Brain state in a convex body

S. Hui

Martin Bohner
Missouri University of Science and Technology, bohner@mst.edu

Follow this and additional works at: http://scholarsmine.mst.edu/faculty_work

Part of the Mathematics Commons, and the Statistics and Probability Commons

Recommended Citation
http://scholarsmine.mst.edu/faculty_work/1889

This Article is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. For more information, please contact weaverjr@mst.edu.
Abstract—We study a generalization of the brain-state-in-a-box (BSB) model for a class of nonlinear discrete dynamical systems where we allow the states of the system to lie in an arbitrary convex body. The states of the classical BSB model are restricted to lie in a hypercube. Characterizations of equilibrium points of the system are given using the support function of a convex body. Also, sufficient conditions for a point to be a stable equilibrium point are investigated. Finally we study the system in polytopes. The results in this special case are more precise and have simpler forms than the corresponding results for general convex bodies. The general results give one approach of allowing pixels in image reconstruction to assume more than two values.

I. INTRODUCTION

The brain-state-in-a-box (BSB) neural model was proposed by Anderson and coworkers in 1977 (see [1]). It can be described by the equation

\[ x_{k+1} = g(x_k + W x_k) \]

where \( x_0 \) is an element of the closed \( n \)-dimensional unit hypercube, \( x_k \) is the state of the system at time \( k \), \( W \) some weight matrix, and the function \( g \) ensures that the states of the system are constrained to be in the unit hypercube. The BSB model has been investigated by many researchers, among them Anderson et al. ([11], [5, chapter 4]), Golden [2], Greenberg [3], Grossberg [4], Hui and Zak [7], and Hui et al. [5, chapter 11].

One of the applications of the BSB model is to store patterns in such a way that when presented with a new pattern \( p \), the system responds by finding the stored pattern most closely resembling \( p \). This problem is known as the associative memory problem (see [5] and [6]). One can study the equilibrium points of the system: the points \( e \) such that \( g(e + We) = e \). Of greater interest is the set of all stable equilibrium points, namely, those points \( s \) where there exists an entire neighborhood around \( s \) with \( g(x + Wx) = s \) for all \( x \) in that neighborhood. We can consider the stable equilibrium points of the system described above as the stored patterns. The neighborhood of attraction then contains the noisy versions of the stored pattern \( q \) which should be identified with \( q \). It is useful to choose the extreme points of the hypercube to be the equilibrium points of the system. Hui and Zak ([5, chapter 11], [7]) were able to give conditions on the matrix \( W \) so that this occurs.

Of course the number of stored patterns is restricted to be \( 2^n \) for some natural number \( n \) in this case. The BSB model also only allows the coordinates to assume two values. For example, if one thinks of each coordinate as the value of a pixel in a two-dimensional image, then the BSB model only allows each pixel to be on or off with no possibility of a gray scale. In the present paper we introduce a generalization of the BSB model which can be used to address these problems.

We fix an arbitrary closed, convex, bounded, and nonempty set \( S \) and consider the system described by

\[ x_{k+1} = g(f(x_k)) \]

where \( f \) is any continuous function and \( g \) maps from \( S \) to itself. The precise descriptions are given in Section II. Of course we are interested in the equilibrium points of the system, and to find them, it is useful to introduce the support function of \( S \) and to look at some properties of convex sets and convex functions. Using the support function of \( S \), we will give a characterization of the set of all equilibrium points which yields a necessary and sufficient condition for the statement

all vertices of \( S \) are equilibrium points. (i)

Moreover, if the support function of \( S \) is differentiable at certain points, then it is even easier to check whether a point is an equilibrium point of the system. This is explored in Section III. In Section IV we look at stability of the equilibrium points. First, a sufficient condition for a point to be a stable equilibrium point is given. We can simplify this condition if \( S \) is a polytope and we give conditions which imply

all vertices of \( S \) are stable equilibrium points. (ii)

For the remainder of Section IV, our system is governed by a linear function \( f(x) = x + Wx \), where \( W \) is a weight matrix. We give conditions on \( W \) for (i) and (ii) which are numerically very easy to check. On the other hand, if we would like to have a finite number of fixed points be the equilibrium points, we can choose \( S \) to be the convex hull of those points. This may be one approach to reducing the number of spurious equilibrium points. Also, a pool of matrices \( W \) which will work for (i) and (ii) is given in this section. We can choose from this pool the matrices that are the best for a particular application. In particular, the results of Hui and Zak [7] for \( S = [-1,1]^n \) will be easy consequences of our general theory. In Section V, we indicate how the results can be applied to the gray scale problem and give a numerical example.

II. DEFINITIONS AND BACKGROUND RESULTS

Definition 1: Let \( \mathcal{H} \) be a Hilbert space over \( R \) with dim \( \mathcal{H} = n \in \mathbb{N} \). A closed, convex, bounded, nonempty, and \( n \)-dimensional subset of \( \mathcal{H} \) is called a convex body.

Let \( S \) denote a convex body. Let \( f \) be a continuous function and let \( x_k = f(x_{k-1}) \). We are interested in the restriction of
the system to $S$. Since we want all the points to stay in $S$, we need to send the points which fall out of $S$ back to $S$. To do that, we need the following lemma.

Lemma 1: For each $y \in \mathcal{H}$ there exists a unique $g(y) \in S$ such that

$$
\|y - g(y)\| = \inf_{s \in S} \|y - s\|.
$$

Furthermore $g : \mathcal{H} \to S$ is continuous.

Proof: A proof can be found, for example, in [11, p. 27].

With the “nearest-point-map” $g$ of the above lemma we can define our system.

Definition 2: Let $x_0 \in S$. Define

$$
x_{k+1} := g(f(x_k)) \quad \forall k \in \mathbb{N} \cup \{0\}
$$

where $f : S \to \mathcal{H}$ is continuous and $g : \mathcal{H} \to S$ is the nearest-point-map. ■

That is, if $x_{k+1} \not\in S$ for some $k$, we take it back to the unique point in $S$ which minimizes the distance to $f(x_k)$.

Definition 3: Let $T(x) := g(f(x))$ for $x \in S$ and let $x^* \in S$.

1) If $T(x^*) = x^*$, then $x^*$ is called an equilibrium point of the system. With Equi($S$) we denote the set of all such points.

2) Let $\Delta(x^*, \delta) := \{ s \in \mathcal{H} | \|x^* - s\| < \delta \}$. If there exists $\delta > 0$ so that $T(\mathcal{H} \cap \Delta(x^*, \delta)) = \{x^*\}$, then $x^*$ is called a stable equilibrium point of the system. The set of all stable equilibrium points is referred to as Equi*(S).

In other words, an equilibrium point is stable if there exists a neighborhood of the point so that all the points in that neighborhood are sent to the equilibrium point in one step. Observe also that Equi($S$) $\neq \emptyset$ by a consequence of Brouwer's fixed point theorem.

Before we can give conditions for a point of $S$ to be an equilibrium point, we need some properties of convex sets. To begin with, we define the support function of a convex body.

Definition 4: The function $h : \mathcal{H} \to \mathbb{R}$ defined by

$$
h(u) := \sup_{s \in S}(s, u), \quad u \in \mathcal{H}
$$

called the support function of $S$. For each $u \in \mathcal{H}$, let

$$
H_u := \{ x \in \mathcal{H} | (x, u) \leq h(u) \}.
$$

Of course, $H_u$ is just the half-space containing $S$ determined by the hyperplane that is orthogonal to $u$ and tangent to $S$. Clearly, if $s_0 \in S$ with $h(u) = (s_0, u)$, then $s_0 \in S \cap \partial H_u$ and $\partial H_u = \{ x \in \mathcal{H} | (x, u) = h(u) \}$ is a hyperplane which supports $S$ at $s_0$.

Some properties of the support function are collected in the following lemma.

Lemma 2: Let $S$ be a convex body and $h$ its support function. Then

1) $h$ is subadditive, positively homogeneous, and convex on $\mathcal{H}$.

Proof: The boundedness of $S$ together with the Cauchy–Schwarz inequality imply 1). Everything else can be verified easily.

To become familiar with the support function, we will give two easy examples on how to compute it. We use $\mathcal{X}$, $\partial \mathcal{X}$, and $\mathcal{X}$ to denote the closure, boundary, and interior of a set $\mathcal{X}$, respectively.

Example 1 (Support Function):

1) Let $S = [-1, 1]^n$ be the closed $n$-dimensional unit hypercube in $\mathbb{R}^n$. We can calculate the support function $h$ of $S$ as follows

$$
h(u) = \sup_{s \in S} (s, u) = \sup_{-1 \leq s_i \leq 1} \sum_{i=1}^n s_i u_i = \sum_{i=1}^n \text{sgn}(u_i) u_i
$$

for all $u \in \mathcal{R}^n$.

2) Let $S = \Delta(x_0, k) = \{ s \in \mathbb{R}^n | \|x_0 - s\| \leq k \}$ be the closed $n$-dimensional ball of radius $k$ in $\mathbb{R}^n$ around $x_0$. For $s \in S$, we can find $u \in \Delta(0, k)$ with $s = x_0 + u$. Thus we have

$$
h(u) = \sup_{s \in S} (s, u) = \langle x_0, u \rangle + \sum_{a \in \Delta(0, k)} \text{sgn}(a, u)
$$

where we applied the equality part of the Cauchy–Schwarz inequality.

Since $(x - y, u) = h(u) - h(u) = 0$ for all $x, y \in \partial H_u$, the vector $u$ is normal to the hyperplane $\partial H_u$. Now the geometric meaning of the following definition, where we denote $\{ x + a | a \in \mathcal{A} \}$ for $x \in \mathcal{H}$ and $\mathcal{A} \subset \mathcal{H}$ by $x + \mathcal{A}$, is clear.

Definition 5: Let $x \in \partial S$.

1) $N(x) := x + \{ u \in \mathcal{H} | x \in \mathcal{H} \cap \partial H_u \}$ is called the normal cone of $S$ at $x$.

2) $N^*(x) := x + \{ u \in \mathcal{H} | x \in \mathcal{H} \cap \partial H_u \}$ is called the absolute normal cone of $S$ at $x$.

3) $x$ is said to be a vertex of $S$, if all affine subspaces containing $N(x)$ have dimension $n$. The set of all vertices of $S$ is denoted by Vert($S$).

We illustrate the above definitions with an example.

Example 2: Let $S \subseteq \mathbb{R}^2$ be the region as depicted in Fig. 1. Let $l_a, l_b, l_c, l_d,$ and $l_e$ be line segments and $R_a, R_b$ be open sectors as shown. We have

$$
N(u) = N^*(u) = \{ l_u \}, \\
N(b) = l_b, \\
N^*(b) = \emptyset, \\
N(u) = l_u \cup \tilde{l}_u \cup R_u, \\
N^*(u) = l_u \cup R_u, \\
N(v) = l_v \cup \tilde{l}_v \cup R_v, \\
N^*(v) = (R_v).
$$

Furthermore, we have Vert($S$) $= \{ u, v, w \}$.

Observe that $N(x) = x + \{ u \in \mathcal{H} | (x, u) = h(u) \}$ and that $N(x)$ is the collection of the outward normal vectors to the supporting hyperplanes at $x$. Moreover, it is easy to verify that $N(x)$ is convex for each $x \in \partial S$. In the next section, the normal cone $N(x)$ will be used to give necessary and
sufficient conditions for \( \text{Vert}(S) \subset \text{Equi}(S) \). In Section IV, we will then use the absolute normal cone \( N^*(x) \) to derive sufficient conditions for \( \text{Vert}(S) \subset \text{Equi}^*(S) \). For both results we need the following two theorems which are well-known results about convex sets and convex functions.

**Theorem 1 (Separating and Supporting Properties):**

1) If \( S \) is convex and \( x_0 \in \partial S \), then there exists at least one hyperplane supporting \( S \) at \( x_0 \).
2) If \( A \) and \( B \) are convex with \( A \neq \emptyset \) and \( A \cap B = \emptyset \), then there exists a hyperplane separating \( A \) and \( B \).
3) If \( S \) is a convex body and \( x_0 \not\in S \), then there exists a hyperplane which strictly separates \( \{x_0\} \) and \( S \).

**Proof:** For a proof see, for example, [8, pp. 36, 38, and 41].

**Theorem 2:** Let \( h : \mathcal{H} \to R \) be convex. Then

1) \( h \) is continuous on \( \mathcal{H} \).
2) \( \delta h(x, y) := \lim_{\epsilon \to 0} (h(x + \epsilon y) - h(x))/\epsilon \) exists \( \forall x, y \in \mathcal{H} \).
3) \( \delta h(x, y) = -\delta h(x, -y) \Rightarrow (\delta h/\delta y)(x) \) exists, and
4) if \( (\delta h/\delta y)(x) \) exists for all \( y \in \mathcal{H} \), then \( h \) is differentiable at \( x \) with

\[
\frac{\partial h}{\partial y}(x) = (\nabla h(x), y).
\]

**Proof:** Again we refer the reader to [9, pp. 93, 101].

### III. Equilibrium Points

Now we can return to the model described in Section I. The goal is now to give a necessary and sufficient condition for a point to belong to \( \text{Equi}(S) \). The corollaries of the following two theorems will give necessary and sufficient conditions for \( \text{Vert}(S) \subset \text{Equi}(S) \). They can be considered as the main results of this paper. Before studying the proofs of Theorems 3 and 4, the reader may also first have a look at Example 3 where the stability results concerning the classical BSB model (see [5, chapter 11] and [7]) are derived as easy consequences of our new general theory.

**Theorem 3:** Let \( x \in S \) and \( s \in \partial S \). Then

\[
f(x) \in N(s) \Leftrightarrow T(x) = s.
\]

**Proof:** First, suppose \( f(x) \in N(s) \), i.e., \( \langle s, f(x) - s \rangle = h(f(x) - s) \) and let \( t \in S \). Then with the aid of the Cauchy–Schwarz inequality we find that

\[
||f(x) - t|| \cdot ||f(x) - s|| \geq (f(x) - t, f(x) - s) = (f(x), f(x) - s) - (t, f(x) - s) \geq (f(x), f(x) - s) - h(f(x) - s) = (f(x) - s, f(x) - s) = ||f(x) - s||^2.
\]

Now we have either \( f(x) \in S \) which implies \( 0 \geq ||f(x) - s||^2 \), i.e., \( s = f(x) = g(f(x)) = T(x) \), or \( f(x) \not\in S \), and then

\[
\inf_{t \in S} ||f(x) - t|| \geq ||f(x) - s||^2
\]

so that again (see Lemma 1) \( s = g(f(x)) = T(x) \) holds.

Now suppose \( T(x) = s \). Since \( h(0) = 0 \) we can assume without loss of generality that \( f(x) \not\in S \). We define a new convex set

\[
B := \Delta(f(x), f(x) - s).
\]

Note that \( B \) is the ball around \( f(x) \) which touches \( S \) at the point \( s \in \partial S \). We have \( B \cap S = \emptyset \) since

\[
||f(x) - s|| = \min_{t \in S} ||f(x) - t||.
\]

Thus we can separate those two convex sets by a hyperplane [see Theorem 1-2)], i.e., there exists \( u \in \mathcal{H}\setminus\{0\} \) such that

\[
\langle t, u \rangle \leq h(u) \leq \langle b, u \rangle \quad \forall t \in S, \quad \forall b \in B.
\]

Thus the hyperplane \( \{u \in \mathcal{H} | \langle a, u \rangle = h(u) \} \) supports \( B \) at \( s \in \partial B \). Looking at Example 1-2), where we computed the support function of a ball, we conclude that

\[
\langle s, u^* \rangle = h(u^*) = \langle f(x), u^* \rangle + \|f(x) - s\| \|u^*\|.
\]

We have then

\[
\langle s - f(x), u^* \rangle = \|f(x) - s\| \|u^*\|
\]

and by the equality part of the Cauchy–Schwarz inequality, there exists \( \alpha > 0 \) so that \( u^* = \alpha(s - f(x)) \). Remembering how \( u^* \) was defined and applying Lemma 2-3), we arrive at

\[
\alpha h(f(x) - s) = h(\alpha f(x) - s) = h(-u^*) = h(u) = \langle s, u \rangle = \langle s, \alpha(f(x) - s) \rangle = \alpha(s, f(x) - s),
\]

which yields \( \langle s, f(x) - s \rangle = h(f(x) - s) \) and \( f(x) \in N(s) \).

**Corollary 1:** For \( x_0 \in \partial S \) we have

\[
f(x_0) \in N(x_0) \Leftrightarrow x_0 \in \text{Equi}(S).
\]

Furthermore, the condition

\[
f(x_0) \in N(x_0) \quad \forall x_0 \in \text{Vert}(S)
\]

is necessary and sufficient for \( \text{Vert}(S) \subset \text{Equi}(S) \).

**Proof:** Let \( s = x = x_0 \) in Theorem 3.

**Theorem 4:** Let \( u_0 \in \mathcal{H} \). Then \( h \) is differentiable at \( u_0 \) if and only if there exists \( x_0 \in \partial S \) such that \( x_0 + u_0 \in N^*(x_0) \) and in this case we have \( x_0 = \nabla h(u_0) \).
Proof: Let us assume that \( h \) is differentiable at \( u_0 \). Let \( \mathbf{x}^* \in S \cap \partial H_{u_0} \). Our goal is to show that \( \mathbf{x}^* = \nabla h(u_0) \). For arbitrary \( u \in \mathcal{H} \) and \( \varepsilon > 0 \) we have

\[
\frac{h(u_0 + \varepsilon u) - h(u_0)}{\varepsilon} \geq \mathbf{x}^*, u_0 + \varepsilon u - \mathbf{x}^*, u_0 \quad \text{for all } u \in \mathcal{H}.
\]

Letting \( \varepsilon \) tend to zero from above, we find

\[
\delta h(u_0, u) \geq \langle \mathbf{x}^*, u \rangle \quad \forall \, u \in \mathcal{H}.
\]

Thus we have for all \( u \in \mathcal{H} \)

\[
\langle \mathbf{x}^*, u \rangle = -\langle \mathbf{x}^*, -u \rangle \geq -\delta h(u_0, -u)
\]

and

\[
\delta h(u_0, u) \geq \langle \mathbf{x}^*, u \rangle.
\]

Note that the last equality is a consequence of Theorem 2.3 since \( h \) is differentiable at \( u_0 \). Now it follows by Theorem 2.4 that

\[
\langle \nabla h(u_0), u \rangle = \frac{\partial h}{\partial u}(u_0) = \delta h(u_0, u)
\]

whenever \( u \in \mathcal{H} \). Since the above is true for all \( u \in \mathcal{H} \), we have \( \|\nabla h(u_0) - \mathbf{x}^*\| = 0 \) which shows that \( S \cap \partial H_{u_0} = \{\nabla h(u_0)\} \) holds.

Conversely suppose \( x_0 \in \partial S \) with \( S \cap \partial H_{x_0} = \{x_0\} \). To compute \( \delta h(u_0, u) \) [which exists by Theorem 2-2] for \( u \in \mathcal{H} \) we begin with

\[
\mathbf{h}(x_0 + \varepsilon u) - h(x_0) \geq \varepsilon \mathbf{x}^*, u_0 + \varepsilon u - x_0, u_0 \quad \forall \varepsilon > 0.
\]

So we have \( \delta h(u_0, u) \geq \langle x_0, u \rangle \) \( \forall \, u \in \mathcal{H} \). Now we turn our attention to the opposite inequality. For each \( \varepsilon > 0 \) there exists \( x_0(\varepsilon) \in S \) with

\[
h(x_0 + \varepsilon u) = \mathbf{x}^*, u_0 + \varepsilon u - x_0(\varepsilon), u_0 \quad \forall \varepsilon > 0.
\]

Since \( S \) is bounded, the sequence \( \{x_0(1/n)\}_{n \in \mathbb{N}} \) is bounded also. Therefore, by the Banach–Alaoglu Theorem for Hilbert spaces (see, for example, [10, p. 77]), it contains a weakly convergent subsequence \( \{x_0(1/n_k)\} \), say

\[
\lim_{k \to \infty} x_0(1/n_k) = \mathbf{x}^*, u \quad \forall \, u \in \mathcal{H}.
\]

Also, \( \mathbf{x}^* \in S \), since closed convex sets are weakly closed (see [10, p. 81]). By the definition of \( \{x_0(1/n_k)\} \), we have

\[
\mathbf{h}(x_0 + u/n_k) - h(x_0) \leq \frac{1}{n_k} x_0(1/n_k), u - \frac{1}{n_k} x_0(1/n_k), u_0
\]

and

\[
\mathbf{h}(x_0 + u/n_k) - h(x_0) = \frac{1}{n_k} x_0(1/n_k), u.
\]

Therefore, letting \( k \to \infty \) in the inequality, it follows that

\[
\delta h(u_0, u) \leq \langle \mathbf{x}^*, u \rangle \quad \forall \, u \in \mathcal{H}.
\]

But since \( h \) is continuous [Theorem 2-1], we can write

\[
\mathbf{h}(x_0 + u/n_k) = \langle x_0(1/n_k), u_0 + 1/n_k, x_0(1/n_k), u \rangle
\]

and let \( k \to \infty \) to arrive at

\[
h(x_0) = \langle \mathbf{x}^*, u_0 \rangle + 0\langle \mathbf{x}^*, u \rangle = \langle \mathbf{x}^*, u \rangle.
\]

By assumption, \( x_0 \) is the only element in \( S \) which satisfies the above equality, therefore \( \mathbf{x}^* = x_0 \) and

\[
\delta h(u_0, u) \leq \langle x_0, u \rangle \quad \forall \, u \in \mathcal{H}.
\]

Combining the above inequalities, we conclude that

\[
\delta h(u_0, u) = \langle x_0, u \rangle \quad \forall \, u \in \mathcal{H}.
\]

Applying finally parts 3) and 4) of Theorem 2, we see that the (two-sided) directional derivative \( \langle \partial h/\partial u(u_0) \rangle \) exists for all \( u \in \mathcal{H} \). Thus \( h \) is differentiable at \( u_0 \) and

\[
\langle x_0, u \rangle = \delta h(u_0, u) = \langle \nabla h(x_0), u \rangle \quad \forall \, u \in \mathcal{H}.
\]

Therefore \( \nabla h(x_0) = x_0 \), and the proof is complete.

Theorem 5: Let \( x_0 \in S \) and \( s \in \partial S \). Then the following are equivalent:

1) \( T(x) = s \) and \( h \) is differentiable at \( f(x) - s \),
2) \( h \) is differentiable at \( f(x) - s \) with \( \nabla h(f(x) - s) = s \), and
3) \( f(x) \in N^*(s) \).

Proof: Suppose 1) holds. Then \( f(x) \in N^*(s) \) by Theorem 3 and

\[
\nabla h(f(x) - s) + f(x) - s \in N^*(\nabla h(f(x) - s))
\]

by Theorem 4 which yields

\[
s \in S \cap \partial H_{f(x) - s} = \{\nabla h(f(x) - s)\}.
\]

Therefore 1) implies 2).

Using Theorem 4, we see that 2) implies

\[
f(x) = \nabla h(f(x) - s) + f(x)
\]

and

\[
N^*(s)
\]

and 3) holds.

From 3) it follows that

\[
s + (f(x) - s) = f(x) \in N^*(s) \subseteq (s)\]

holds which implies 1) by Theorem 3 and Theorem 4.

Corollary 2: Let \( x_0 \in \partial S \). Then the following are equivalent:

1) \( x_0 \in \text{Equi}(S) \) and \( h \) is differentiable at \( f(x_0) - x_0 \),
2) \( h \) is differentiable at \( f(x_0) - x_0 \) with \( \nabla h(f(x_0) - x_0) = x_0 \), and
3) \( f(x_0) \in N^*(x_0) \).

Proof: Theorem 5 with \( s = x = x_0 \).

To see how applicable the condition given in Corollary 2 is, we will now give two examples. The first deals with the \( n \)-dimensional hypercube and the second with the \( n \)-dimensional unit ball. The first example contains a derivation of a well-
known result from the study of the BSB model using the
techniques presented above.

Example 3: Let \( S = [-1, 1]^n \) and define
\[
E := \{ c = (c_i)_{1 \leq i \leq n} \mid |c_i| = 1 \quad \forall i \in \{1, \ldots, n\} \}.
\]
The set \( E \) is the collection of all extreme points of \( S \). Let us assume that \( E \subset \text{Equi}(S) \) for \( f(x) = x + Wx + b \), where \( f_i(e) \neq e_i \quad \forall \in E \). Since \( h(u) = \sum_{i=1}^{n} |u_i| \) by Example 1-1), we have
\[
\nabla h(u) = (\text{sgn} u_i)_{1 \leq i \leq n}, \quad u = (u_i)_{1 \leq i \leq n},
\]
\( u_i \neq 0 \quad \forall i \in \{1, \ldots, n\} \).

We have \( E \subset \text{Equi}(S) \) and so the following holds by
Corollary 2 for each \( e = (e_i)_{1 \leq i \leq n} \in E \)
\[
e = \nabla h(f(e) - e) = (\text{sgn}(b_i + (We_i)))_{1 \leq i \leq n} = (\text{sgn}(w_{i1}e_i + b_i + \sum_{j=1, j \neq i}^{n} w_{ij}e_j))_{1 \leq i \leq n} .
\]
By a suitable choice of the vector \( e \) we see that a necessary
condition of the required equation is given by
\[
|w_{ii}| > b_i + \sum_{j=1, j \neq i}^{n} |w_{ij}| \quad \text{and}
\]
\[|w_{ii}| > b_i + \sum_{j=1, j \neq i}^{n} |w_{ij}| \quad \forall i \in \{1, \ldots, n\}
\]
that is
\[
|w_{ii}| > b_i + \sum_{j=1, j \neq i}^{n} |w_{ij}| \quad \forall i \in \{1, \ldots, n\} \quad (S)
\]
We will show later that this condition is also sufficient for the
stability of the vertices.

Observe that a matrix which satisfies condition (S) is
necessarily strongly row dominant, that is
\[
|w_{ii}| > \sum_{j=1, j \neq i}^{n} |w_{ij}| \quad \forall i \in \{1, \ldots, n\}
\]
holds. For some properties of strongly row dominant matrices see for example [7].

Example 4: Let \( S = \Delta(0,1) \). By Example 1-2), we know that
\( h(u) = ||u|| \) for \( u \in \mathcal{H} \). Thus \( h \) is differentiable whenever \( u \subset \mathcal{H} \{0\} \) and we can calculate the partial derivatives
\[
\frac{\partial h}{\partial u}(u) = \frac{1}{2||u||} 2u = \frac{u}{||u||} \quad \forall u \subset \mathcal{H} \{0\}.
\]
Therefore
\[
\nabla h(u) = \frac{u}{||u||} \quad \forall u \subset \mathcal{H} \{0\}.
\]
We conclude that \( h \) is differentiable at \( f(z_0) - x_0 \) provided
\( f(z_0) \neq x_0 \) in this case we have by Corollary 2 that
\[
x_0 \in \text{Equi}(S) \Leftrightarrow \nabla h(f(z_0) - x_0) = x_0
\]
\[
\Rightarrow x_0 = \frac{f(z_0) - x_0}{||f(z_0) - x_0||}
\]
\[
\Rightarrow f(z_0) = x_0 ||f(z_0) - x_0|| + 1.
\]
So we have: \( \text{Equi}(S) = \partial S \Leftrightarrow \exists \alpha \geq 1 \) with \( f(x) = \alpha x \).
Now we compute for each \( j \in \{1, \ldots, m\} \backslash \{i\} \)
\[
\begin{align*}
\h(u_0 + \varepsilon u) &> \h(u_0) - \varepsilon \star = \h(u_0) - \min_{1 \leq j \leq m, j \neq i} \delta_j \\
&+ \rho \max_{1 \leq j \leq m} \|x_j\| \\
&\geq \h(u_0) - \delta_j + \rho \|x_j\| \\
&= (\delta_j, \rho) \\
&\geq (x_j, u_0) + \varepsilon (x_j, u) \\
&= (x_j, u_0 + \varepsilon u).
\end{align*}
\]

Note that \( \|u\| \leq 1 \) and so the last inequality is just the Cauchy–Schwarz inequality. Now take an arbitrary \( z \in S \backslash \{z_i\} \).
Then there exists \( \{\alpha_j\}_{1 \leq j \leq m} \subseteq [0,1] \) with \( \sum_{j=1}^{m} \alpha_j = 1 \), and \( \varepsilon \in \{1, \ldots, m\} \backslash \{i\} \) with \( \alpha_i > 0 \) so that \( x = \sum_{j=1}^{m} \alpha_j x_j \). We have therefore by the above estimate that
\[
\begin{align*}
\langle x, u_0 + \varepsilon u \rangle &= \sum_{j=1}^{m} \alpha_j \langle x_j, u_0 + \varepsilon u \rangle < \sum_{j=1}^{m} \alpha_j \h(u_0 + \varepsilon u) \\
&= \sum_{j=1}^{m} \alpha_j \h(u_0 + \varepsilon u).
\end{align*}
\]

But since equality has to hold for at least one element of \( S \), this element must be \( x_i \) itself and we have immediately
\[
S \cap \partial H_{u_0 + \varepsilon u} = \{x_i\} \quad \forall u \in \Delta(0,1).
\]

Therefore
\[
S \cap \partial H_0 = \{x_i\} \quad \forall u \in \Delta(u_0, \varepsilon).
\]

This shows that \( \Delta(0, \varepsilon) \subseteq N^\ast (x_i) - x_i \), and \( N^\ast (x_i) \) is open.

It is not hard to show that for polytopes \( S \) with minimal representation \( V = \{x_i\}_{1 \leq i \leq m} \) we have \( \text{Vert}(S) = V \). Using the above lemma, we can now give immediately the following corollary (compare also Corollary 1).

**Corollary 3:** Let \( V = \{x_i\}_{1 \leq i \leq m} \) be a minimal representation of the polytope \( S \). Then
\[
f(x_i) \in N^\ast (x_i) \Rightarrow x_i \in \text{Equi}^\ast (S).
\]

Also, a sufficient condition for \( \text{Vert}(S) \subseteq \text{Equi}^\ast (S) \) is
\[
f(x_i) \in N^\ast (x_i) \quad \forall i \in \{1, \ldots, m\}.
\]

**Proof:** This is clear by Theorem 6 and Lemma 3.

**Corollary 4:** If \( h \) is differentiable at \( f(x_i) - x_i \) for all \( i \in \{1, \ldots, m\} \), then
\[
\nabla h(f(x_i) - x_i) = x_i \quad \forall i \in \{1, \ldots, m\}
\]
\[
\Rightarrow \text{Vert}(S) \subseteq \text{Equi}^\ast (S).
\]

**Proof:** This is Corollary 2 with Corollary 3.

A demonstration of the practicality of condition \((A^\ast_p)\) follows.

**Example 5:** Let \( S = [-1,1]^n, f(x) = x + W x + b \). Suppose that \( W \) and \( b \) satisfy condition \((S)\) given in Example 3. Then \( \text{Vert}(S) \subseteq \text{Equi}^\ast (S) \) (This is a result of Hui and Zak from [7]). We have shown in Example 3 that condition \((S)\) is necessary.

To show this assertion recall condition \((S)\)
\[
w_i > |b_i| + \sum_{j=1, j \neq i}^{n} |w_{ij}| \quad \forall i \in \{1, \ldots, n\}, (S)
\]

We have \( V = E \) (see Example 3). Now assume that \( (f(x) - e)_i = 0 \) for some \( e \in \text{Vert}(S) \) and \( i^* \in \{1, \ldots, m\} \). But then we have
\[
w_{i^*} = \frac{b_{i^*}}{e_{i^*}} + \sum_{j=1, j \neq i^*}^{n} \frac{w_{i^*j}}{e_{i^*}} \leq |b_{i^*}| + \sum_{j=1, j \neq i^*}^{n} |w_{i^*j}|
\]
contradicting condition \((S)\). Thus the support function \( h \) is differentiable at any \( e \in \text{Vert}(S) \) with
\[
\nabla h(f(x) - e) = \nabla h(b + We)
\]
\[
= \left( \text{sgn} \left( w_i e_i + b_i + \sum_{j=1, j \neq i}^{n} w_{ij} e_j \right) \right)_{1 \leq i \leq n}.
\]

We now claim that the last expression is equal to \( e \). Note first that by the triangle inequality
\[
-b_i - \sum_{j=1, j \neq i}^{n} |w_{ij}| \leq b_i + \sum_{j=1, j \neq i}^{n} w_{ij} e_j
\]
\[
\leq |b_i| + \sum_{j=1, j \neq i}^{n} |w_{ij}|.
\]

Now, if \( e_i = 1 \), we have by condition \((S)\) and the left part of the above inequality
\[
\text{sgn} \left( w_i e_i + b_i + \sum_{j=1, j \neq i}^{n} w_{ij} e_j \right) = 1 = e_i
\]
and if \( e_i = -1 \), we multiply condition \((S)\) by \((-1)\) and use the right part of the above inequality to obtain
\[
\text{sgn} \left( w_i e_i + b_i + \sum_{j=1, j \neq i}^{n} w_{ij} e_j \right) = -1 = e_i.
\]

This proves the claim and we have
\[
\nabla h(f(x) - e) = e \quad \forall e \in \text{Vert}(S)
\]
and therefore \( \text{Vert}(S) \subseteq \text{Equi}^\ast (S) \) by Corollary 4.

Our last goal is now to give explicit conditions on the matrix \( W \) such that \( \text{Vert}(S) \subseteq \text{Equi}(S) \) or \( \text{Vert}(S) \subseteq \text{Equi}^\ast (S) \) is true if \( S \) is a polytope and if \( f(x) = x + W x \). To find such conditions, we first need to compute the support function of \( S \) in the case when \( S \) is a polytope.

**Lemma 4:** Let \( V = \{x_i\}_{1 \leq i \leq m} \) be a minimal representation of the polytope \( S \).

\[
h(u) = \max_{1 \leq i \leq m} \langle x_i, u \rangle \quad \forall u \in \mathcal{H}.
\]
Proof: A simple calculation shows that for \( u \in H \)
\[
h(u) = \sup_{s \in S} \left\{ \sum_{i=1}^{m} \alpha_i x_i, u \right\} = \sup_{s \in S} \left\{ \sum_{i=1}^{m} \alpha_i x_i \right\} = \sup_{s \in S} \left\{ \sum_{i=1}^{m} \alpha_i = 1, 0 \leq \alpha_i \leq 1 \right\}
\]
\[
= \sup_{1 \leq i \leq m} \left\{ x_i, u \right\} \leq \left( \sum_{i=1}^{m} \alpha_i = 1, 0 \leq \alpha_i \leq 1 \right) \leq \max_{1 \leq i \leq m} \left\{ x_i, u \right\}
\]
\[
= \max_{1 \leq i \leq m} \left\{ x_i, u \right\} \leq h(u).
\]
Thus it follows that \( h(u) = \max_{1 \leq i \leq m} \left\{ x_i, u \right\} \).

With Lemma 4 we can now rewrite the conditions given in the last section. This is done in the following:

**Corollary 5.** Consider the conditions:

\[
(x_j, f(x_j) - x_j) = \max_{i \in \{1, \ldots, m \}} \{ x_i, f(x_j) - x_j \}
\]
\[\forall j \in \{1, \ldots, m \}. \quad (A)
\]
\[
(x_j, f(x_j) - x_j) > \max_{i \in \{1, \ldots, m \}\backslash\{j\}} \{ x_i, f(x_j) - x_j \}
\]
\[\forall j \in \{1, \ldots, m \}. \quad (A'_p)
\]

Then condition (A) is equivalent to \( \text{Vert}(S) \subseteq \text{Equi}(S) \) and condition \( (A'_p) \) implies \( \text{Vert}(S) \subseteq \text{Equi}'(S) \).

**Proof:** Lemma 4 with Corollary 1 and Corollary 3.

Finally, let us consider linear functions of the form \( f(x) = x + Wx \) where \( W \) is a linear operator on \( \mathcal{F} \). In this case we have immediately from the above corollary:

**Corollary 6.** Assume \( f(x) = x + Wx \). Then conditions (A) and \( (A'_p) \) have the form

\[
(x_j, Wx_j) = \max_{i \in \{1, \ldots, m \}} \{ x_i, Wx_j \} \quad \forall j \in \{1, \ldots, m \}. \quad (A)
\]
\[
(x_j, Wx_j) > \max_{i \in \{1, \ldots, m \}\backslash\{j\}} \{ x_i, Wx_j \} \quad \forall j \in \{1, \ldots, m \}. \quad (A'_p)
\]

be the vertices of a regular polytope and let

\[
X = \begin{pmatrix}
0 & 1 & 1 & 1 & 0 & -1 & -1 & 1 & 0 & 1 \\
1.1 & 0 & 1 & 1.1 & 1 & 0 & -1 & -1.1 & -1 & 0 & 1
\end{pmatrix}
\]

\[
= \begin{pmatrix}
14.52 & 14.3 & 12.1 & -12.1 & -14.52 & -14.3 & -12.1 & 12.1 \\
14.3 & 26 & 14.3 & 0 & -14.3 & -26 & -14.3 & 0 \\
1.21 & 14.3 & 14.52 & 12.1 & -12.1 & -14.3 & -14.52 & -12.1 \\
-12.1 & 0 & 12.1 & 22 & 12.1 & 0 & -12.1 & -22 \\
-14.52 & -14.3 & -12.1 & 14.52 & 14.3 & 12.1 & -12.1 \\
-14.3 & -26 & -14.3 & 0 & 14.3 & 26 & 14.3 & 0 \\
-12.1 & -14.3 & -14.52 & -12.1 & 12.1 & 14.3 & 14.52 & 12.1 \\
12.1 & 0 & -12.1 & -22 & -12.1 & 0 & 12.1 & 22
\end{pmatrix}
\]
be the matrix whose columns are $x_1, \ldots, x_8$. With 

$$W = \begin{pmatrix} 1 & 2 & 12 \\ 1 & 12 \end{pmatrix}$$

we can compute $X^T W X$ to be the matrix shown at the bottom of the preceding page. We see that condition

$$\langle x_j, W x_j \rangle > \max_{i \in \{1, \ldots, m\} \setminus \{j\}} \langle x_i, W x_j \rangle \; \forall j \in \{1, \ldots, m\}$$

(A6)

of Corollary 6 is satisfied. Therefore, all vertices are stable equilibrium points of our system.

VI. CONCLUSION

We studied the BSB model on general convex bodies. We gave necessary and sufficient conditions for vertices to be equilibrium points and sufficient conditions for vertices to be stable equilibrium points in the generalized system. These results can be used in the study of associative memory problems as shown in Section V, where we proposed an approach to allow a gray scale in the pixels. The results here also contain as special cases the main results in [3] and [7].

REFERENCES


