. Then, the variables $D(t)$ and $M(t)$ can be represented on a standard Cartesian plane (2-D plot), with the horizontal being $D(t)$ and the vertical $M(t)$. Then, all combinations of severity levels for depression and mania can be represented via $D(t)$ and $M(t)$, with values between zero and one (i.e., $D(t) \in [0, 1]$ and $M(t) \in [0, 1]$) representing, in general, the mixed state when the two variables are nonzero (i.e., $D(t) > 0$ and $M(t) > 0$). In Fig. 1, we rotate the standard 2-D axis by $-45°$ to obtain a "mood plane." There are three advantages to this rotation: 1) the resulting mood plane generally corresponds to traditional descriptions of mood disorders, such as BD-I, where mania corresponds to "up" and depression corresponds to "down"; 2) the mood plane highlights not only "bipolarity" of BD-I but also the mixed state and flat affect poles; and 3) it conceptualizes euthymia as a type of mixed state where there is a normal mix of emotions that create that state.

Next, consider the mood plane in Fig. 1(a). Different mood states are represented with black dots. Mood change directions are represented by the arrows (vectors). For example, if the mood state represents low depression and hypomania (dot, upper left), the vector represents that depression stays constant, but mania increases. The black dot (again, upper left plot) in the center/bottom represents that someone is in a mixed state with more depression than mania, and the vector pointing up means that depression is decreasing at the same rate as mania is increasing. In Fig. 1(b), the mood state is in a euthymic region (black circle centered at the green dot), with "normal" defined for a person or population (refer the following). A "mood trajectory" is a time-sequence of mood
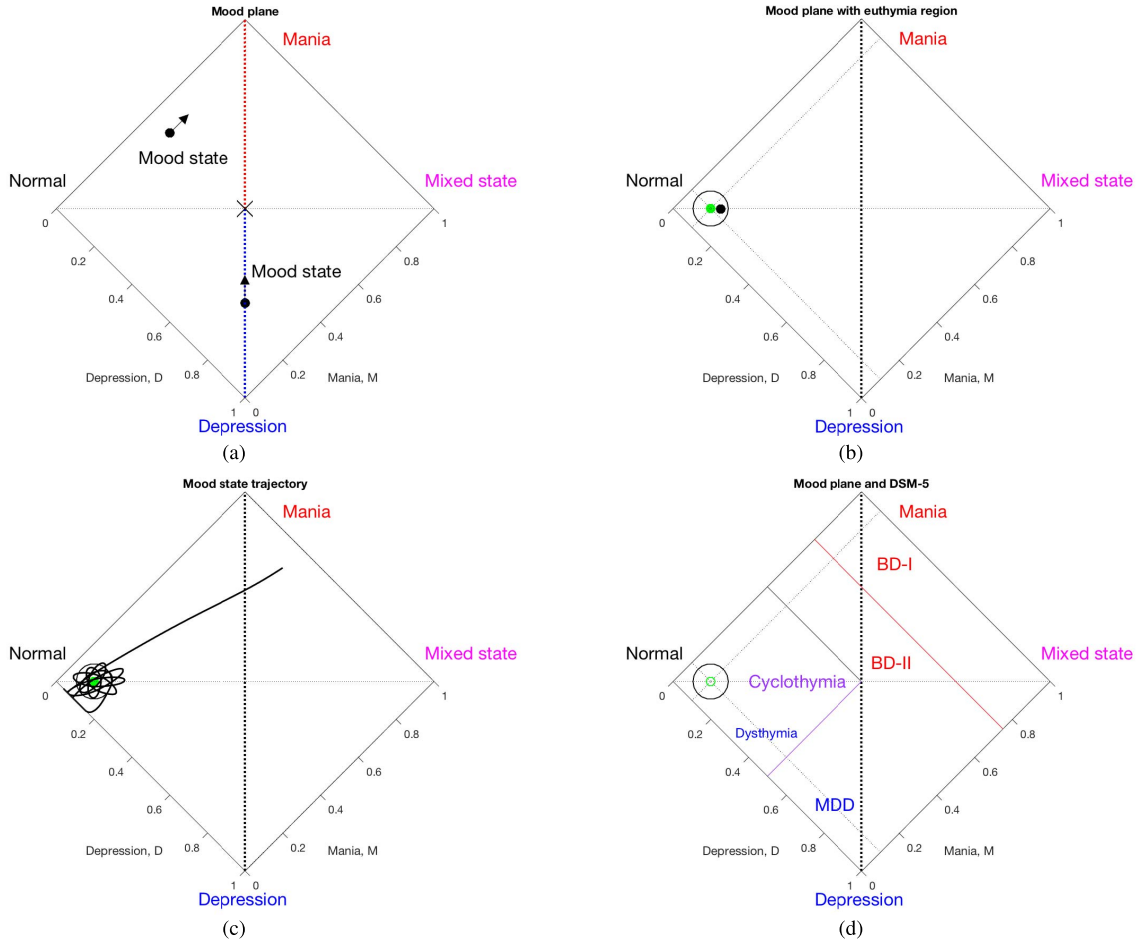
Fig. 1. For each figure, we rotate the standard 2-D axis by −45° to obtain a "mood plane." (a) 1-D and 2-D representations of mood. (b) Mood state is in a euthymic region (black circle centered at the green dot). (c) Mood trajectory example with mood starting in a manic/mixed state and decreasing to euthymia with mood lability. (d) Regions on the mood plane of maximal mood variations for each of the bipolar spectrum disorders [1], [33].

states. Fig. 1(c) shows a mood trajectory example, with mood starting in a manic/mixed state (upper right) and decreasing to euthymia, but with mood lability (see [37] for a discussion on this case); many other mood trajectories will be considered in the following.

Fig. 1(d) shows regions on the mood plane of maximal mood variations for each of the bipolar spectrum disorders [1], [33]. The general description of disorders on the mood spectrum as vertical lines representing mood variations has, e.g., a longer vertical line for BD-I than, e.g., cyclothymia, such as mood swings over a wider range for BD-I (e.g., see [1, pp. 22 and 23, Fig. 1-1]). When considering the mood plane, the mood variation region for one disorder (e.g., cyclothymia) is a subset of another disorder (BD-I), but this is only in terms of mood variation. An "ordering" of the disorders on the spectrum analogous to the one in [1] holds but now in terms of subsets. The dimensional characterization of mood disorders in Fig. 1(d) is related to the one used in [38, p. 145, Fig. 8.1].

Let the fixed constants $n_d \in [0, 1]$ and $n_m \in [0, 1]$ represent a point in the mood plane. Referencing this point, other mood features can be added to the mood plane cases in Fig. 1:

1) *Euthymia:* For $n_d \in [0, 1]$ and $n_m \in [0, 1]$, we assume that, if $D(t) = n_d$ and $M(t) = n_m$, this describes

"normal" or "euthymic" mood [1], [34]. The values of n_d and n_m could be specified for an individual via assessment or a population by averaging individual assessments. As an example, in Fig. 1, $n_d = n_m = 0.1$ is used to represent the center of an euthymic region.

2) *Extreme Mood States:* $D(t) = 1$ or $M(t) = 1$ represents maximally severe depression (respectively, mania), and if $D(t) = M(t) = 1$, this represents a maximally severe mixed state [33].

3) *Mixed States:* Intermediate values of $D(t)$ and $M(t)$ represent mixed mood states. If $n_d = n_m = 0.1$, mixed state examples include the following.
   a) $D(t) = 0.75$ and $M(t) = 0.1$ representing very depressed but no mania compared with normal (e.g., an MDD state).
   b) $D(t) = 0.2$ and $M(t) = 0.1$ representing light depression but no mania compared with normal (e.g., a dysphoric state).
   c) $D(t) = 0.25$ and $M(t) = 0.25$ representing moderate depression compared with normal and moderate mania compared with normal (e.g., as in a mixed state in cyclothymia).
   d) $D(t) = 0.1$ and $M(t) = 0.4$ representing no depression compared with normal but hypomania (e.g., as in BD-II).

e) $D(t) = 0.75$ and $M(t) = 0.75$ representing a mixed state (e.g., in BD-I).

4) *Absence of Depression and/or Mania:* $D(t) = 0$ ($M(t) = 0$) represents the total absence of depression (respectively, mania), and if $D(t) = M(t) = 0$, this represents "flat affect" [33]. $D(t) \geq 0$ with $M(t) = 0$ represents "anhedonia." $D(t) = 0$ with $M(t) \geq 0$ represents "hedonia" [1].

B. Mood Dynamics and Equilibria

The time units adopted in this work are days. We are considering adults who do not experience full-range mood swings within 24 h [1]. Even though mood can vary with time, for simplicity, we frequently drop the notation for time dependence and use D and M (similarly, for other variables).

Mood dynamics are represented by the differential equations

$$\frac{dD}{dt} = S_d(D, M, u_d)$$
$$\frac{dM}{dt} = S_m(D, M, u_m) \quad (1)$$

with nonlinear functions $S_d(D, M, u_d)$ and $S_m(D, M, u_m)$ specifying the rates of change of mood (derivatives, (dD/dt) and (dM/dt)), where $u_d(t)$ and $u_m(t)$ are the internal/external inputs to the depressive and manic dynamics, respectively, e.g., from stimuli psychophysiological response systems.

Consider the "unforced" mania mood dynamics, that is, without the influence of depressive mood or other external inputs and outputs so that $\dot{M} = (dM/dt) = S_m(0, M, 0)$. Define

$$\dot{M} = b_m \frac{a_m(M - n_m + c_m)^2}{a_m(M - n_m + c_m)^2 + 1}$$
$$- d_m \frac{f_m(M - n_m - g_m)^2}{f_m(M - n_m - g_m)^2 + 1} - h_m(M - n_m). \quad (2)$$

In the depressive case, $\dot{D} = S_d(D, 0, 0)$ is defined in an analogous manner. Solutions to these differential equations exist and are unique since S_m and S_d are continuous and satisfy the Lipschitz conditions. The double-sigmoid concept has been employed to model systems with multiple equilibria, of which possibly the most representative and relevant to our work is the Wilson–Cowan model of the interaction between populations of excitatory and inhibitory neurons [39]. To illustrate why the shape of the nonlinear function $S_m(0, M, 0)$ represents key features of mood dynamics, consider an example. Let $n_m = 0.5$, $b_m = 0.07$, $d_m = b_m$, $c_m = 0.34$, $g_m = c_m$, $a_m = 29$, $f_m = a_m$, and $h_m = 0.19$. Fig. 2(a) shows the linear decay line (blue, diagonal, see the third term in (2) versus M), the sum of the rational functions that compose the double-sigmoid (red, see first two terms in (2)) versus M), and $\dot{M} = S_m(0, M, 0)$ versus M [magenta, right-hand side of (2))]. For the linear decay (blue) plot, for a given value of $M \geq n_m$ ($M < n_m$) on the horizontal axis, there is a negative (positive) value moving mood down (respectively, up); that is, $-h_m(M - n_m)$ tries to stabilize mood to normal. The first two terms in (2) versus M, the red line, have: 1) no influence at $M = n_m$; 2) an increasing positive (negative)
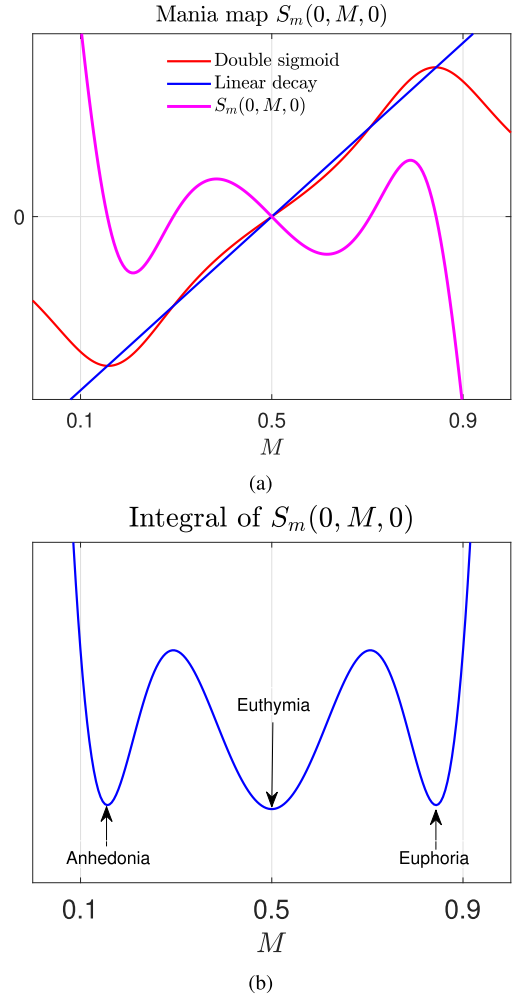

Fig. 2. (a) Unforced manic mood dynamics functions. (b) Basins of attraction for equilibria via integration of $S_m(0, M, 0)$ and interpretation of each attractor.

influence on mood change as M moves to intermediate values above (below) $M = n_m$ representing destabilizing effects on mood (e.g., making $\dot{M}$ positive when $M > n_m$ so it increases further); and 3) a lower positive (negative) influence on mood change as M moves above the peaks in the red line, representing destabilizing effects on mood that are weaker for high values of $\pm M$. The $\dot{M} = S_m(0, M, 0)$ versus M case, the magenta line, is the sum of all three right-hand side terms in (2), which are the red line and the negative of the blue line. Notice that, by plotting the linear decay versus M, we can see the intersection points in Fig. 2(a), which are the five points on the magenta line that cross zero. These zero points identify M values where $\dot{M} = S_m(0, M, 0) = 0$, that is, where there is no change in mood, up or down (these are "equilibria"). For instance, at $M = n_m$, $\dot{M} = S_m(0, M, 0) = 0$, so that when the person is at normal, and there are no influences from depression or internal/external inputs, and then, the mood will stay at normal.

C. Basins of Attraction for Mood

To visualize the dynamics in the vicinity of the five equilibria, imagine drawing arrows on the horizontal axis of Fig. 2(a), with the directions indicating how M will change, as specified

by the sign of $\dot{M}$, for each value of $M \in [0, 1]$. For example, for M values just above (below) $n_m = 0.5$, the arrow will point to the left (right) since $\dot{M} = S_m(0, M, 0) < 0$ ($\dot{M} = S_m(0, M, 0) > 0$, respectively), and this shows graphically that $n_m = 0.5$ is an asymptotically stable equilibrium point. Since such arrows will move away from the point to the right of $n_m = 0.5$ where the bottom of the valley exists, it is called an unstable equilibrium point (if the M value is to the left of the bottom of that valley, M will decrease toward n_m, but, if it is to the right of the bottom of the valley, M will increase, moving away from the valley bottom). Similar analyses work for the other three cases where $S_m(0, M, 0) = 0$.

For another way to view the dynamics and equilibria, consider (1) and (2), and taking a continuous-time gradient optimization perspective on the mania dynamics, we let, for any $M(0) \in [0, 1]$ and all $t \geq 0$

$$\frac{dM(t)}{dt} = -\alpha \left. \frac{\partial J(M)}{\partial M} \right|_{M=M(t)} = S_m(0, M(t), 0)$$

where $\alpha > 0$ is a constant "step size" and $J(M)$ is the ("potential") function to be minimized; one that must be chosen so that the dynamics specified by (1) are matched by this equation. Independent of time $t \geq 0$, for any $M \in [0, 1]$, integrating the two right-hand side terms of this equation, we get

$$J(M) = \int_0^M \frac{\partial J(\lambda)}{\partial \lambda} d\lambda = -\frac{1}{\alpha} \int_0^M S_m(0, \lambda, 0) d\lambda.$$

For $\alpha = 1$, the plot of $J(M)$ is shown in Fig. 2 where there are "basins of attraction" (valleys) for euthymia, anhedonia, and euphoria (in the depression case, there are basins for euthymia, hedonia, and dysphoria, with some justification for this in [21]). The gradient optimization perspective says that, if mood is perturbed from the bottom of one of these basins, it will move to go "down hill" until it reaches the bottom of the basin. Hence, the bottoms of these basins "attract" the mood trajectory $M(t)$ if it is in its vicinity. The peaks on the two hills represent "unstable" points where, if M is perturbed even slightly to the left or right, the mood will tend to move to the left or right more, and hence, M will move away from the peak—the tendency is always to move down the $J(M)$ function if it is not at a peak.

Mood shifts between basins in Fig. 2 has been linked to external and internal stimuli, such as stressors, sleep, and seasonal patterns, and the response to these stimuli by other internal processes, such as physiological arousal or BAS/BIS. For instance, in Fig. 2, if mood M starts at euthymia, it is important to know if other variables (e.g., stress and sleep deficits) can move it out of the euthymia basin, to the right, over the hill, and then down/farther to the right to end up at euphoria. Alternatively, in the analogous diagram to Fig. 2 for depression, if mood M starts at dysphoria, it is important to know if other variables (e.g., an antidepressant) can influence it to move it out of the dysphoria basin, to the right, over the hill, and then down/farther to the right to end up at euthymia. In [40], two main processes are identified in the dynamics of depression: neurobiological processes responsible for mood-congruent cognitive biases in attention, processing, rumination, self-referential schemes, and attenuated cognitive

control to correct these biases. Along those lines, we argue that a basin's size is mostly affected by the degree of biased processing of relevant stimuli. For example, increased processing of negative stimuli by limbic structures in major depressive disorder, which contributes to a reduced stressor tolerance and a reduced threshold for mood switching, is represented by a narrow normal equilibrium basin and wide basin for dysphoria, in the depressive mood case. Also, a basin's depth could be correlated with the reinforcing elements of mood disorders that maintain mood in the basin, increasing abnormal episodes' duration. In this case, biased self-reference schemas, mood-congruent attention, and maladaptive strategies, such as rumination, could increase the abnormal equilibria's basin's depth.

Decreasing the parameters a_m and f_m has a larger effect in increasing the depth ratio between the abnormal and the normal basins' depth. Therefore, low values of these parameters indicate a higher effect of mood-congruency on the cognitive biases, in the abnormal basins. This could increase the duration of episodes of abnormal mood compared with the duration in normal levels. Decreasing c_m and g_m has a greater effect in displacing the unstable equilibria, decreasing the euthymic basin's size. Thus, low values of c_m and g_m reflect biased processing and low switching thresholds to abnormal mood. An increase in b_m and d_m increases both the abnormal basins' depth and size. Thus, high values in these parameters represent high severity of the cognitive biases and self-referential schemas that attract mood to the abnormal equilibria, leading to the long duration or even chronic episodes due to the difficulty of escaping the basin. Low h_m can be associated with attenuated regulatory processes in the prefrontal cortex that regulate mood [40].

III. RESULTS

A. Stabilization to Euthymia

Here, we provide a narrative on the main theoretical result of this work.

Theorem 1 Equivalence of Qualitative Properties: Consider the unforced and coupled mood dynamics given by

$$\dot{D} = b_d \frac{a_d(D - n_d + c_d)^2}{a_d(M - n_d + c_d)^2 + 1} - d_d \frac{f_d(D - n_d - g_d)^2}{f_d(D - n_d - g_d)^2 + 1} - h_d(D - n_d) + q_d(M - n_m)$$

$$\dot{M} = b_m \frac{a_m(M - n_m + c_m)^2}{a_m(M - n_m + c_m)^2 + 1} - d_m \frac{f_m(M - n_m - g_m)^2}{f_m(M - n_m - g_m)^2 + 1} - h_m(M - n_m) + q_m(D - n_d). \tag{3}$$

and assume that the mood profile is symmetric as in Fig. 2(a), with $b_i = d_i$, $f_i = a_i$, and $g_i = c_i$, with $f_i > 0$ and $g_i \neq 0$ for $i = d, m$. Then, the point (n_d, n_m) at euthymia is globally asymptotically stable if

$$\left(h_d - \frac{d_d}{g_d} \right) \left(h_m - \frac{d_m}{g_m} \right) > |q_d||q_m|. \tag{4}$$

Proof: We begin by showing that the point in euthymia is an equilibrium of the system. At (n_d, n_m), the rate of change

of the unforced mood dynamics in (3) is

$$\dot{D} = b_d \frac{a_d c_d^2}{a_d c_d^2 + 1} - d_d \frac{f_d g_d^2}{f_d g_d^2 + 1}$$

$$\dot{M} = b_m \frac{a_m c_m^2}{a_m c_m^2 + 1} - d_m \frac{f_m g_m^2}{f_m g_m^2 + 1}.$$

When $b_i = d_i$, $f_i = a_i$, and $g_i = c_i$ for $i = d, m$, then $\dot{D} = \dot{M} = 0$, and (n_d, n_m) is an equilibrium point. To analyze the stability of this equilibrium point, we propose the change of variables $x_d = D - n_d$ and $x_m = M - n_m$, obtaining

$$\dot{x}_d = -z_d(x_d) + q_d x_m$$
$$\dot{x}_m = -z_m(x_m) + q_m x_d \qquad (5)$$

where

$$z_i(x_i) = \left(h_i - \frac{4 \, d_i f_i g_i}{p_i(x_i)}\right) x_i$$

and

$$p_i(x_i) = (f_i(x_i - g_i)^2 + 1)(f_i(x_i + g_i)^2 + 1)$$

for $i = d, m$. We will proceed in two steps to prove global asymptotic stability of euthymia. First, we will find a Lyapunov function for the uncoupled case when $q_d = q_m = 0$, and in a second step, we will prove stability of the coupled system using the M-matrix results from [41]. Note that when $q_i = 0$, then, by the assumption in (4), we have that $h_i > d_i/g_i$ for $i = d, m$. Furthermore, we have that $p_i(x_i) \geq 1$, and it is a polynomial of degree four, implying that the vector field in (5) is continuously differentiable everywhere. Furthermore, using the fact that $\min_{x_i \in \mathbb{R}} p_i(x_i) = 4 f_i g_i^2 > 0$, we have for $i = d, m$

$$\left(h_i - \frac{4 \, d_i f_i g_i}{p_i(x_i)}\right) \geq \left(h_i - \frac{4 \, d_i f_i g_i}{\min_{x_i \in \mathbb{R}} p_i(x_i)}\right)$$
$$= \left(h_i - \frac{4 \, d_i f_i g_i}{4 \, f_i g_i^2}\right) = \left(h_i - \frac{d_i}{g_i}\right).$$

Then, when $q_i = 0$, we obtain that $x_i z_i(x_i) > 0$, $\forall x_i \neq 0$, and $z_i(0) = 0$ for $i = d, m$. Consider the function $V_i : \mathbb{R} \to \mathbb{R}$

$$V_i(x_i) = \int_0^{x_i} z_i(y) dy$$

for $i = d, m$. The function $V_i(x_i)$ is continuously differentiable, $V_i(0) = 0$ for $i = d, m$, and $|x_i| \to \infty$ implies that $V_i(x_i) \to \infty$. Furthermore, using the fact that $x_i z_i(x_i) > 0$ for all $x_i \neq 0$ and

$$(x_i - 0) \min_{x_i \in \mathbb{R}} z_i(x_i) \qquad \int_0^{x_i} z_i(y) dy$$

we obtain that $V_i(x_i) > 0$ for all $x_i \neq 0$, for $i = d, m$. Therefore, $V_i(x_i)$ is a Lyapunov function candidate for the corresponding subsystem in (5) when $q_i = 0$, for $i = d, m$. Furthermore, the derivative of $V_i(x_i)$ along the trajectories of the corresponding subsystem for $i = d, m$ is

$$\dot{V}_i(x_i) = \frac{\partial V_i}{\partial x_i}[-z_i(x_i)] = -z_i^2(x_i) \leq -\left(h_i - \frac{d_i}{g_i}\right)^2 x_i^2 < 0$$

for all $x_i \in \mathbb{R} - \{0\}$. Therefore, the point $x = (x_d, x_m) = 0$, which corresponds to the equilibrium at euthymia in (n_d, n_m),

is globally asymptotically stable when $q_i = 0$ for $i = d, m$. The second step is to employ the later result to prove that $x = 0$ is globally asymptotically stable when $q_i \neq 0$ for $i = d, m$ in (5). To achieve that, we will employ the result [41, Th. 9.2] for interconnected systems. We choose $\phi_i(x_i) = |x_i|$ and obtain, for $i = d, m$

$$\frac{\partial V_i}{\partial x_i}[-z_i(x_i)] = -z_i^2(x_i) \leq -\left(h_i - \frac{d_i}{g_i}\right)^2 |x_i|^2 = -\alpha_i \phi_i^2(x_i)$$

$$\left\|\frac{\partial V_i}{\partial x_i}\right\| = |z_i(x_i)| \leq \left(h_i - \frac{d_i}{g_i}\right)|x_i| = \beta_i \phi_i(x_i)$$

where we define

$$\alpha_i = \left(h_i - \frac{d_i}{g_i}\right)^2$$

$$\beta_i = h_i - \frac{d_i}{g_i}.$$

Furthermore

$$|q_d x_m| = |q_d|\phi_m(x_m) = \gamma_{dd}\phi_d(x_d) + \gamma_{dm}\phi_m(x_m)$$
$$|q_m x_d| = |q_m|\phi_d(x_d) = \gamma_{mm}\phi_m(x_m) + \gamma_{md}\phi_d(x_d)$$

where

$$\gamma_{dd} = \gamma_{mm} = 0, \quad \gamma_{dm} = |q_d|, \quad \gamma_{md} = |q_m|.$$

Therefore, the matrix S is given by

$$S = \begin{vmatrix} \alpha_d - \beta_d \gamma_{dd} & -\beta_d \gamma_{dm} \\ -\beta_m \gamma_{md} & \alpha_m - \beta_m \gamma_{mm} \end{vmatrix}$$

$$= \begin{vmatrix} \left(h_d - \dfrac{d_d}{g_d}\right)^2 & -\left(h_d - \dfrac{d_d}{g_d}\right)|q_d| \\ -\left(h_m - \dfrac{d_m}{g_m}\right)|q_m| & \left(h_m - \dfrac{d_m}{g_m}\right)^2 \end{vmatrix}.$$

The point $x = 0$ is globally asymptotically stable for the system in (5) if the determinant of matrix S is positive. Using the assumption in (4), we obtain

$$\det(S) = \left(h_d - \frac{d_d}{g_d}\right)^2 \left(h_m - \frac{d_m}{g_m}\right)^2$$
$$- \left(h_d - \frac{d_d}{g_d}\right)|q_d|\left(h_m - \frac{d_m}{g_m}\right)|q_m| > 0.$$

$\square$

Theorem 1 states that a sufficient condition for the stability of the equilibrium at euthymia of the unforced mood dynamics represented in (3) is that there exists a lower bound on the regulation rate h_d and h_m, given by

$$\left(h_d - \frac{d_d}{g_d}\right)\left(h_m - \frac{d_m}{g_m}\right) > |q_d||q_m| \qquad (6)$$

which is independent of f_d and f_m. The parameters q_d and q_m represent the strength of the coupling between manic and depressive mood. This result shows that the maximum slope of the sigmoids in the equilibrium corresponding to euthymia is given by d_i/g_i for $i = m, d$. Hence, this is analogous to saying that, when the mood regulation rate is large enough that it can regulate the cognitive biases, the multistability (e.g., equilibria at euphoria and anhedonia) disappears. Larger h_i for $i = d, m$ can be obtained via psychotherapy that promotes
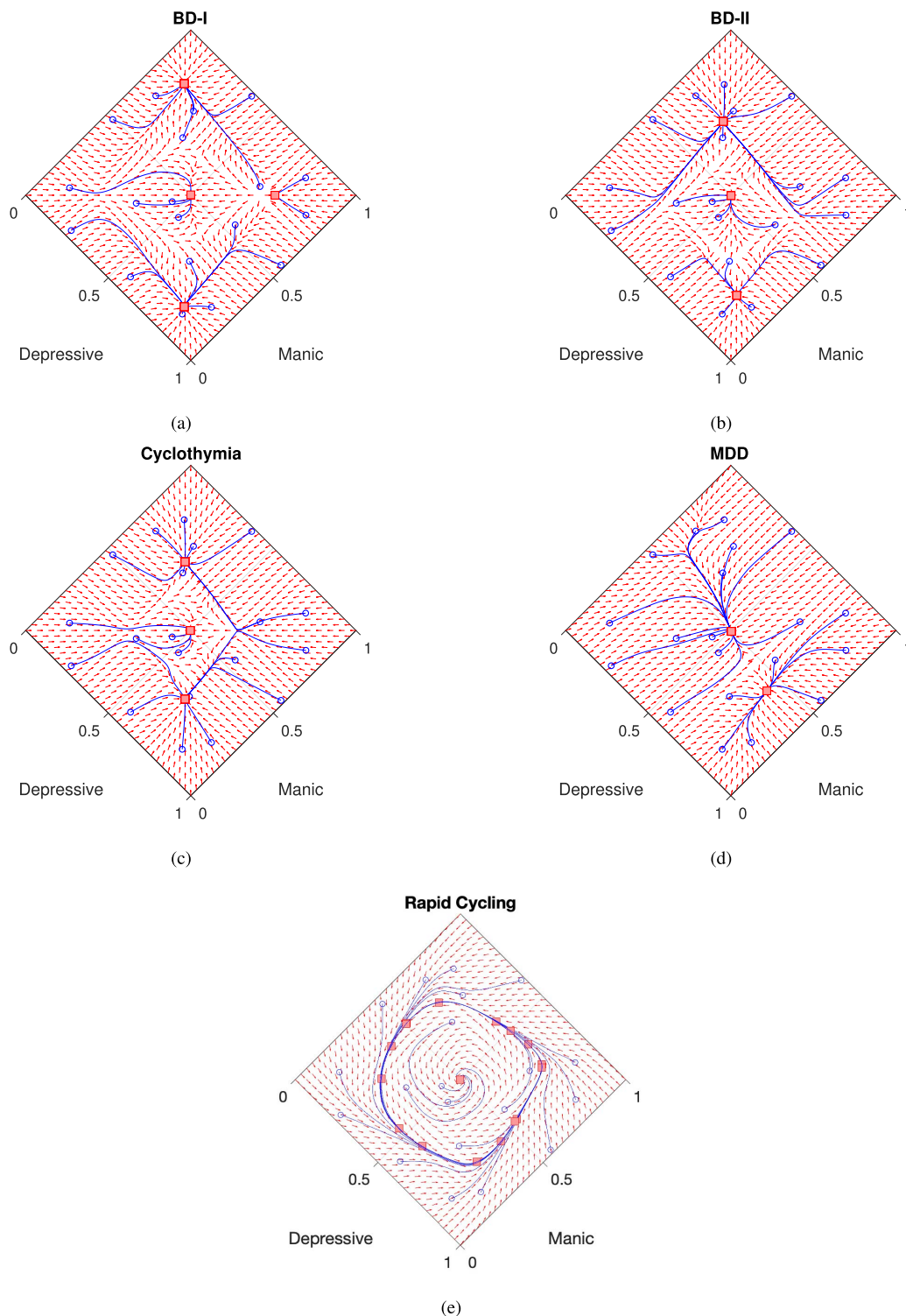
Fig. 3. Vector field diagrams along the depressive and mood axes of unforced dynamics in (2) for (a) BD-I, (b) BD-II, (c) Cyclothymia, (d) MDD, and (e) rapid cycling. Red arrows indicating the direction of mood change for different points in the mood plane. The blue lines represent trajectories with circles at their initial conditions and squares as their final state at time $T = 10$ days. Squares correspond roughly to stable equilibria of the unforced dynamics, except for rapid cycling in case (e), where mood oscillates with periods of approximately six days.

awareness of cognitive biases and tools to regulate emotions and mood, as well as pharmacotherapy that targets biological determinants of the prefrontal cortex activity, such as serotonin levels. The meaning and effects of these parameters are discussed more in [40], [42], [43].

B. Mood Trajectories for Various Mood Disorders

The differential equation model is flexible enough to represent several of the most important mood disorders in the DSM-5. Fig. 3 shows the vector field diagram of the unforced

TABLE I
PARAMETER VALUES IN FIG. 3

	BD-I	BD-II	Cyclothymia	MDD
a_d	90	220	280	54
a_m	90	270	280	90
b_d	0.45	0.4	0.26	0.45
b_m	0.45	0.3	0.26	0.28
c_d	0.3	0.2	0.2	0.4
c_m	0.3	0.25	0.2	0.3
d_d	0.45	0.38	0.28	0.45
d_m	0.45	0.35	0.28	0.28
f_d	130	145	270	130
f_m	130	225	270	130
g_d	0.25	0.3	0.15	0.25
g_m	0.25	0.15	0.15	0.25
h_d	1.3	1.32	1.37	1.3
h_m	1.3	1.38	1.37	2
q_d	-0.35	-0.35	-0.4	-0.35
g_m	-0.35	-0.35	-0.4	-0.35

dynamics in (2) for BD-I, BD-II, cyclothymia, and MDD, with red arrows indicating the direction of mood change for different points in the mood plane. The blue lines represent trajectories with circles at their initial conditions and squares as their final state at time $T = 10$ days. By altering the size and depth of the basins of attraction for the equilibria, for both the depressive and manic dimensions (as in Fig. 2), we can create equilibria at asymmetric locations in the mood plane. The model parameters employed for each diagnosis can be found in Table I. For BD-I in Fig. 3(a), there are four stable equilibria marked with squares, namely, normal/euthymia at $n_m = n_d = 0.5$, depression, mania, and mixed states where the trajectories may converge. Note that the rate of convergence matches the onset times given in the literature, which are, on average, three days for mania and seven days for depression [1]. The vector field diagram of BD-II features an equilibrium at hypomania and major depression and is generated by reducing the severity of mania, i.e., increasing the value of the difference $h_m - d_m/g_m$, while the cyclothymia vector field depicts equilibria at hypomania, and mild depression, as well as in the mixed state, but with increased basin depths to reflect the extended episodic durations in these disorders. MDD features only two equilibria in the normal and major depressive episodes, generated by eliminating the abnormal equilibria along with mania.

The coupling parameters $q_d \in \mathbb{R}$ and $q_m \in \mathbb{R}$ provide the existence of a mixed state, as well as a way to model the cyclic nature of bipolar disorders. Simulations revealed that, by setting opposite signs and increasing the magnitude of the coupling parameters, the reinforcement effect introduces oscillations analogous to what is seen in rapid cycling. The vector field diagram in Fig. 3(e) was obtained based on the BD-I parameters but increasing the coupling parameter three times, i.e., $q_d = -q_m = 1.05$.

C. Pharmacotherapy and Triggering Mania

The effects of pharmacotherapy are modeled as an external input $u(t) = [u_d(t), u_m(t)]$ in (1). For example, the effect of a mood stabilizer, such as lithium, can be modeled as an additive term in (2) with $u_d(t) = K_d(D(t) - n_d)$ and $u_m(t) =$
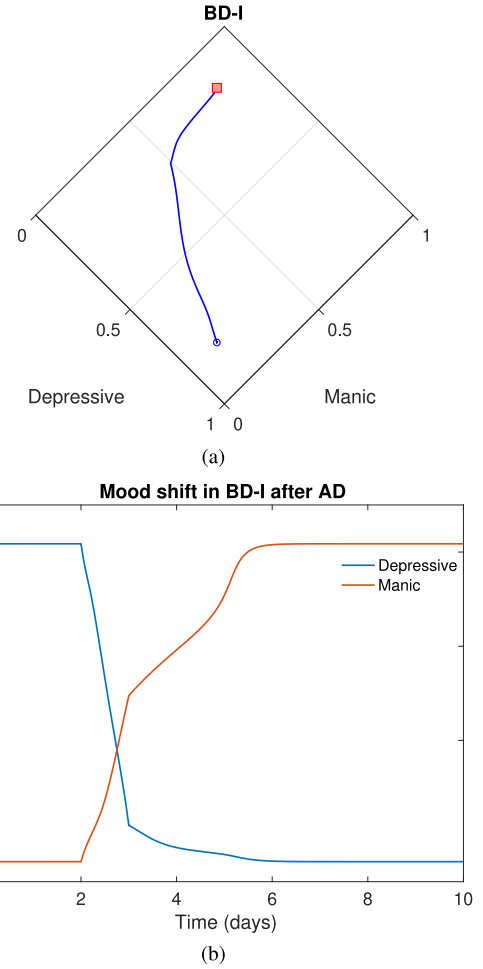

(a)

(b)

Fig. 4. (a) Trajectory (blue) in the BD-I mood plane. (b) Mood episode shift from depression to mania after treatment with antidepressant.

$K_m(M(t) - n_m)$. Simulation results presented in Fig. 4 show the effects of administering antidepressants to a BD-I patient, misdiagnosed with MDD during a depressive episode. Fig. 4(a) shows the mood trajectory in the mood plane, while Fig. 4(b) shows the time trajectories for depressive mood in blue and manic mood in red. The initial mood state corresponds to the depression episode equilibrium, and on the second day, an antidepressant is administered for two days, modeled here as an additive, constant negative input in depressive mood and a smaller, additive positive input in manic mood in (2). Even after stopping the medication, manic and depressive mood are attracted to the equilibrium representing mania, which, according to the DSM-5, constitutes sufficient evidence for a BD-I diagnosis.

IV. CONCLUSION

We introduced a general ODE model for disorders on the bipolar spectrum. Via the Lyapunov stability analysis, we show conditions by which mood is guaranteed to return to euthymia if perturbed off this mood state. We discussed the stability result from the perspective of psychodynamics and psychotherapy. Via computational analysis, we simulated the mood switch from depression to mania when bipolar

disorder is misdiagnosed as major depressive disorder, and an antidepressant is administered without a mood stabilizer. The most critical and challenging direction is to expand the model and validate it against experimental data. Parameter estimation can be performed on more tractable approximations of the model presented here, including piecewise linear affine or the Markov jump linear systems. To estimate these models, intensive longitudinal data on bipolar disorder symptoms, e.g., from ecological momentary assessments or daily logs, can be employed [44], [45].

References

[1] F. K. Goodwin and K. R. Jamison, *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression*, vol. 1. London, U.K.: Oxford Univ. Press, 2007.

[2] R. DeRubeis and D. Strunk, Eds., *The Oxford Handbook of Mood Disorders*. Oxford, U.K.: Oxford Univ. Press, 2017.

[3] H. C. Sabelli, L. Carlson-Sabelli, and J. I. Javaid, "The thermodynamics of bipolarity: A bifurcation model of bipolar illness and bipolar character and its psychotherapeutic applications," *Psychiatry*, vol. 53, no. 4, pp. 346–368, Nov. 1990.

[4] A. Gottschalk, M. S. Bauer, and P. C. Whybrow, "Evidence of chaotic mood variation in bipolar disorder," *Arch. Gen. Psychiatry*, vol. 52, no. 11, pp. 947–959, 1995.

[5] M. L. Pe, J. Vandekerckhove, and P. Kuppens, "A diffusion model account of the relationship between the emotional flanker task and rumination and depression," *Emotion*, vol. 13, no. 4, p. 739, 2013.

[6] D. G. Dillon *et al.*, "A computational analysis of flanker interference in depression," *Psychol. Med.*, vol. 45, no. 11, pp. 2333–2344, Aug. 2015.

[7] M. T. Huber, J.-C. Krieg, H. A. Braun, X. Pei, A. Neiman, and F. Moss, "Noisy precursors of bifurcations in a neurodynamical model for disease states of mood disorders," *Neurocomputing*, vols. 32–33, pp. 823–831, Jun. 2000.

[8] M. T. Huber, H. A. Braun, K. Voigt, and J.-C. Krieg, "Some computational aspects of the kindling model for neuropsychiatric disorders," *Neurocomputing*, vols. 38–40, pp. 1297–1306, Jun. 2001.

[9] A. Goldbeter, "A model for the dynamics of bipolar disorders," *Prog. Biophys. Mol. Biol.*, vol. 105, nos. 1–2, pp. 119–127, 2011.

[10] A. Goldbeter, "Origin of cyclicity in bipolar disorders: A computational approach," *Pharmacopsychiatry*, vol. 46, no. S01, pp. S44–S52, Apr. 2013.

[11] T. D. Frank, "A limit cycle oscillator model for cycling mood variations of bipolar disorder patients derived from cellular biochemical reaction equations," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 18, no. 8, pp. 2107–2119, Aug. 2013.

[12] A. Steinacher and K. A. Wright, "Relating the bipolar spectrum to dysregulation of behavioural activation: A perspective from dynamical modelling," *PLoS ONE*, vol. 8, no. 5, pp. 633–645, 2013.

[13] E. Koutsoukos and E. Angelopoulos, "Mood regulation in bipolar disorders viewed through the pendulum dynamics concept," *Int. J. Bipolar Disorders*, vol. 2, no. 1, p. 9, Dec. 2014.

[14] F. Hadaeghi, M. R. H. Golpayegani, and G. Murray, "Towards a complex system understanding of bipolar disorder: A map based model of a complex winnerless competition," *J. Theor. Biol.*, vol. 376, pp. 74–81, Jul. 2015.

[15] A. K. Wittenborn, H. Rahmandad, J. Rick, and N. Hosseinichimeh, "Depression as a systemic syndrome: Mapping the feedback loops of major depressive disorder," *Psychol. Med.*, vol. 46, no. 3, pp. 551–562, Feb. 2016.

[16] N. Z. Bielczyk, J. K. Buitelaar, J. C. Glennon, and P. H. E. Tiesinga, "Circuit to construct mapping: A mathematical tool for assisting the diagnosis and treatment in major depressive disorder," *Frontiers Psychiatry*, vol. 6, p. 29, Feb. 2015.

[17] S. Demic and S. Cheng, "Modeling the dynamics of disease states in depression," *PLoS ONE*, vol. 9, no. 10, Oct. 2014, Art. no. e110358.

[18] S. Y. van der Werf, K. I. Kaptein, P. de Jonge, J. Spijker, R. de Graaf, and J. Korf, "Major depressive episodes and random mood," *Arch. Gen. Psychiatry*, vol. 63, no. 5, p. 509, May 2006.

[19] A. Bystritsky, A. Nierenberg, J. Feusner, and M. Rabinovich, "Computational non-linear dynamical psychiatry: A new methodological paradigm for diagnosis and course of illness," *J Psychiatric Res.*, vol. 46, no. 2, pp. 428–435, 2011.

[20] R. Ratcliff, T. VanZandt, and G. McKoon, "Connectionist and diffusion models of reaction time," *Psychol. Rev.*, vol. 106, no. 2, pp. 231–300, 1999.

[21] S. L. Johnson and A. Nowak, "Dynamical patterns in bipolar depression," *Personality Social Psychol. Rev.*, vol. 6, no. 4, pp. 380–387, Nov. 2002.

[22] J. J. Gross, Ed., *Emotion Regulation: Conceptual and Empirical Foundations*, 2nd ed. New York, NY, USA: Guilford Press, 2014.

[23] J. Joorman and M. Siemer, "Emotion regulation in mood disorders," in *Handbook of Emotion Regulation*, J. J. Gross, Ed., 2nd ed. New York, NY, USA: Guildford Press, 2013.

[24] A. Gershon, S. L. Johnson, and I. Miller, "Chronic stressors and trauma: Prospective influences on the course of bipolar disorder," *Psychol. Med.*, vol. 43, no. 12, pp. 2583–2592, Dec. 2013.

[25] A. D. Peckham, S. L. Johnson, and I. H. Gotlib, "Attentional bias in euthymic bipolar I disorder," *Cognition Emotion*, vol. 30, no. 3, pp. 472–487, Apr. 2016, doi: 10.1080/02699931.2015.1014313.

[26] J. Gruber, P. Eidelman, S. L. Johnson, B. Smith, and A. G. Harvey, "Hooked on a feeling: Rumination about positive and negative emotion in inter-episode bipolar disorder," *J. Abnormal Psychol.*, vol. 120, no. 4, pp. 956–961, 2011.

[27] J. Linke, C. Sönnekes, and M. Wessa, "Sensitivity to positive and negative feedback in euthymic patients with bipolar I disorder: The last episode makes the difference," *Bipolar Disorders*, vol. 13, nos. 7–8, pp. 638–650, Nov. 2011.

[28] S. L. Johnson, M. D. Edge, M. K. Holmes, and C. S. Carver, "The behavioral activation system and mania," *Annu. Rev. Clin. Psychol.*, vol. 8, no. 1, pp. 243–267, Apr. 2012.

[29] L. B. Alloy *et al.*, "Behavioral approach system and behavioral inhibition system sensitivities and bipolar spectrum disorders: Prospective prediction of bipolar mood episodes," *Bipolar Disorders*, vol. 10, no. 2, pp. 310–322, Mar. 2008.

[30] D. Fulford, S. L. Johnson, M. M. Llabre, and C. S. Carver, "Pushing and coasting in dynamic goal pursuit: Coasting is attenuated in bipolar disorder," *Psychol. Sci.*, vol. 21, no. 7, pp. 1021–1027, Jul. 2010.

[31] S. L. Johnson and D. Fulford, "Preventing mania: A preliminary examination of the GOALS program," *Behav. Therapy*, vol. 40, no. 2, pp. 103–113, Jun. 2009.

[32] R. M. Post, "Transduction of psychosocial stress into the neurobiology of recurrent affective disorder," *Amer. J. Psychiatry*, vol. 149, pp. 999–1010, Aug. 1992.

[33] *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*, 5th ed., Amer. Psychiatric Assoc., Washington, DC, USA, 2013.

[34] I. Gotlib and C. Hammen, Eds., *Handbook of Depression*, 3rd ed. New York, NY, USA: Guilford Press, 2014.

[35] J. Cavanagh, M. Schwannauer, M. Power, and G. M. Goodwin, "A novel scale for measuring mixed states in bipolar disorder," *Clin. Psychol. Psychotherapy*, vol. 16, no. 6, pp. 497–509, Nov. 2009.

[36] C. Henry *et al.*, "Inhibition/activation in bipolar disorder: Validation of the multidimensional assessment of thymic states scale (MAThyS)," *BMC Psychiatry*, vol. 13, no. 1, Dec. 2013, Art. no. 79.

[37] C. Henry *et al.*, "Affective lability and affect intensity as core dimensions of bipolar disorders during euthymic period," *Psychiatry Res.*, vol. 159, nos. 1–2, pp. 1–6, May 2008.

[38] E. Youngstrom and G. P. Algorta, "Features and course of bipolar disorder," in *Handbook of Depression*, I. Gotlib and C. Hammen, Eds., 3rd ed. New York, NY, USA: Guilford Press, 2014.

[39] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophys. J.*, vol. 12, no. 1, pp. 1–24, Jan. 1972.

[40] S. G. Disner, C. G. Beevers, E. A. Haigh, and A. T. Beck, "Neural mechanisms of the cognitive model of depression," *Nature Rev. Neurosci.*, vol. 12, no. 8, p. 467, 2011.

[41] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[42] R. Davidson, "Affective style, psychopathology, and resilience: Brain mechanisms and plasticity," *Amer. Psychol.*, vol. 55, no. 11, p. 1196, 2000.

[43] T. M. Leyro, M. J. Zvolensky, and A. Bernstein, "Distress tolerance and psychopathological symptoms and disorders: A review of the empirical literature among adults," *Psychol. Bull.*, vol. 136, no. 4, pp. 576–600, 2010.

[44] E. L. Hamaker, R. P. P. P. Grasman, and J. H. Kamphuis, "Modeling BAS dysregulation in bipolar disorder: Illustrating the potential of time series analysis," *Assessment*, vol. 23, no. 4, pp. 436–446, Aug. 2016.

[45] M. A. Walsh, A. Royal, L. H. Brown, N. Barrantes-Vidal, and T. R. Kwapil, "Looking for bipolar spectrum psychopathology: Identification and expression in daily life," *Comprehensive Psychiatry*, vol. 53, no. 5, pp. 409–421, Jul. 2012.

Hugo Gonzalez Villasanti (Member, IEEE) received the bachelor's degree in electromechanical engineering from the National University of Asuncion, San Lorenzo, Paraguay, in 2011, and the Ph.D. degree in electrical and computer engineering from The Ohio State University, Columbus, OH, USA, in 2019.

He is currently a Post-Doctoral Scholar with the Crane Center for Early Childhood Research and Policy, The Ohio State University. His current research interests include stability theory and multimedia sensing and actuation for sociotechnological systems.

Kevin M. Passino (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN, USA, in 1989.

He is currently a Professor of electrical and computer engineering and the Director of the Humanitarian Engineering Center, The Ohio State University, Columbus, OH, USA. He is a coeditor of the book, *An Introduction to Intelligent and Autonomous Control* (with P. J. Antsaklis, Kluwer Academic Press, 1993), a coauthor of the books, *Fuzzy Control* (with S. Yurkovich, Addison Wesley Longman Publishing, 1998, *Stability Analysis of Discrete Event Systems* (with K. L. Burgess, NY, USA: John Wiley and Sons, 1998), *The RCS Handbook: Tools for Real Time Control Systems Software Development* (with V. Gazi, M. L. Moore, W. Shackleford, F. Proctor, and J. S. Albus, NY: John Wiley and Sons, 2001), *Stable Adaptive Control and Estimation for Nonlinear Systems: Neural and Fuzzy Approximator Techniques* (with J. T. Spooner, M. Maggiore, and R. Ordonez, NY: John Wiley and Sons, 2002), and *Swarm Stability and Optimization* (with V. Gazi, Heidelberg, Germany: Springer-Verlag, 2011), and the author of *Biomimicry for Optimization, Control, and Automation* (London, U.K.: Springer-Verlag, 2005) and *Humanitarian Engineering: Creating Technologies That Help People* (OH, USA: Bede Pub., second edition, 2015).

Dr. Passino was an Elected Member of the IEEE Control Systems Society Board of Governors. He was the Program Chair of the 2001 IEEE Conf. on Decision and Control. He has served as the Vice-President of Technical Activities of the IEEE Control Systems Society (CSS). He is also a Distinguished Lecturer of the IEEE Society on Social Implications of Technology. For more information, see http://www.ece.osu.edu/p̄assino/.