Convergence of map seeking circuits

Tomáš Gedeon
Department of Mathematical Sciences
Montana State University
Bozeman, MT 59715
gedeon@math.montana.edu
and
David Arathorn
General Intelligence Corp. and
Center for Computational Biology
Montana State University
dwa@giclab.com

August 7, 2007

Abstract

The map-seeking circuit (MSC) is an explicit biologically-motivated computational mechanism which provides practical solution of problems in object recognition, image registration and stabilization, limb inverse-kinematics and other inverse problems which involve transformation discovery. We formulate this algorithm as discrete dynamical system on a set $\Delta = \Pi_{\ell=1}^{L} \Delta^{(\ell)}$, where each $\Delta^{(\ell)}$ is a compact subset of a nonnegative orthant of $\mathbb{R}^n$, and show that for an open and dense set of initial conditions in Δ the corresponding solutions converge to either a vector with unique nonzero element in each $\Delta^{(\ell)}$ or to a zero vector. The first result implies that the circuit finds a unique best mapping which relates reference pattern to a target pattern; the second result is interpreted as "no match found". These results verify numerically observed behaviour in numerous practical applications.

Key words: Map seeking circuits, computer vision, dynamical systems, Lyapunov function.

1 Introduction

This paper studies the behavior of an algorithm using ideas and methods from dynamical systems theory. The algorithm, called a Map-Seeking Circuit (MSC), was developed by D.Arathorn [1] and has been applied by him to a variety of theoretical and practical problems in biological vision [2, 3, 4] and machine vision [5, 6], inverse kinematics and route planning [2, 7], and in cooperation with other investigators to dynamic image processing [12, 13], and high degree-of-freedom robotic motion control [14]. The MSC algorithm is applicable to a variety of inverse problems that can be posed as transformation-discovery problems, where the goal is to find the best transformation that maps a reference pattern to a target pattern.

The MSC algorithm was motivated by the structure and function of the cortical visual processing streams. A number of visual tasks such as stereo vision, determining shape from motion and recognition of rigid and articulated objects can be posed as transformation-discovery problems and can be readily solved by the MSC algorithm. The solution of more complex problems like object recognition involves a decomposition of an aggregate transformation into a sequence of component transformations. For example, the recognition of a known 3D object in an image which contains other objects (see Figure 1), involves discovering the transformations involved in image formation: the location of the projection of the object in the scene, the magnification of that projection and the orientation angles which produced the particular 2D projection of physical 3D object. For objects whose recognition requires determining interior surface shape rather than just occluding contour, lighting direction becomes an additional factor in the image formation transformations.

For objects which are not rigid, physical articulation or morphing transformations are composed with the image formation transformations and must be discovered in the process of recognition [4, 6].

The composition of transformations applies readily to inverse kinematics, in which the unknown transformations consist of a sequence of projections from the limb root via each limb segment to the target location for the end effector [2, 7, 14]. Similarly, for route finding and/or motion planning, the transformations to be discovered are the sequence of movements which will take the animal or robot from its current location and velocity to the target location and velocity [7, 14].

We illustrate the algorithm on a problem of recognition of a rigid 3D object in Figure 1. The problem, often referred to in machine and biological vision circles as the correspondence problem, is to identify the transformation which maps the model to the projection of the object in the scene, ignoring all the distracting objects in the scene. The model, or a target pattern, in this case is a 3D model of the surface of the pig defined by normal vectors located in space, Figure 1c. The 2D projection of the pig (reference pattern) appears in the input image Figure 1a. The MSC algorithm solves the inverse problem of finding the transformation that takes 2D image to the 3D model. It seeks the unknown transformation as a composition of (1) translation in the image plane, (2) rotation in the image plane, (3) scaling in the image plane, and (4) projections between 2D and 3D parameterized by azimuth and elevation. MSC arrives at a solution by a process of convergence which involves competitive culling of linear combinations (superpositions) of all the possible transformations. A graphical presentation of this process on the inverse of 3D-2D projections is seen in Figure 1d-f. This example ignores occlusion, background noise and image degradation, all of which are dealt with in [4, 5, 6].

The behavior of the convergence of MSC in discovering these transformations, regardless of the application, is the subject of this paper.

2 Results

We now describe the problem in more detail where we follow the exposition in [9]. Denote the reference (input) pattern by I, denote the target (memory) pattern by M, and let I, M lie in $\mathbb{R}^p$. In order to simplify the exposition we consider I and M of the same dimension, but our results extend to the case where I and M have different dimensions. For a particular transformation T in a given class of transformations $\mathcal{T}$ we define the *correspondence* associated with T to be

$$c(T) = \langle T(I), M \rangle, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on $\mathbb{R}^p$.

We assume that each $T \in \mathcal{T}$ is a composition of L maps

$$T = T^{(L)}_{i_L} \circ \ldots \circ T^{(2)}_{i_2} \circ T^{(1)}_{i_1}. \tag{2}$$

For each index ℓ between 1 and L, the maps $T^{(\ell)}_{i_\ell}$, $i_\ell = 1, \ldots, n_\ell$, are taken from a collection of transformation termed a *layer*. The layer terminology reflects the data flow organization of the algorithm and is not intended as an analogy of anatomical "layers" in the visual cortex, but is more likely to correspond to cortical anatomical areas (e.g. V1, V2, etc) in which are believed to implement stages of transformation to the visual signal. We also require each component transformation for layer ℓ, $T^{(\ell)}_{i_\ell}$, to be linear and to be discretely indexed so that $1 \leq i_\ell \leq n_\ell$. While linearity may seem to be a severe restriction, it holds in many important applications. For example, the component transformations in visual pattern recognition—translations, rotations, and rescalings—are each linear. The number of layers L and the number of transformations n_ℓ in each layer are problem specific and are determined by the user of the algorithm. Large number of transformations will potentially provide a better solution, but the computational cost will increase.

The task of maximizing the correspondence then reduces to selecting a particular transformation of the form (2) to maximize (1). Equivalently, one can select the indices $(i_1^*, i_2^*, \ldots, i_L^*)$ which maximize the *correspondence array*,

$$c(i_1, i_2, \ldots, i_L) := \langle T^{(L)}_{i_L} \circ \ldots \circ T^{(2)}_{i_2} \circ T^{(1)}_{i_1}(I), M \rangle. \tag{3}$$

Hence one can solve the correspondence problem simply by constructing the $N := n_1 \cdot n_2 \cdot \ldots \cdot n_L$ components of the L-dimensional array in (3) and then finding its maximum entry. For most of the interesting applications the number of components N is extremely large, so this approach is impractical.

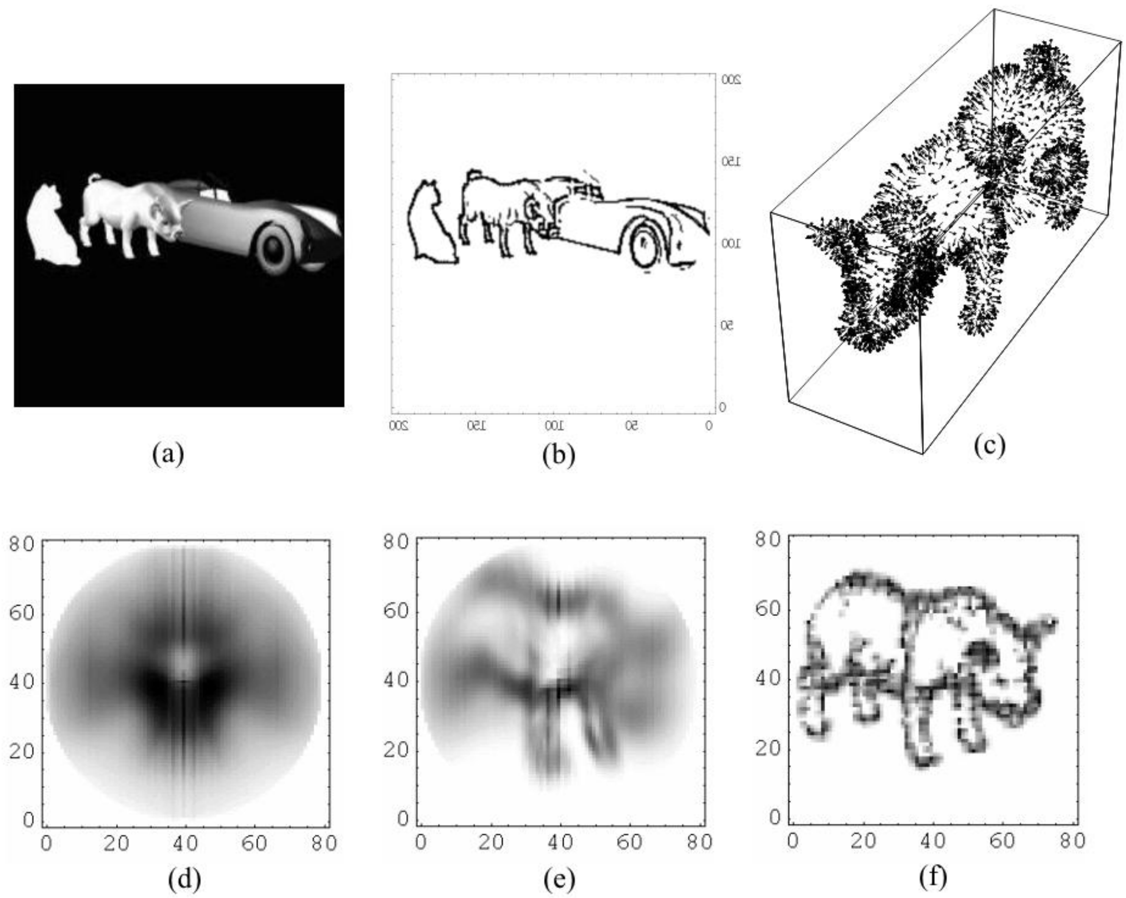

Figure 1: MSC algorithm in image recognition. (a) input image; (b) edge filtered signal into layer 1 forward; (c) 3D surface normal model in memory; (d - f) convergence of superpositions of transformations, iterations 1, 8, and 25. (*Figure taken from [3];3D models courtesy www.3DCafe.com*).

A key idea in [2], which Arathorn refers to as *ordering property of superpositions*, allows the MSC algorithm to perform correspondence maximization iteratively with a cost per iteration that is proportional to the $\sum_{\ell=1}^{L} n_\ell$. The idea is to embed the discretely parameterized linear transformations (2) in a family of continuously parameterized transformations. For each layer ℓ, take

$$T^{(\ell)}_{\mathbf{x}^{(\ell)}} = \sum_{i_\ell=1}^{n_\ell} x^{(\ell)}_{i_\ell} T^{(\ell)}_i, \tag{4}$$

where $x^{(\ell)}_{i_\ell} \leq 1$ are *gain coefficients*. If we replace the individual maps $T^{(\ell)}_{i_\ell}$ in the right hand side of (3) by the linear combinations (4) we obtain the *correspondence function*

$$\begin{aligned} f(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots, \mathbf{x}^{(L)}) &:= \langle T^{(L)}_{\mathbf{x}^{(L)}} \circ \ldots \circ T^{(2)}_{\mathbf{x}^{(2)}} \circ T^{(1)}_{\mathbf{x}^{(1)}}(I), M\rangle \\ &= \sum_{i_1=1}^{n_1} \cdots \sum_{i_L=1}^{n_L} c(i_1, \ldots, i_L) x^{(1)}_{i_1} \ldots x^{(L)}_{i_L}. \end{aligned} \tag{5}$$

The goal is to maximize the function f on the set where $\mathbf{x}^{(\ell)}_i \leq 1$ for all ℓ.

We now describe the Map Seeking Circuit (MSC) algorithm [2]. We set

$$I^{(0)} := I \ \text{ and } \ I^{(\ell)} := T^{(\ell)}_{\mathbf{x}^{(\ell)}}(I^{(\ell-1)}) := \sum_{i=1}^{n_\ell} x_i^{(\ell)} T^{(\ell)}_{i_\ell}(I^{(\ell-1)}),$$

where $I^{(\ell)}$ is a linear combination (superposition) of maps $T^{(\ell)}_{i_\ell}$ applied to the input to the ℓ-th layer $I^{(\ell-1)}$. Similarly, we set

$$M^{(L)} := M \ \text{ and } \ M^{(\ell-1)} := T^{(\ell)*}_{\mathbf{x}^{(\ell)}}(M^{(\ell)}) := \sum_{i=1}^{n_\ell} x_i^{(\ell)} T^{(\ell)*}_{i_\ell}(M^{(\ell)}),$$

where $M^{(\ell)}$ is the backward input to the ℓ-th layer and $T^{(\ell)*}_{i_\ell}$ are Hermitian conjugates of maps $T^{(\ell)}_{i_\ell}$. Therefore

$$I^{(\ell)} \ := \ T^{(\ell)}_{\mathbf{x}^{(\ell)}} \cdots T^{(1)}_{\mathbf{x}^{(1)}}(I), \quad \ell = 1, 2, \dots, L, \tag{6}$$

$$M^{(\ell-1)} \ := \ T^{(\ell)*}_{\mathbf{x}^{(\ell)}} \cdots T^{(L)*}_{\mathbf{x}^{(L)}}(M), \quad \ell = L, ..., 2, 1. \tag{7}$$

Using (6) and (7), the objective function (5) can be expressed for any $\ell = 1, \dots, L$, as

$$f(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(L)}) = \langle I^{(\ell)},\ M^{(\ell)} \rangle = \sum_{i=1}^{n_\ell} x_i^\ell \ \langle T_i^\ell(I^{(\ell-1)}),\ M^{(\ell)} \rangle. \tag{8}$$

The first equality follows by taking adjoints in (5) and substituting (6) and (7). The second equality follows from (4) and the bilinearity of the inner product. Since (8) holds for at every layer ℓ we can dynamically update the coefficients $x_i^{(\ell)}$ synchronously on all layers. First a vector of matches

$$L^{(\ell)} := \Big(\langle T_1^{(\ell)}(I^{(\ell-1)}), M^{(\ell)} \rangle, \langle T_2^{(\ell)}(I^{(\ell-1)}), M^{(\ell)} \rangle, \dots, \langle T_{n_\ell}^{(\ell)}(I^{(\ell-1)}), M^{(\ell)} \rangle \Big) \tag{9}$$

is computed, where $\langle \cdot, \cdot \rangle$ is a dot product. The greatest entry in this vector represents the best match between transformed input and the transformed memory. The weight $x^{(\ell)}_{i_\ell}$ of the map $T^{(\ell)}_{i_\ell}$ that produced the best match should be retained while other weights should be suppressed. Therefore we update the vector of gating coefficients $\mathbf{x}^{(\ell)}$ using a competition function $C(\cdot)$

$$\mathbf{x}^{(\ell)}(n+1) = C^{(\ell)}(\mathbf{x}^{(\ell)}(n), L^{(\ell)}),$$

where the i-th component of $C^{(\ell)}$ is defined by

$$C_i^{(\ell)}(\mathbf{u}, \mathbf{v}) := \begin{cases} max(0, \mathbf{u}_i - \kappa^{(\ell)}(1 - \frac{\mathbf{v}_i}{\max(\mathbf{v})})) & \text{if } \max(\mathbf{v}) \geq \epsilon^{(\ell)} \\ 0 & \text{if } \max(\mathbf{v}) < \epsilon^{(\ell)} \end{cases} \tag{10}$$

and $\max(\mathbf{v})$ is the maximal component of the vector $\mathbf{v}$, see Figure 2. The functions $C^{(\ell)}$ for different ℓ may differ in the choice of the constants $\kappa^{(\ell)}$ and $\epsilon^{(\ell)}$. However, the different choices of $\epsilon^{(\ell)}$ do not significantly affect our argument and thus we simplify our bookkeeping by assuming $\epsilon = \epsilon^{(\ell)}$ for all $\ell = 1, \dots, L$. Observe, that the function $C^{(\ell)}$ preserves the value of the maximal weight $x^{(\ell)}_{i_\ell}$ and lowers other weights $x^{(\ell)}_{j_\ell}, j_{(\ell)} \neq i_{(\ell)}$ towards zero. If these weights are driven below the threshold ϵ without convergence, they are all set to zero.

With the updated gating constants $\mathbf{x}^{(\ell)}(n+1)$ we compute updated values of $I^{(\ell)}$ and $M^{(\ell)}$ in (6) and (7) and iterate the whole process. We take the initial gating constants $x^{(\ell)}_{i_\ell}$ to be equal to a small random perturbation of the value 1.

We now formulate the updates of the algorithm as iterations of a map on a space of all feasible weights $x^{(\ell)}_{i_\ell}$. Let $\Delta^{(\ell)} := \{\mathbf{x}^{(\ell)} \in \mathbb{R}^{n_\ell +} \mid \sum x^{(\ell)}_{i_\ell} \leq n_\ell\}$ and let

$$\Delta = \Delta^{(1)} \times \ldots \times \Delta^{(L)}.$$

The dynamics of each layer is described by

$$\mathbf{x}^{(\ell)}(k+1) := C^{(\ell)}(\mathbf{x}^{(\ell)}(n), L^{(\ell)}(\mathbf{x}^{(1)}(k), \mathbf{x}^{(2)}(k), \dots, \mathbf{x}^{(L)}(k))), \tag{11}$$

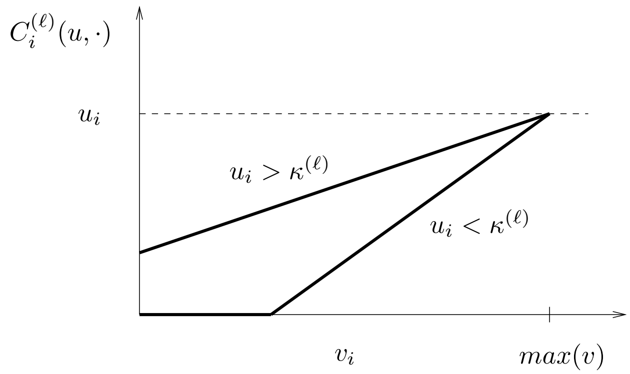

Figure 2: The function $C_i^{(\ell)}(\mathbf{u},\mathbf{v})$ for a fixed value of u_i. The function is linear and intersects the positive v axis when $u_i < \kappa^{(\ell)}$.

where k denotes the iteration number and where $L^{(\ell)} : \mathbb{R}^{n+} \to \mathbb{R}^{n_\ell+}$ is the ℓ-th *layer transfer function* (see (9)), defined by

$$L_i^{(\ell)}(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(L)}) = \langle T_i^{(\ell)}(I^{\ell-1}), M^\ell \rangle. \tag{12}$$

Let $\mathbf{x} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(L)})$ be the concatenation of vectors $\mathbf{x}^{(\ell)} \in \mathbb{R}^{n_\ell+}$, let C be a concatenation of functions $C^{(\ell)}$ and let $L(\mathbf{x})$ denote the collection of layer transfer functions $L^{(\ell)}(\mathbf{x})$. The dynamics of the whole circuit can be then expressed as

$$\mathbf{x}(k+1) := C(\mathbf{x}(k), L(\mathbf{x}(k))) = C(\mathbf{x}(k)), \tag{13}$$

where $C, L : \mathbb{R}^{n+} \to \mathbb{R}^{n+}$ and $n := \sum_{\ell=1}^{L} n_\ell$ is the total number of linear transformations in all L layers.. Let $\mathbf{e}_{i_\ell} \in \mathbb{R}^{n_\ell+}$ be a vector whose $i(\ell)$-th coordinate is 1 and other coordinates are zero. We formulate our main result.

Theorem 2.1 *Consider a discrete dynamical system generated on Δ by (13). Assume that $c(i_1, \ldots, i_L) \geq 0$ and set $c_{min} := \min_{c(\cdot)\neq 0} c(i_1, \ldots, i_L)$, $c_{max} := \max c(i_1, \ldots, i_L)$. Fix a set of constants $\kappa^{(1)}, \ldots, \kappa^{(L)}$ in competition functions $C^{(1)}, \ldots, C^{(L)}$ with the property $\kappa^{(\ell)} \leq (\frac{c_{min}}{c_{max}})^2$ for all ℓ.*

Then for a generic correspondence array $\{c(i_1, \ldots, i_L)\}$ there is an open and dense set $G \subset \Delta$ with the following property. If initial condition $\mathbf{x}(0) \in G$ then iterations $\mathbf{x}(n)$ of (13) converge either to the zero vector or to a vector $(a_1\mathbf{e}_{i_1}^{(1)}, a_2\mathbf{e}_{i_2}^{(2)}, \ldots, a_L\mathbf{e}_{i_L}^{(L)})$, for some positive numbers $a_1, \ldots, a_L$.

The expression "generic correspondence array $c(i_1, \ldots, i_L)$" means that there is an open and dense set in the space of all collections for which our results are true. The necessary condition for being in the generic set is that all elements of the collections are distinct, see Lemma 3.8 and Lemma 3.9, but it may not be sufficient (Lemma 3.10).

The assumptions for the main result are mild and are satisfied in all (known to us) implementations. The condition $c(i_1, \ldots, i_n) \geq 0$ is not very restrictive. Starting with an arbitrary set of coefficients $c(i_1, \ldots, i_L)$ we can satisfy the positivity condition by adding a constant to the nonzero coefficients.

The condition $\kappa^{(\ell)} \leq (\frac{c_{min}}{c_{max}})^2$ is relates the step size of the algorithm $\kappa^{(\ell)}$ to the set of coefficients $c(i_1, \ldots, i_L)$.

Note that the set Δ is a closed subset of $\mathbb{R}^{n+}$ with a non-empty boundary. As the algorithm eliminates weights $x_i^{(\ell)}$ by setting them to zero, it enters the boundary of Δ. We can strengthen the result of Theorem 2.1 to state that the set G is actually open and dense in majority of the boundary subsets of Δ. Since the formulation of this results requires an additional notation, we have delegated its formulation to the section 3 (see Theorem 3.1).

We now outline the argument of the proof. We first characterize the fixed points of the map C and then show that for each ℓ the function $\sum_{i_\ell=1}^{n_\ell} x_{i_\ell}$ is a Lyapunov function for the map $C^{(\ell)}$ (Lemma 3.6). Hence the function $\sum_{\ell=1}^{L} \sum_{i_\ell=1}^{n_\ell} x_{i_\ell}$ is a Lyapunov function for the map C. By LaSalle invariance principle [11] existence of a Lyapunov function in a continuous time dynamical system (i.e a set ordinary differential equations) has strong implications for the character of the omega limit sets of a system. With mild regularity assumptions on the Lyapunov function the omega limit set of each point is in the set of equilibria. For discrete dynamical system, such as studied in this paper, the existence of the Lyapunov function does not apriori rule out periodic points as omega limit sets, provided that the value of the Lyapunov function is constant along such periodic orbit. However, as we show in Corollary 3.7 in our system all omega limit sets lie in the set of fixed points of the map C. Furthermore, each fixed point is either of the form $(a_1\mathbf{e}^{(1)}_{i_1}, \ldots, a_L\mathbf{e}^{(L)}_{i_L})$ for some collection $a_\ell > 0$, or it is a zero solution, or it is an *internal fixed point*. Next we show that for a generic $c(i_1, \ldots, i_L)$ the set of initial conditions that converge to internal fixed points is nowhere dense. The primary mathematical difficulties in proving the last result stem from two facts. The first is that the correspondence function (5) is multi-linear and thus the fixed points of the map are not isolated, but form L dimensional families. Further, the map itself is not invertible. The key results addressing these issues are Lemma 3.10 and Theorem 3.14, respectively.

Finally we comment on the relationship of the nonzero solution to which MSC algorithm converges and solution of the maximization problem (5). By Theorem 4.1 [9] all the local maxima of (5) on the smaller space $\sum_{i_\ell=1}^{n_\ell} x^{(\ell)}_i = 1$ for all ℓ, are solutions of the form $(\mathbf{e}^{(1)}_{i_1}, \ldots, \mathbf{e}^{(L)}_{i_L})$ where $\mathbf{e}^{(j)}_{i_j}$ is a unit vector in direction i_j in the j-th layer. In other words, in each layer there is precisely one weight $x^{(j)}_{i_j} = 1$ and the others are zero. The MSC algorithm generically converges to a point $(a_1\mathbf{e}^{(1)}_{i_1}, \ldots, a_L\mathbf{e}^{(L)}_{i_L})$ for some collection $a_\ell > 0$, or to a zero solution. There is an obvious one-to-one correspondence between the vectors of this form and vectors $(\mathbf{e}^{(1)}_{i_1}, \ldots, \mathbf{e}^{(L)}_{i_L})$. The advantage of setting up the problem on a space $\sum_{i_\ell} x^l_{i_\ell} \leq 1$ for all i rather then on $\sum_{i_\ell} x^{(\ell)}_{i_\ell} = 1$ for all ℓ, is the possibility of the "no match found" outcome.

We close this section by providing details about the application of the map-seeking circuit algorithm illustrated in Figure 1. The memory in this test consists of a 3 dimensional surface normal model of one of the objects presented in the image. The implementation uses $L = 4$ layers. These layers, in sequential order along the forward pathway, comprise a full set (120×120) of translations in increments of one pixel ($n_1 = 14,400$), a set of scalings spanning ratios from 0.7 to 1.4 times the linear dimensions of the input image in steps of factor 1.025 ($n_2 = 29$), rotations in the viewing plane from $-30°$ to $30°$ by increments of 1 degree ($n_3 = 61$). In the fourth layer the forward transformations are line-of-view projections of the 2D space into the 3D space of the model and in the backward pathway orthographic 3D projections of the model into 2D corresponding to all viewpoints from $-90°$ to $+90°$ in azimuth and $0°$ to $90°$ of elevation in $5°$ degree increments ($n_4 = 703$). The input image is converted into an edge filtered representation and the data on the forward path remains in the 2D image domain until layer 4 where it is projected into the 3D model space in order to locate corresponding normals perpendicular to the projected line of view. On the backward path the model normals perpendicular to each line of view are projected to form a collection of edge rendered 2D views from different angles, and these form the superposition on the backward path. Hidden edges are suppressed where normals are not perpendicular to the line of view. In this way, proximal non-tangent surfaces suppress distal tangent surfaces which otherwise would have produced edges in the 2D projection. The total number of transformations ($n_1 + n_2 + n_3 + n_4$) implemented are $15,193$. These in composition ($n_1n_2n_3n_4$) comprise 1.79×10^{10} possible aggregate transformations.

The circuit converges very quickly as can be seen in Figure 1, having found an approximate solution by iteration 8 and fully converged by iteration 25.

3 Proof of the main result

The section is organized as follows. We first carefully define boundary subsets of Δ and formulate a stronger version of Theorem 2.1. In section 4.1 we characterize the fixed points of the map C and in section 4.2 we find a Lyapunov function for the system. The key result of this section is Corollary 3.7, which shows that all solutions either converge to a point $(a_1\mathbf{e}^{(1)}_{i_1}, \ldots, a_L\mathbf{e}^{(L)}_{i_L})$ for some collection $a_\ell > 0$, to a zero solution, or to an *internal fixed point*. The sections 4.3 and 4.4 are devoted to an argument showing that for a generic MSC the set of initial conditions that converge to internal fixed points is nowhere dense.

The set Δ is a closed subset of $\mathbb{R}^{n+}$ with a non-empty boundary. As the algorithm eliminates weights $x^{(\ell)}_i$ by setting them to zero, it enters the boundary of Δ, which we now describe. We define for each layer

ℓ a non-empty collection of integers $\omega^{(\ell)} = (i_1, \ldots, i_{q(\ell)})$, $q(\ell) \geq 1$, $i_j \in \{1, \ldots, n_\ell\}$. For any such $\omega^{(\ell)}$ we denote by $\mathbb{R}^{n_\ell +}_{\omega}$, the boundary part of $\mathbb{R}^{n_\ell +}$ consisting of vectors of the form $(0, u_1, 0, \ldots, u_{q(\ell)}, 0)$, where the nonzero elements are in positions $i_1, \ldots, i_{q(\ell)}$. We set

$$\Delta^{(\ell)}_{\omega} := \mathbb{R}^{n_\ell +}_{\omega} \cap \Delta^{(\ell)}.$$

Similarly, for $\Omega = (\omega^{(1)}, \ldots, \omega^{(L)})$ we denote $\mathbb{R}^{n+}_{\Omega}$ the boundary part of $\mathbb{R}^{n+}$ consisting of vectors of the form $(\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(L)})$ such that $\mathbf{u}^{(\ell)} = (0, u^{(\ell)}_1, 0, \ldots, u^{(\ell)}_{q(\ell)}, 0)$, where the nonzero elements of $\mathbf{u}^{(\ell)}$ are in positions specified by $\omega^{(\ell)} = (i^{(\ell)}_1, \ldots, i^{(\ell)}_{q(\ell)})$. We set

$$\Delta_\Omega := \mathbb{R}^{n+}_{\Omega} \cap \Delta = \Pi^L_{\ell=1} \Delta^{(\ell)}_{\omega}.$$

Let $q = \sum_{\ell=1}^L q(\ell)$ be the total sum of nonzero components in $\mathbb{R}^{n+}_{\Omega}$. We define the set

$$\Xi := \{\Omega = (\omega^{(1)}, \ldots, \omega^{(L)}) \mid \text{ there are at least two layers } \ell, r \text{ with } |\omega^{(\ell)}| \geq 2, |\omega^{(r)}| \geq 2\}. \tag{14}$$

Note that the space $\mathbb{R}^{n+}_{\Omega}$ with $\Omega \notin \Xi$ is a product of half-lines and at most one set isomorphic to the positive cone of $\mathbb{R}^2$. We now formulate an extension of Theorem 2.1

Theorem 3.1 *Assume all assumptions of Theorem 2.1.*

Then for a generic correspondence array $\{c(i_1, \ldots, i_L)\}$ there is an open and dense set $G \subset \Delta$, which is also open and dense in every boundary set Δ_Ω for all $\Omega \in \Xi$, with the following property. If initial condition $x(0) \in G$ then iterations $\mathbf{x}(n)$ of (13) converge either to the zero vector or to a vector $(a_1 \mathbf{e}^{(1)}_{i_1}, a_2 \mathbf{e}^{(2)}_{i_2}, \ldots, a_L \mathbf{e}^{(L)}_{i_L})$, for some positive numbers $a_1, \ldots, a_L$.

3.1 Fixed points of C

We start with a technical Lemma that provides the bridge between the MSC algorithm and the correspondence function f in (5). Fix $\Omega \in \Xi$ and let L_Ω and f_Ω be restrictions of the functions L and f to the set $\mathbb{R}^{n+}_{\Omega(q)}$.

Lemma 3.2 ([9, 8]) *The function L_Ω is the gradient of the cost function f_Ω*

$$\begin{aligned} &[L^{(\ell)}_\Omega]_i(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(L)}) = [\nabla^{(\ell)} f_\Omega]_i \\ = &\sum_{i_1 \in \omega^1} \cdots \sum_{i_{\ell-1} \in \omega^{\ell-1}} \sum_{i_{\ell+1} \in \omega^{\ell+1}} \cdots \sum_{i_L \in \omega^L} c(i_1, \ldots, i_{\ell-1}, i, i_{\ell+1} \ldots i_L) x^{(1)}_{i_1} \ldots x^{(\ell-1)}_{i_{\ell-1}} x^{(\ell+1)}_{i_{\ell+1}} \ldots x^{(L)}_L \end{aligned} \tag{15}$$

Further,

$$\left[L^{(\ell)}_\Omega(\mathbf{x})\right]_i > 0 \quad \textit{if} \quad i \in \omega^{(\ell)} \;\; \textit{and} \;\; \left[L^{(\ell)}_\Omega(\mathbf{x})\right]_i = 0 \quad \textit{if} \quad i \notin \omega^{(\ell)}.$$

Proof. Differentiating (8) we obtain the components of the gradient:

$$\frac{\partial f}{\partial x^{(\ell)}_i} = \langle T^{(\ell)}_i(I^{(\ell-1)}), \quad M^{(\ell)} \rangle = [L^{(\ell)}]_i(\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(L)}). \tag{16}$$

Restriction to the subset Ω finishes the first result. To show the second part we observe that the right hand side of (15) is positive since all $x^{(\ell)}_{i_\ell} > 0$ and the coefficients $c(i_1, \ldots, i_L) > 0$. The result follows. □

Lemma 3.3 *A point $\mathbf{x} \in \mathbb{R}^{n+}_{\Omega}$, $\mathbf{x} \neq \mathbf{0}$ is a fixed point of the map (13), if, and only if,*

$$L(\mathbf{x}) = (a_1 \mathbf{1}^{(1)}_\omega, \ldots, (a_\ell \mathbf{1}^{(\ell)}_\omega, \ldots, a_L \mathbf{1}^{(L)}_\omega),$$

where $\mathbf{1}^{(\ell)}_\omega := \mathbf{e}^{(\ell)}_{i_1} + \mathbf{e}^{(\ell)}_{i_2} + \ldots + \mathbf{e}^{(\ell)}_{i_q}$ $i_j \in \omega^{(\ell)}$, be the vector of 1's in all directions in $\omega^{(\ell)}$.

In particular, every point of the form $\mathbf{e} = (a_1 \mathbf{e}^{(1)}_{i_1}, a_2 \mathbf{e}^{(2)}_{i_2}, \ldots, a_L \mathbf{e}^{(L)}_{i_L})$ for any positive constants a_i, is a fixed point of the map C.

Proof. Take $\mathbf{x} \in \mathbb{R}^{n+}_{\Omega}$ with $\mathbf{x} \neq \mathbf{0}$. Then $\mathbf{x}$ is a fixed point if it satisfies $\mathbf{x} = C(\mathbf{x}, L(\mathbf{x}))$. From the form of the competition functions $C^{(\ell)}$ follows that for all ℓ we must have $L^{(\ell)}_i(\mathbf{x}) = \max L^{(\ell)}(\mathbf{x}) =: K^{(\ell)}$ for all i where $L^{(\ell)}_i(\mathbf{x}) > 0$. Thus $L^{(\ell)}(\mathbf{x})$ has the form above with $a_\ell = K^{(\ell)}$.

To show the second part take $\Omega = (\omega^{(1)}, \ldots, \omega^{(L)})$ and each $\omega^{(\ell)} = \{m_\ell\}$ contains exactly one element. Then a nonzero $\mathbf{x} \in \mathbb{R}_\Omega$ has the form $\mathbf{e} = (a_1 \mathbf{e}^{(1)}_{i_1}, \ldots, a_L \mathbf{e}^{(L)}_{i_L})$. Applying (15) to such $\mathbf{e}$ we get

$$[L^{(\ell)}_\Omega]_i(\mathbf{e}) = (\Pi_{i \neq \ell} a_i) c(m_1, \ldots, m_L) \text{ for } i = m_\ell \quad \text{and} \quad [L^{(\ell)}_\Omega]_i(\mathbf{e}) = 0 \text{ if } i \neq m_\ell.$$

Since for each ℓ the vector $L^{(\ell)}_\Omega$ has a single non-zero element, its maximum is achieved at such element and $K^{(\ell)} = (\Pi_{i \neq \ell} a_i) c(m_1, \ldots, m_L)$. By the part one of this Lemma $\mathbf{e}$ is a fixed point of C. □

Lemma 3.4 *Let $\mathbf{u} = C(\mathbf{x})$. Let $\zeta^{(\ell)}(\mathbf{x}) := \{i \,|\, \mathbf{x}^{(\ell)}_i \neq 0\}$ be the set of indices of nonzero elements of $\mathbf{x}^{(\ell)}$ and let $\eta^{(\ell)}(\mathbf{x}) := \{i \,|\, C^{(\ell)}_i(\mathbf{x}) \neq 0\}$ be the set of indices of nonzero elements of $C^{(\ell)}(\mathbf{x})$. Then for all $\ell = 1, \ldots, L$*

$$\eta^{(\ell)}(\mathbf{x}) \subset \zeta^{(\ell)}(\mathbf{x})$$

.

Proof. We will consider each part of the function C separately. Recall that $K^{(\ell)} = \max L^{(\ell)}(\mathbf{x})$. Fix ℓ and rewrite the i-th component of the function $C^{(\ell)}$ as

$$\begin{aligned} C^{(\ell)}_i(\mathbf{x}^{(\ell)}, L^{(\ell)}(\mathbf{x}))) &= \max(0, x^{(\ell)}_i - \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}})) \\ &= \begin{cases} 0 & \text{if } x^{(\ell)}_i \leq \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}}) \\ x^{(\ell)}_i - \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}}) & \text{if } x^{(\ell)}_i \geq \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}}) \end{cases} . \end{aligned} \tag{17}$$

Assume that $x^{(\ell)}_i = 0$. Then $\kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}}) \geq 0$, since $K^{(\ell)} = \max L^{(\ell)}(\mathbf{x})$. Since $x^{(\ell)}_i = 0$, this implies $x^{(\ell)}_i \leq \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}})$. Therefore by (17) we have $u^{(\ell)}_i = C^{(\ell)}_i(\mathbf{x}) = 0$. □

We can rewrite the previous result as $\zeta^{(\ell)}(C\mathbf{x}) \subset \zeta^{(\ell)}(\mathbf{x})$ which is equivalent to the statement $x^{(\ell)}_i = 0$ implies $C^{(\ell)}_i(\mathbf{x}) = 0$ for all $\ell = 1, \ldots, L$. This yields the following Corollary.

Corollary 3.5 *For any Ω the boundary set $\mathbb{R}^{n+}_\Omega$ is positively invariant under the map C.*

3.2 The Lyapunov function

The key result is a construction of a Lyapunov function [10, 11]. The following Lemma establishes that the l_1 norm $|\mathbf{x}| := \sum_{\ell=1}^{n} \sum_{i_\ell=1}^{n_\ell} x^{(\ell)}_{i\ell}$ of the vector $\mathbf{x}$ is a Lyapunov function for the map C.

Lemma 3.6 *If $\mathbf{u} = (\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(L)}) \in \mathbb{R}^{n+}$ then*

1. *$|C(\mathbf{u}, L(\mathbf{u}))| \leq |\mathbf{u}|$, where $|\mathbf{v}| = \sum_i \mathbf{v}_i$ denotes the sum of the elements of the vector $\mathbf{v}$;*
2. *The previous inequality is strict, unless $\mathbf{u}$ is a fixed point of C.*

Proof. Observe that by (17) the set $\eta^{(\ell)}(\mathbf{x})$ is the set of indices where $x^{(\ell)}_i \geq \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}})$.

We compute the sum of elements of the vector $C^{(\ell)}(\mathbf{x})$:

$$\begin{aligned} |C^{(\ell)}(\mathbf{x}^{(\ell)}, L^{(\ell)}(\mathbf{x}))| &= \sum_{i \in \eta^{(\ell)}(\mathbf{x})} x^{(\ell)}_i - \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}}) \\ &\leq \sum_{i \in \eta^{(\ell)}(\mathbf{x})} x^{(\ell)}_i \\ &\leq \sum_{i \in \zeta^{(\ell)}(\mathbf{x})} x^{(\ell)}_i \\ &= |\mathbf{x}^{(\ell)}|. \end{aligned}$$

Here the first inequality holds because the maximum of rescaled vector $\frac{L^{(\ell)}(\mathbf{x})}{K^{(\ell)}}$ is 1. The second inequality follows from Lemma 3.4. Since $\mathbf{x}$ is a concatenation of vectors $\mathbf{x}^{(\ell)}$ and $|\mathbf{x}| = |\mathbf{x}^{(1)}| + \ldots |\mathbf{x}^{(\ell)}|$ by the definition of $|\cdot|$, this proves the first part of the Lemma.

The equality $|C(\mathbf{x}, L(\mathbf{x})| = |\mathbf{x}|$ happens when for all $\ell = 1, \ldots, L$, both of the inequalities above are in fact equalities. For a fixed ℓ the equality $|C^{(\ell)}(\mathbf{x}, L^{(\ell)}(\mathbf{x}))| = |\mathbf{x}^{(\ell)}|$ implies that, first, for all $i \in \eta^{(\ell)}(\mathbf{x}^{(\ell)})$ we have $[L^{(\ell)}(\mathbf{x})]_i = K^{(\ell)}$, and, second, that

$$\zeta^{(\ell)}(\mathbf{x}) = \eta^{(\ell)}(\mathbf{x}).$$

Since this holds for every $\ell = 1, \ldots, L$, $\mathbf{x}$ is a fixed point of C by Lemma 3.3. □

Corollary 3.7 *1. For all Ω the set Δ_Ω is positively invariant (i.e. $C(\Delta_\Omega) \subset \Delta_\Omega$) under the map (13).*

2. There is an integer N such that for any initial condition $\mathbf{x}(0) \in \Delta$, the N-th iterate $\mathbf{x}(N)$ is either

- **(a)** $\mathbf{x}(N) = \mathbf{0}$*; or*
- **(b)** $\mathbf{x}(N) = \mathbf{e} = (a_1\mathbf{e}^{(1)}_{i_1}, \ldots, a_L\mathbf{e}^{(L)}_{i_L})$ *for some collection of $a_i > 0$; or*
- **(c)** *the trajectory $\{\mathbf{x}(k)\}_{k=1}^{\infty} \to \mathbf{u}$ where $L(\mathbf{u}) = (a_1\mathbf{1}^{(1)}_\omega, \ldots, a_L\mathbf{1}^{(1)}_\omega)$ is a fixed point with at least one ℓ with $|\omega^{(\ell)}| \geq 2$.*

Proof. The first statement is a corollary of Lemma 3.6 and Corollary 3.5.

To show the second part, observe that the number of nonzero components $\zeta^{(\ell)}(\mathbf{x}(k))$ is a non-increasing function of the iteration number k by Corollary 3.5. Since this function has also discrete set of values, it must be eventually constant. Let $N(\mathbf{x})$ be such that for all $k \geq N$ the number of nonzero components $\zeta^{(\ell)}(\mathbf{x}(k))$ of $\mathbf{x}(k)$ is constant. Since $N(\mathbf{x})$ depends continuously on $\mathbf{x}$ and Δ is compact, there exists a uniform N valid for all $\mathbf{x} \in \Delta$. If $\zeta^{(\ell)}(\mathbf{x}(N))$ has a single component for all ℓ then by Corollary 3.5 $\mathbf{x}(N) = \mathbf{e}$, which satisfies (b). It follows from the form of f (see 5) and Lemma 3.2 that if there is an ℓ such that $|\zeta^{(\ell)}(\mathbf{x}(N))| = 0$ then $\mathbf{x}(N) = \mathbf{0}$.

Finally, assume there is an ℓ such that $|\zeta^{(\ell)}(\mathbf{x}(N))| = s \geq 2$ and $|\zeta^{(i)}(\mathbf{x}(N))| \geq 1$ for $i \neq \ell$. Then by construction of the number N we have $|\zeta^{(\ell)}(\mathbf{x}(k))| = s$ for all $k \geq N$. Since the Lyapunov function is bounded below by zero, we must have that $\mathbf{x}(k) \to \mathbf{u}$ and $|C(\mathbf{u}, L(\mathbf{u})| = |\mathbf{u}|$. Further, by continuity we have $|\zeta^{(\ell)}(\mathbf{u})| = s \geq 2$. By Lemma 3.6.2 $\mathbf{x}(n)$ converges to a fixed point which by Lemma 3.3 has the advertised form. □

3.3 The outline of the argument

In this brief section we outline the rest of the argument.

As a first step we precisely characterize the set of the internal fixed points Fix (see (21) below). Because the correspondence function is multi-linear internal fixed points are not isolated. However we show in Lemma 3.9) that set Fix is nowhere dense in each $\mathbb{R}^{n+}_\Omega$ with $\Omega \in \Xi$.

The next step is to show that the set of all points that converge to Fix is also nowhere dense. The main obstacle is that the map C is not one-to-one: in the neighbourhood of the boundary it maps multiple points to the same point on the boundary, see (17). Therefore we divide the argument into two parts, see Figure 3. We first define the set W_Ω which is the set of all points in a particular $\mathbb{R}^{n+}_\Omega$ that converge to the internal fixed points in the same $\mathbb{R}^{n+}_\Omega$. Restriction of C to such $\mathbb{R}^{n+}_\Omega$ is invertible and we show that generically there is an eigenvalue of the linearization at every internal fixed point with modulus larger then 1. Using the stable manifold theory we conclude that W_Ω is nowhere dense in $\mathbb{R}^{n+}_\Omega$ (Lemma 3.10).

To start the second part of the argument we set

$$W := \bigcup_{\Omega \in \Xi} W_\Omega, \tag{18}$$

be the collection of all stable manifolds W_Ω of all internal fixed points $\mathbf{u} \in \Delta_\Omega$, and define for $k = 1, 2, \ldots$

$$X_k := \{\mathbf{x} \in \mathbb{R}^{n+} \mid C^k(\mathbf{x}) \in W\} \tag{19}$$

be the set of points which map after k iterates to some stable manifold of an internal fixed point. Here we have to face the non-uniqueness of the map C since it collapses entire intervals in X_1 onto points in W. In spite of this we show in Theorem 3.14 that given an arbitrary nowhere dense set D in $\mathbb{R}^{n+}$ the (generalized) inverse $C^{-1}(D)$ is also nowhere dense $\mathbb{R}^{n+}$.

In the final section 3.6 we use the Theorem 3.14 inductively to show that X_k is a nowhere dense set for each k and thus the set

$$U_k := \mathbb{R}^{n+} \setminus X_k$$

is open and dense for each k. Therefore for N specified by Lemma 3.7 the set

$$G := \bigcap_{k=1}^{N} U_N \cap \Delta$$

is an open and dense set of initial conditions, which converge to either to $\mathbf{0}$ or to a vector $\mathbf{e}$. This will conclude the proof of Theorem 3.1

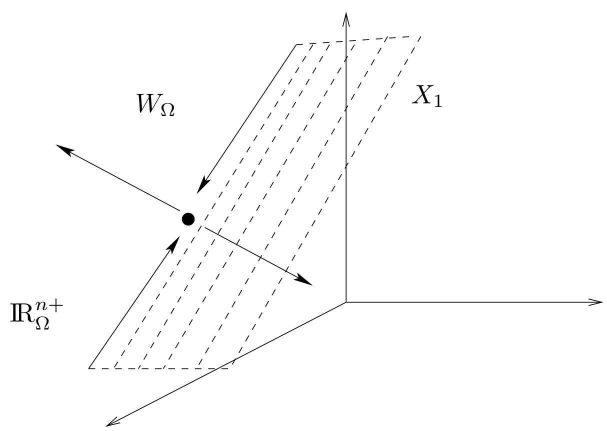

Figure 3: The function C is invertible on $\mathbb{R}^{n+}_\Omega$ and the stable manifold W_Ω can be constructed. On the other hand C maps the dashed set onto W_Ω in one iteration. This set is part of X_1.

3.4 Internal fixed points and their stable sets

As a consequence of Corollary 3.7, to prove Theorem 2.1 we need to show that there is an open and dense set W of initial conditions $\mathbf{x}(0)$, such that the iterations $\mathbf{x}(n)$ do not converge to a fixed point $\mathbf{x}$ satisfying (c) of the Corollary 3.7. Then the proof of Theorem 2.1 will follow from Lemma 3.6 and Corollary 3.7.

The fixed points $\mathbf{u}$ of C which satisfy condition (c) above will be called *internal fixed points*, since there must be at least one layer ℓ where $\mathbf{u}^{(\ell)}$ is in the interior of $\Delta_{\omega^{(\ell)}}$. We now look more closely at these internal fixed points. Recall that $\omega^{(\ell)} = (i_1, \ldots, i_{q(\ell)})$, $i_j \in \{1, \ldots, n_\ell\}$ is a non-empty collection of integers and $\Omega = (\omega^{(1)}, \ldots, \omega^{(L)})$ is a collection of $\omega^{(\ell)}$. Let

$$Z_\Omega := \{\mathbf{x} \in \mathbb{R}^{n+}_\Omega \mid \mathbf{x} = (a_1 \mathbf{1}_{\omega^{(1)}}, a_2 \mathbf{1}_{\omega^{(2)}}, \ldots, a_L \mathbf{1}_{\omega^{(L)}}), a_1, \ldots, a_L > 0\}$$

and

$$B_\Omega = \{\mathbf{x} \in \mathbb{R}^{n+} \mid L(\mathbf{x}) \in Z_\Omega\}.$$

In view of Lemma 3.3, B_Ω is the set of fixed points of C. The next Lemma justifies the definition of the class Ξ of Ω's (see 14), since only the boundary sets Δ_Ω with $\Omega \in \Xi$ may contain internal fixed points.

Lemma 3.8 *Assume that the collection $c(i_1, \ldots, i_L)$ has distinct elements. If $\Omega = (\omega^{(1)}, \ldots, \omega^{(L)})$ has a unique ω^ℓ such that $|\omega^{(\ell)}| = 2$ and $|\omega^{(i)}| = 1$ for all $i \neq \ell$, then $B_\Omega = \emptyset$.*

Proof. By assumption there is a unique layer ℓ such that $|\omega^{(\ell)}| = 2$. Assume without loss that $\omega^{(\ell)} = \{1,2\}$. Since $L(\mathbf{x}) \in Z_\Omega$ implies $L_1^{(\ell)}(\mathbf{x}) = a_\ell = L_2^{(\ell)}(\mathbf{x})$, by (15) we get

$$\begin{aligned} L_1^{(\ell)}(\mathbf{x}) &= c(m_1, m_2, \dots, 1, \dots, m_L) x^{(1)}_{m_1} x^{(2)}_{m_2} \dots \hat{\mathbf{x}}^{(\ell)} \dots x^{(L)}_{m_L} \\ &= c(m_1, m_2, \dots, 2, \dots, m_L) x^{(1)}_{m_1} x^{(2)}_{m_2} \dots \hat{\mathbf{x}}^{(\ell)} \dots x^{(L)}_{m_L} = L_2^{(\ell)}(\mathbf{x}) \end{aligned}$$

where notation $\hat{\mathbf{x}}^{(\ell)}$ indicates that there are no $x_i^{(\ell)}$ in the expression. This implies $c(m_1, m_2, \dots, 1, \dots, m_L) = c(m_1, m_2, \dots, 2, \dots, m_L)$, contradicting our assumption. $\square$

Let

$$Z := \bigcup_{\Omega \in \Xi} Z_\Omega, \tag{20}$$

where the union is over the collection of all $\Omega \in \Xi$. By Lemma 3.3 and Corollary 3.7 the set

$$Fix := \{\mathbf{x} \in \mathbb{R}^{n+} \mid L(\mathbf{x}) \in Z\} \tag{21}$$

is the set of internal fixed points of the map C.

Lemma 3.9 *If the collection $c(i_1, \dots, i_L)$ has distinct elements, then the set Fix is closed and nowhere dense in $\mathbb{R}^{n+}$ and the intersection $Fix \cap \mathbb{R}^{n+}_\Omega$ is closed, nowhere dense subset of $\mathbb{R}^{n+}_\Omega$, for every $\Omega \in \Xi$.*

Proof. We first observe that since the function L is continuous and Z and $\mathbb{R}^{n+}_\Omega$ are closed, the set $Fix \cap \mathbb{R}^{n+}_\Omega$ is closed for each Ω.

We now prove the density of the complement of Fix in every $\mathbb{R}^{n+}_\Omega$ with $\Omega \in \Xi$. Fix Ω and assume, contrary to our assertion, that there is an open set $D \subset \mathbb{R}^{n+}_\Omega$ such that $L(D) \subset Z$. By definition of Z this means that there exists Ω' with $\Omega' \in \Xi$ such that $L(D) \subset Z_{\Omega'}$, that is, for all $\ell = 1, \dots, L$,

$$L^{(\ell)}(D) \subset Z^{(\ell)}_{\omega^{(\ell)\prime}}.$$

Choose ℓ with $|\omega^{(\ell)\prime}| \geq 2$. Then there must exist two coordinates $i, j \in \{1, \dots, n_\ell\}$ such that

$$L_i^{(\ell)}(D) = L_j^{(\ell)}(D).$$

By (15) this is equivalent to

$$\sum_{\omega^{(k)} \neq \omega^{(\ell)}} c(i_1, \dots, i, \dots, i_L) x^{(1)}_{i_1} \dots x^{(L)}_{i_L} = \sum_{\omega^{(k)} \neq \omega^{(\ell)}} c(i_1, \dots, j, \dots, i_L) x^{(1)}_{i_1} \dots x^{(L)}_{i_L}$$

for all $\mathbf{x} \in D$. Since D is open, we have $c(i_1, \dots, i, \dots, i_L) = c(i_1, \dots, j, \dots, i_L)$ for all $i_k \in \omega^{(k)}$ with $\omega^{(k)} \neq \omega^{(\ell)}$. This contradicts our assumption and finishes the proof of the Lemma. $\square$

Now we show that every internal fixed point $\mathbf{x} \in Fix$ is unstable, i.e. it has at least one eigenvalue with modulus greater then 1. Notice, that since the function L has the form described in (15), it has the following scaling property. If $\mathbf{x} = (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(L)})$ and $\mathbf{y} = (b_1 \mathbf{x}^{(1)}, \dots, b_L \mathbf{x}^{(L)})$ then

$$L^{(\ell)}(\mathbf{y}) = (\Pi_{j \neq \ell} b_j)\, L^{(\ell)}(\mathbf{x})$$

for all ℓ. We call $\lambda := (b_1, \dots, b_L)$, where all $b_j > 0$, a *multi-scaling factor* and write $\mathbf{y} = \lambda \mathbf{x}$. With this notation, if $\mathbf{x}$ is an internal fixed point that satisfies $L(\mathbf{x}) = (a_1 \mathbf{1}_{\omega^{(1)}}, \dots, a_L \mathbf{1}_{\omega^{(L)}})$ then

$$L(\lambda \mathbf{x}) = (a_1 (\Pi_{j \neq 1} b_j) \mathbf{1}_{\omega^{(1)}}, \dots, a_L (\Pi_{j \neq L} b_j) \mathbf{1}_{\omega^{(L)}})$$

and hence it is again an internal fixed point. Let

$$\bar{\mathbf{u}} := \{\mathbf{v} \in \Delta \mid \mathbf{v} = \lambda \mathbf{u} \text{ for some multi-scaling factor } \lambda\}$$

be the set of all fixed points related to $\mathbf{u}$ by scaling. We see that each internal fixed point belongs to an L dimensional cone-like space of fixed points. Therefore a linearization at each internal fixed point has eigenvalue 1 with a multiplicity at least L. However, for any $\Omega \in \Xi$ the set $\Delta_\Omega \subset \mathbb{R}^{n+}_\Omega$ has dimension $q \geq L+2$. Therefore for a generic $c(i_1,\ldots,i_L)$ the set $Fix \cap \mathbb{R}^{n+}_\Omega$ should have codimension 2. The proof of the next Lemma is based on the fact that for a generic collection $\{c(i_1,\ldots,i_L)\}$ there is always at least one eigenvalue of the linearization of C at an internal fixed point with the modulus greater then 1, and one eigenvalue with the modulus smaller then 1.

Lemma 3.10 *There exists an open and dense set of coefficients $c(i_1,\ldots,i_L)$ such that for all $\Omega \in \Xi$ and all internal fixed points $\mathbf{u} \in Fix \cap \mathbb{R}^{n+}_\Omega$, satisfying $L(\mathbf{u}) = (a_1\mathbf{1}_{\omega^{(1)}},\ldots,a_L\mathbf{1}_{\omega^{(L)}})$ with $\Omega = (\omega^{(1)},\ldots,\omega^{(L)})$, the set*

$$W_\Omega(\bar{\mathbf{u}}) := \{\mathbf{x} \in \mathbb{R}^{n+}_\Omega \mid \lim_{k\to\infty} C^k(\mathbf{x}) = \lambda\mathbf{u} \text{ for some } \lambda\} \tag{22}$$

is closed and nowhere dense in Δ_Ω.

Proof. Take a point $\mathbf{u} \in Fix$ with $L(\mathbf{u}) = (a_1\mathbf{1}_{\omega^{(1)}},\ldots,a_L\mathbf{1}_{\omega^{(L)}})$, set $\Omega = (\omega^{(1)},\ldots,\omega^{(L)})$, and assume a sequence of iterates $\{\mathbf{x}(n)\}_{n=1}^\infty \in \mathbb{R}^{n+}_\Omega$, $\mathbf{x}(k+1) = C(\mathbf{x}(k)) = C^k(\mathbf{x})$, converges to $\mathbf{u}$. We assume without loss of generality that $\Omega = (\omega^{(1)},\ldots,\omega^{(L)})$ where $\omega^{(\ell)} = (1,2,3,\ldots,q(\ell))$ for all $\ell = 1,\ldots,L$. To each point $\mathbf{x}(k)$ we can assign a collection $\alpha(k) = (m_1(k), m_2(k),\ldots,m_L(k))$, $m_\ell(k) \in \omega^{(\ell)}$, with the property that $L^{(\ell)}_i(\mathbf{x}(n))$ has the maximal element $L^{(\ell)}_{m_\ell}(\mathbf{x}(n))$. Since the number of distinct collections α is finite, there is a subsequence $\mathbf{x}(k_n)$ with $k_n \to \infty$ such that $\alpha(\mathbf{x}(k_n))$ is constant. We rename the subsequence to be again $\{\mathbf{x}(k)\}$.

By Corollary 3.5 the space $\mathbb{R}^{n+}_\Omega$ is positively invariant. Let $C_\Omega : \mathbb{R}^{n+}_\Omega \to \mathbb{R}^{n+}_\Omega$ be the restriction of the map C to $\mathbb{R}^{n+}_\Omega$ defined by

$$\mathbf{u}^{(\ell)}_i := \mathcal{C}^{(\ell)}_{\Omega,i}(\mathbf{x}) = X^{(\ell)}_i + \kappa^{(\ell)}\left(\frac{[L^{(\ell)}(X)]_i}{K^{(\ell)}} - 1\right),$$

where $X := (X^{(1)},\ldots,X^{(L)})$, $X^{(\ell)} = (x^{(\ell)}_1,\ldots,x^{(\ell)}_{q(\ell)},0,0,\ldots,0)$, and $K^{(\ell)}$ is the maximal element of the vector $L^{(\ell)}(\mathbf{x})$. Similarly, we will denote by $\mathcal{L}^{(\ell)}_\Omega$ the restriction of $\mathcal{L}$ to $\mathbb{R}^{n+}_\Omega$.

Since $\mathbf{x}(n) \to \mathbf{u}$ the sequence of unit vectors

$$\mathbf{v}(n) := \frac{C(\mathbf{x}(n+1)) - C(\mathbf{x}(n))}{|C(\mathbf{x}(n+1)) - C(\mathbf{x}(n))|}$$

converges to an eigenvector $\mathbf{v}$ of the derivative matrix $\frac{d\mathcal{C}_\Omega}{dx}(\mathbf{u})$ with the corresponding eigenvalue with the modulus less or equal to 1. The derivative matrix $\frac{d\mathcal{C}_\Omega}{dx}(\mathbf{u})$ is a $q\times q$ matrix of the form $I + A$, where I is the $q\times q$ identity matrix and A is a block matrix with $l\times l$ blocks, where (ℓ,s)-block, $\ell = 1,\ldots,L$, $s = 1,\ldots,L$, has the size $q(\ell)\times q(s)$. The (i,t) element of the (ℓ,s) block of A is

$$[A(\mathbf{u})]^{(\ell),(s)}_{i,t} = \frac{\kappa^{(\ell)}}{(\mathcal{L}^{(\ell)}_{m_\ell})^2}\left(\frac{\partial\mathcal{L}^{(\ell)}_i}{\partial x^s_t}\mathcal{L}^{(\ell)}_{m_\ell} - \frac{\partial\mathcal{L}^{(\ell)}_{m_\ell}}{\partial x^s_t}\mathcal{L}^{(\ell)}_i\right), \tag{23}$$

for all $1 \leq i \leq q(\ell)$ and $1 \leq t \leq q(s)$. Notice that this is well defined since the sequence of vectors $L^{(\ell)}(\mathbf{x}(n))$ has the same maximal element $L^{(\ell)}_{m_\ell}(\mathbf{x}(n))$ for all ℓ. By (15) each (ℓ,ℓ) block of the matrix A is zero. The trace of $I + A$ is therefore $q = \sum_{\ell=1}^L q_\ell$, which is the sum of all eigenvalues. Since there are q eigenvalues, either all eigenvalues are equal to 1, or there is a pair of eigenvalues λ_1,λ_2 with $|\lambda_1| > 1$ and $|\lambda_2| < 1$. Thus all we need to show is that for all $\Omega \in \Xi$, not all eigenvalues of $\frac{d\mathcal{C}_\Omega}{dx}(\mathbf{x})$ are equal to 1. Assume to the contrary, that all eigenvalues of $I + A$ are equal to 1. Then by the Jordan normal form it follows that A is nilpotent, i.e. there exists a power N such that A^N is the zero matrix. For any $\Omega \in \Xi$ consider the corresponding matrix $A = A_\Omega(\mathbf{u})$, where we emphasize the dependence of the matrix A on both Ω and the internal fixed point $\mathbf{u}$. The proof of the Lemma will be complete if we prove the following claim, since it implies that the matrix $A_\Omega(\mathbf{u})$ is not nilpotent and thus not all eigenvalues of $I + A$ are on the unit circle.

Claim 3.11 *For an open and dense set of coefficients $c(i_1,\ldots,i_L)$ there is a nonzero diagonal element $a^{ii}_\Omega(\mathbf{u})$ of the matrix $A^2_\Omega(\mathbf{u})$.*

Before we prove the claim, observe that

1. each diagonal (ℓ, ℓ) block of A_Ω is zero; and
2. for each ℓ the $m(\ell)$-th row is zero by the formula (23).

This second fact implies that for each ℓ with $|\omega_\ell| = 1$, all corresponding blocks (ℓ, s) for all s are zero. Observe, that this implies that $I + A$ has at least L eigenvalues 1. These correspond to the directions along the family of internal fixed points related by a multi-factor scaling. This opens a real possibility that the matrix A_Ω may be nilpotent. Define $\bar{A}_\Omega$ to be the matrix which has 1 in each position of A, that is different then positions forced to be zero by (1) and (2) above. The following result shows that the form of the matrix A_Ω for $\Omega \in \Xi$ is compatible with A_Ω being not nilpotent.

Claim 3.12 *For all $\Omega \in \Xi$, there is a diagonal element of the matrix $\bar{A}^2$ that is non-zero.*

Proof. Since $\Omega \in \Xi$, there are ℓ and s such that ω_ℓ and ω_s have at least two elements. Recall that m_ℓ denotes the index of the maximal element of $\mathcal{L}^{(\ell)}$, that is

$$\mathcal{L}^{(\ell)}_{m_\ell} = \max_i \mathcal{L}^{(\ell)}_i.$$

Take $i \neq m_\ell$ and $j \neq m_s$ and consider $[\bar{a}^{ii}_\Omega]^{(\ell),(\ell)}$ the (i,i)-th element of the matrix $\bar{A}^2$ in the (ℓ,ℓ) block. Since both the (i,j) element of the (ℓ, s) block and the (j,i) element of the (s,ℓ) block of $\bar{A}$ are 1, $[\bar{a}^{ii}_\Omega]^{(\ell),(\ell)} \neq 0$.
□

Proof of Claim 3.11 We consider the term $[a^{ii}_\Omega]^{(\ell),(\ell)}$ at the same position as the non-zero term $[\bar{a}^{ii}_\Omega]^{(\ell),(\ell)}$ in the previous Claim. To simplify notation we will use a^{ii}_Ω to denote this term. We start with a formula for a^{ii}_Ω which follows from (23):

$$\begin{aligned} a^{ii}_\Omega(\mathbf{u}) &= \sum_{s,t} \frac{\kappa^{(\ell)}}{(\mathcal{L}^{(\ell)}_{m_\ell})^2} \frac{\kappa^{(s)}}{(\mathcal{L}^{(s)}_{m_s})^2} \left(\frac{\partial \mathcal{L}^{(\ell)}_i}{\partial x^s_t} \mathcal{L}^{(\ell)}_{m_\ell} - \frac{\partial \mathcal{L}^{(\ell)}_{m_\ell}}{\partial x^s_t} \mathcal{L}^{(\ell)}_i \right) \left(\frac{\partial \mathcal{L}^{(s)}_t}{\partial x^\ell_i} \mathcal{L}^{(s)}_{m_s} - \frac{\partial \mathcal{L}^{(s)}_{m_s}}{\partial x^\ell_i} \mathcal{L}^{(s)}_t \right) \\ &= \sum_{s \neq \ell, t} \frac{\kappa^{(\ell)} \kappa^{(s)}}{\mathcal{L}^{(\ell)}_{m_\ell} \mathcal{L}^{(s)}_{m_s}} \left(\frac{\partial \mathcal{L}^{(\ell)}_i}{\partial x^s_t} - \frac{\partial \mathcal{L}^{(\ell)}_{m_\ell}}{\partial x^s_t} \right) \left(\frac{\partial \mathcal{L}^{(s)}_t}{\partial x^\ell_i} - \frac{\partial \mathcal{L}^{(s)}_{m_s}}{\partial x^\ell_i} \right) \end{aligned} \tag{24}$$

where the second equality follows from the assumption that $L(\mathbf{u}) = (a_1 \mathbf{1}_{\omega^{(1)}}, \ldots, a_L \mathbf{1}_{\omega^{(L)}})$ and thus $\mathcal{L}^{(\ell)}_{m_\ell} = \mathcal{L}^{(\ell)}_i$ and $\mathcal{L}^{(s)}_{m_\ell} = \mathcal{L}^{(s)}_i$, whenever these values are nonzero. We compute the functions in (24) using (15)

$$\begin{aligned} \mathcal{L}^{(\ell)}_{m_\ell}(\mathbf{x}) &= \sum_{\omega^{(j)} \neq \omega^{(\ell)}} c(i_1, \ldots, i_{\ell-1}, m_\ell, i_{\ell+1} \ldots i_L) x^{(1)}_{i_1} \ldots \hat{\mathbf{x}}^{(\ell)} \ldots x^{(L)}_{i_L} \qquad (25) \\ \frac{\partial \mathcal{L}^{(\ell)}_i}{\partial x^s_t}(\mathbf{x}) &= \sum_{\omega^{(j)} \neq \{\omega^{(\ell)}, \omega^{(s)}\}} c(i_1, \ldots, i_{\ell-1}, i, i_{\ell+1}, \ldots, i_{s-1}, t, i_{s+1}, \ldots, i_L) x^{(1)}_{i_1} \ldots \hat{\mathbf{x}}^{(\ell)} \hat{\mathbf{x}}^{(s)} \ldots x^{(L)}_{i_L}, \\ \mathcal{L}^{(\ell)}_i(\mathbf{x}) &= \sum_{\omega^{(j)} \neq \omega^{(\ell)}} c(i_1, \ldots, i_{\ell-1}, i, i_{\ell+1} \ldots i_L) x^{(1)}_{i_1} \ldots \hat{\mathbf{x}}^{(\ell)} \ldots x^{(L)}_{i_L}, \\ \frac{\partial \mathcal{L}^{(\ell)}_{m_\ell}}{\partial x^s_t}(\mathbf{x}) &= \sum_{\omega^{(j)} \neq \{\omega^{(\ell)}, \omega^{(s)}\}} c(i_1, \ldots, i_{\ell-1}, m_\ell, i_{\ell+1} \ldots i_{s-1}, t, i_{s+1} \ldots i_L) x^{(1)}_{i_1} \ldots \hat{\mathbf{x}}^{(\ell)} \hat{\mathbf{x}}^{(s)} \ldots x^{(L)}_{i_L} \end{aligned}$$

where we use notation $\hat{\mathbf{x}}^{(\ell)}$ to denote the fact that the variables $\mathbf{x}^{(\ell)}$ are missing in a given expression.

We multiply all elements $c(i_1, \ldots, c_\ell, \ldots, i_L)$ with $c_\ell \neq m_\ell$ by a constant b and observe how the function $a^{ii}_\Omega(\mathbf{u}, b)$ behaves under such scaling. Since $a^{ii}_\Omega(\mathbf{u}, b)$ is an analytic function of b, it is either identically zero, or, except for a finite number of exceptional values of b, we have $a^{ii}_\Omega(\mathbf{u}, b) \neq 0$. Observe that the functions

$\mathcal{L}^{(\ell)}_{m_\ell}$ and $\frac{\partial \mathcal{L}^{(\ell)}_{m_\ell}}{\partial \mathbf{x}^s_t}$ do not contain $c(i_1,\ldots,c_\ell,\ldots,i_L)$ with $c_\ell \neq m_\ell$. On the other hand every summand in functions $\mathcal{L}^{(s)}_{m_s}$, $\frac{\partial \mathcal{L}^{(s)}_t}{\partial \mathbf{x}^\ell_i}$ and $\frac{\partial \mathcal{L}^{(s)}_{m_s}}{\partial \mathbf{x}^\ell_i}$ is being scaled by b. We write (24) as

$$a^{ii}_\Omega(\mathbf{u},b) = \sum_{s\neq \ell,t} \kappa^{(\ell)}\kappa^{(s)} \left(\frac{\frac{\partial \mathcal{L}^{(\ell)}_i}{\partial x^s_t} - \frac{\partial \mathcal{L}^{(\ell)}_{m_\ell}}{\partial x^s_t}}{\mathcal{L}^{(\ell)}_{m_\ell}} \right) \left(\frac{\frac{\partial \mathcal{L}^{(s)}_t}{\partial x^\ell_i} - \frac{\partial \mathcal{L}^{(s)}_{m_s}}{\partial x^\ell_i}}{\mathcal{L}^{(s)}_{m_s}} \right).$$

Under the above scaling, the second term remains unchanged since both the numerator and the denominator are scaled by b, while in the first term only one term in the numerator is scaled

$$\frac{b\frac{\partial \mathcal{L}^{(\ell)}_i}{\partial x^s_t} - \frac{\partial \mathcal{L}^{(\ell)}_{m_\ell)}}{\partial x^s_t}}{\mathcal{L}^{(\ell)}_{m_\ell}}.$$

Since $\frac{\partial \mathcal{L}^{(\ell)}_i}{\partial x^s_t} \neq 0$ by the same argument as in Lemma 3.2, for $b >> 1$ we get that $a^{ii}_\Omega(\mathbf{u},b) \neq 0$. Therefore there is an open and dense set of values of b for which for which $a^{ii}_\Omega(\mathbf{u},b) \neq 0$. Since the function L is multi-linear, we also have $a^{ii}_\Omega(\lambda\mathbf{u},b) \neq 0$ for any multi-scaling factor λ and the same b. Furthermore by continuity and for a fixed b, there is an open set of $\mathbf{y}$ in a neighbourhood $N(\mathbf{u})$ of $\mathbf{u}$ such that $a^{ii}_\Omega(\mathbf{y},b) \neq 0$. Since Δ_Ω is compact, there is a finite cover by such neighbourhoods and thus there is an open and dense set V_Ω of b such that if $b \in V_\Omega$ then $a^{ii}_\Omega(\mathbf{u},b) \neq 0$ for all $\mathbf{u}$ internal fixed points in Δ_Ω. If we repeat the same argument for all $\Omega \in \Xi$ we get a set

$$V := \bigcap_{\Omega\in\Xi} V_\Omega$$

with the property that if $\{c(i_1,\ldots,i_L)\} \in V$ then $a^{ii}_\Omega(\mathbf{u},b) \neq 0$ for all $\Omega \in \Xi$ and all internal fixed points $\mathbf{u} \in \Delta_\Omega$. Since the collection Ξ is finite, the set V is open and dense. $\square$

The Claim 3.11 implies that the matrix $A_\Omega(\mathbf{u})$ is not nilpotent and thus not all eigenvalues of $I + A$ are equal to 1 for all $c(i_1,\ldots,i_L)$ in an open and dense set U. This finishes the proof of Lemma 3.10. $\square$

3.5 The competition map $C(\mathbf{x})$ and its inverse

As outlined in the section 3.3 the next step is to show that the sets X_k (see (19)) are nowhere dense. The major problem is that the map C is not one-to-one: in the neighbourhood of the boundary it maps multiple points to the same point on the boundary, see (17). Thus we need to closely investigate the map $C(\mathbf{x})$ and its inverse.

We first investigate the inverse of the map C on $\mathbb{R}^{n+}_\Omega$. We assume without loss of generality that $\Omega = (\omega^{(1)},\ldots,\omega^{(L)})$ where $\omega^{(\ell)} = (1,2,3,\ldots,q(\ell))$ for all $\ell = 1,\ldots,L$. We fix $\mathbf{u} = (\mathbf{u}^{(1)},\ldots,\mathbf{u}^L) \in \mathbb{R}^{n+}_\Omega$, where $\mathbf{u}^{(\ell)} = (u^{(\ell)}_1,\ldots,u^{(\ell)}_{q(\ell)},0,\ldots,0)$ with $u^{(\ell)}_i > 0$. In order to solve for (the set of) $\mathbf{x}$ such that $\mathbf{u} = C(\mathbf{x})$ we have to solve

$$\begin{aligned} u^{(\ell)}_i &= x^{(\ell)}_i - \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}}) \quad \text{for } i = 1,\ldots,q(\ell) \\ 0 &\geq x^{(\ell)}_i - \kappa^{(\ell)}(1 - \frac{[L^{(\ell)}(\mathbf{x})]_i}{K^{(\ell)}}) \quad \text{for } i = q(\ell)+1,\ldots,n_\ell \end{aligned}$$

for all $\ell = 1,\ldots,L$. This can be rewritten

$$\begin{aligned} u^{(\ell)}_i + \kappa^{(\ell)} &= x^{(\ell)}_i + \frac{\kappa^{(\ell)}}{K^{(\ell)}}[L^{(\ell)}(\mathbf{x})]_i \quad \text{for } i = 1,\ldots,q(\ell) \\ \kappa^{(\ell)} &\geq x^{(\ell)}_i + \frac{\kappa^{(\ell)}}{K^{(\ell)}}[L^{(\ell)}(\mathbf{x})]_i \quad \text{for } i = q(\ell)+1,\ldots,n_\ell \end{aligned} \tag{26}$$

The second set of equations in (26) demonstrate clearly that the map $C(\mathbf{x})$ is not one-to-one. If the first set of the equations in (26) can be inverted by

$$\mathbf{x}^{(\ell)} = \varphi_\Omega^{(\ell)}(\mathbf{u}^{(\ell)}) \quad \text{for all } \ell = 1, \ldots, L,$$

then all solutions of (26) lie in the set

$$S(\mathbf{u}) = \{\mathbf{y} \in \mathbb{R}^{n+} \mid \mathbf{y}_i^{(\ell)} = \varphi_\Omega^{(\ell)}(\mathbf{u}^{(\ell)}) \text{ for } i \in \omega^{(\ell)}, \mathbf{y}_i^{(\ell)} \leq \kappa^{(\ell)} \text{ for } i \notin \omega^{(\ell)}\}. \tag{27}$$

The key observation is that if $\mathbf{u} \in D$, an arbitrary set that is nowhere dense in $\mathbb{R}_\Omega^{n+}$ for all $\Omega \in \Xi$, then the set

$$S(D) := \bigcup_{\mathbf{u} \in D} S(\mathbf{u}) \tag{28}$$

is also nowhere dense in $\mathbb{R}_\Omega^{n+}$ for all $\Omega \in \Xi$.

To show this we first turn our attention to invertibility of the first set of equations in (26)

$$u_i^{(\ell)} + \kappa^{(\ell)} = x_i^{(\ell)} + \frac{\kappa^{(\ell)}}{K^{(\ell)}}[L^{(\ell)}(\mathbf{x})]_i \quad \text{for } i = 1, \ldots, q(\ell),\ \ell = 1, \ldots, L. \tag{29}$$

We note that in (29) the expression $L^{(\ell)}(\mathbf{x})$ involves $x_{q(\ell)+1}^{(\ell)}, \ldots, x_{n_L}^{(L)}$. If we restrict our search to $\mathbf{x} \in \mathbb{R}_\Omega^{n+}$, i.e. to those $\mathbf{x}$ with $x_i^{(\ell)} = 0$ for all $i > q(\ell)$, then the set of equations (29) defines a function

$$\mathcal{C}_\Omega : \mathbb{R}_\Omega^{n+} \to \mathbb{R}_\Omega^{n+} \tag{30}$$

by

$$u_i^{(\ell)} := \mathcal{C}_{\Omega,i}^{(\ell)}(\mathbf{x}) = X_i^{(\ell)} + \kappa^{(\ell)}(\frac{[L^{(\ell)}(X)]_i}{K^{(\ell)}} - 1),$$

where $X := (X^{(1)}, \ldots, X^{(L)})$ and $X^{(\ell)} = (x_1^{(\ell)}, \ldots, x_{q(\ell)}^{(\ell)}, 0, 0, \ldots, 0)$.

Lemma 3.13 *If $\kappa^{(\ell)} \leq \frac{c_{min}^2}{c_{max}^2}$ for all ℓ, then the maps $C_\Omega(\mathbf{x})$ are invertible as functions from $\mathbb{R}_\Omega^{n+}$ to $\mathbb{R}_\Omega^{n+}$ for all $\Omega \in \Xi$.*

Proof. We fix $\Omega \in \Xi$. We have computed the derivative matrix $\frac{d\mathcal{C}_\Omega}{dx}(\mathbf{x})$ in (23) and its components in (25). Since all values of $\mathbf{x}_{i_\ell}^{(\ell)} \leq 1$ we can estimate

$$[A(\mathbf{x})]_{i,t}^{\ell,s} \leq \kappa^{(\ell)} \frac{c_{max}^2}{c_{min}^2} \leq 1.$$

Therefore

$$det(\frac{d\mathcal{C}_\Omega}{dx}(\mathbf{x})) = det(I + A(\mathbf{x})) \neq 0.$$

$\square$

Let $\varphi_\Omega : \mathbb{R}_\Omega^{n+} \to \mathbb{R}_\Omega^{n+}$ denote the inverse of C_Ω. The following Theorem addresses the non-uniqueness of the inverse to $C(\mathbf{x})$.

Theorem 3.14 *Let $D \subset \mathbb{R}^{n+}$ be a nowhere dense closed set in $\mathbb{R}_\Omega^{n+}$ such that $D \cap \mathbb{R}_\Omega^{n+}$ is nowhere dense and closed in $\mathbb{R}_\Omega^{n+}$ for all $\Omega \in \Xi$. Then the set*

$$C^{-1}(D) := \{\mathbf{x} \in \mathbb{R}^{n+} \mid C(\mathbf{x}) \in D\}$$

is nowhere dense and closed in $\mathbb{R}^{n+}$ and $C^{-1}(D) \cap \mathbb{R}_\Omega^{n+}$ is nowhere dense and closed in $\mathbb{R}_\Omega^{n+}$ for all $\Omega \in \Xi$.

Proof. Observe that the set $C^{-1}(D) = \bigcup_{\mathbf{u}\in D} S(\mathbf{u})$ (compare (28), where $S(\mathbf{u})$ has the form (27). Therefore $C^{-1}(D)$ is homotopic to $\varphi_\Omega(D) \times \Pi_{\ell=1}^{L}[0,\kappa^{(\ell)}]$. Since φ_Ω is a continuous function and D is closed, $C^{-1}(D)$ is closed. Now $\mathbb{R}^{n+}_\Omega$ is closed for every Ω and so $C^{-1}(D)$ is closed in $\mathbb{R}^{n+}_\Omega$.

Now we show that the complement of $C^{-1}(D)$ in $\mathbb{R}^{n+}_\Omega$ is dense in $\mathbb{R}^{n+}_\Omega$. Select an arbitrary point $\mathbf{x}_0 \in C^{-1}(D)$ in $\mathbb{R}^{n+}_\Omega$ for some Ω. We allow $\Omega = (\omega^1, \ldots, \omega^l)$ with $\omega^{(\ell)} = \{1, \ldots, n_\ell\}$ for all m, in which case all components of $\mathbf{x}_0$ are nonzero. Take an ϵ-neighbourhood $N_\epsilon \subset \mathbb{R}^{n+}_\Omega$ of $\mathbf{x}_0$. To show that $C^{-1}(D)$ is nowhere dense in $\mathbb{R}^{n+}_\Omega$ we need to find a point $\mathbf{x} \in N_\epsilon$ which does not belong to $C^{-1}(D)$. Since we restrict to $\mathbf{x} \in \mathbb{R}^{n+}_\Omega$ we have

$$C^{-1}(D) \cap \mathbb{R}^{n+}_\Omega = \varphi_\Omega(D).$$

Take $\mathbf{y} \in \varphi_\Omega(D) \cap N_\epsilon$ and choose δ such that a N_δ neighbourhood of $\mathbf{y}$ lies in N_ϵ. Since φ_Ω is a C^1 function with non-singular derivative at $\mathbf{z} := C(\mathbf{y})$ by the Inverse Mapping Theorem, $\varphi_\Omega^{-1}(N_\delta)$ is an open neighbourhood of the point $\mathbf{z} := C(\mathbf{y}), \mathbf{z} \in D$. Since D is nowhere dense in $\mathbb{R}^{n+}_\Omega$, there is a point $\mathbf{w} \in \mathbb{R}^{n+}_\Omega$ in this image with $\mathbf{w} \notin D$. Then there is $\mathbf{x} := \varphi_\Omega(\mathbf{w}) \in N_\delta \subset N_\epsilon$ for which we have $\mathbf{x} \notin C_\Omega^{-1}(D) \cap N_\epsilon$. Since ϵ was arbitrary, the complement of $C_\Omega^{-1}(D)$ is dense in $\mathbb{R}^{n+}_\Omega$.

If $\Omega \subset \Omega'$ and $C^{-1}(D)$ is nowhere dense in $\mathbb{R}^{n+}_\Omega$ then clearly $C^{-1}(D)$ is nowhere dense in $\mathbb{R}^{n+}_{\Omega'}$.

□

3.6 Proof of Theorem 3.1.

Recall from (19) that

$$X_i := \{\mathbf{x} \in \mathbb{R}^{n+} \mid C^k(\mathbf{x}) \in W\}, \quad k = 1, 2, \ldots$$

and from (18) that $W = \bigcup_{\Omega\in\Xi} W_\Omega$ is the set of points that converge to a set of internal fixed points Fix.

By Lemma 3.10 the set W is closed and nowhere dense in $\mathbb{R}^{n+}$ and $W \cap \mathbb{R}^{n+}_\Omega$ is nowhere dense closed in $\mathbb{R}^{n+}_\Omega$ for all $\Omega \in \Xi$.

By Theorem 3.14 with $D := W$ the set X_1 is also a closed and nowhere dense set in $\mathbb{R}^{n+}$ such that $X_1 \cap \mathbb{R}^{n+}_\Omega$ is nowhere dense closed in $\mathbb{R}^{n+}_\Omega$ for $\Omega \in \Xi$.

We proceed by induction on iteration step k, where Theorem 3.14 provides the induction step. Assume X_k is closed and nowhere dense set in $\mathbb{R}^{n+}$ such that $X_k \cap \mathbb{R}^{n+}_\Omega$ is nowhere dense closed in $\mathbb{R}^{n+}_\Omega$ for $\Omega \in \Xi$. Then by Theorem 3.14 X_{k+1} is closed and nowhere dense in $\mathbb{R}^{n+}$, and $X_{k+1} \cap \mathbb{R}^{n+}_\Omega$ is closed and nowhere dense for all $\Omega \in \Xi$.

By induction we conclude that X_k is closed and nowhere dense in $\mathbb{R}^{n+}$ and $X_k \cap \mathbb{R}^{n+}_\Omega$ is closed and nowhere dense in $\mathbb{R}^{n+}_\Omega$ for all $\Omega \in \Xi$. Therefore the set $U_k = \mathbb{R}^{n+} \setminus X_k$ is open and dense for all k. The set $G = \bigcap_{k=1}^{N} U_N \cap \Delta$ where N is selected by Lemma 3.7 is an open and dense set in Δ and $G \cap \Delta_\Omega$ is open and dense in in Δ_Ω for any $\Omega \in \Xi$. The set G represents a set of initial conditions whose iterations will never enter the set W, and therefore do not converge to any fixed point in the set of internal fixed points Fix. By Lemma 3.7 the corresponding trajectory then converges to either $\mathbf{0}$ vector or to a vector $\mathbf{e} = (a_1\mathbf{e}^{(1)}_{i_1}, \ldots, a_L\mathbf{e}^{(L)}_{i_L})$ for some positive collection of a_i and for some choice of vectors $\mathbf{e}^{(\ell)}_{i_\ell}$.

□

4 Discussion

The main result of this paper shows that for a generic correspondence array $c(i_1, \ldots, i_L)$ there is an open and dense set G of initial gating constants $x^{(\ell)}_{i_\ell}$ such the map-seeking circuit always converge to a either a zero solution (i.e. $x^{(\ell)}_{i_\ell} = 0$ for all layers ℓ and all mappings i_ℓ) or it converges to a solution where on each layer ℓ there exists precisely one weight $x^{(\ell)}_{i_\ell}$ which is nonzero and all other weights are equal to zero. The first result is interpreted as "no match found", while the second implies that circuit finds unique composition map

$$T = T^{(L)}_{i_L} \circ \ldots \circ T^{(\ell)}_{i_\ell} \circ \ldots \circ T^{(1)}_{i_1}.$$

This confirms numerical observations of the behavior of the algorithm. Our result is independent of the choice of mappings and of input and reference images or models. In practice thresholds or sigmoid-like non-linearities are used to separate what is considered to be a useful match from what is not. This is a problem encountered in all recognition problems and is independent of the algorithm used to discover the transformations that map the input image to the model or memory template. In real vision problems part of the target is often occluded so the matching threshold must be set to accommodate the fact that only part of the template is matched. On the other hand, the threshold cannot be set so low that trivial matches (e.g. random straight lines in the input image) will be judged a successful match. From the discussion above it is apparent that the MSC algorithm will find the best match in the image, so that even with low thresholds trivial matches will only be reported if a more substantial potential match is absent anywhere in the image. In difficult problems which include high degrees of occlusion, either by solid or scattered (e.g. foliage) occluders it is necessary to use sophisticated, non-linear criteria to distinguish between meaningful and non-meaningful matches when the thresholds are low and low correlation matches are encountered [6].

Acknowledgement. The research of T. G. was partially supported by NSF-BITS grant 0129895, NIH-NCRR grant PR16445, NSF/NIH grant W0467 and NSF-CRCNS grant W0577.

References

[1] D. Arathorn, “Map-Seeking: Recognition Under Transformation Using A Superposition Ordering Property,” *Electronics Letters* 37(3):164-165, February 2001.

[2] D. Arathorn, Map-Seeking Circuits in Visual Cognition, Palo Alto, Stanford University Press 2002.

[3] D. Arathorn, “Memory-driven visual attention: an emergent behavior of map-seeking circuits”, in *Neurobiology of Attention*, Eds. Itti L, Rees G, Tsotsos J, Academic/Elsevier, 2004.

[4] D. Arathorn, “A Cortically-Plausible Inverse Problem Solving Method Applied to Recognizing Static and Kinematic 3D Objects,” Neural Information Processing Systems (NIPS) 18, 2005.

[5] D. Arathorn, “Computation in the Higher Visual Cortices: Map-Seeking Circuit Theory and Application to Machine Vision,” Proceedings of IEEE AIPR 2004, October 2004.

[6] D. Arathorn, “Cortically plausible inverse problem method applied to complex perceptual and planning tasks”, Proceedings SPIE Defense and Security Symposium, April 2006.

[7] D. Arathorn, “From Wolves Hunting Elk to Rubik’s Cubes: Are the Cortices Composition / Decomposition Engines?,” Proceedings AAAI Symposium on Connectionist Compositionality, October 2004.

[8] D. Arathorn, “The Map-seeking Method: From Cortical Theory to Inverse-Problem Method, 2005-2006 IMA/MCIM Industrial Problems Seminar,” http://www.ima.umn.edu/industrial/2005-2006/abstracts.html#arathorn

[9] S. Harker, T. Gedeon and C. Vogel, “Analysis of Constrained Optimization Variants of the Map-Seeking Circuit Algorithm”, to appear in Journal of Mathematical Imaging and Vision.

[10] J. K. Hale and H. Kocak, Dynamics and Bifurcations, Springer-Verlag, New York, 1991.

[11] J. La Salle, “The extent of asymptotic stability,” Proc. Natl. Acad. Sci. USA, 46, 363-365, (1960).

[12] C. Vogel, D. Arathorn, A. Parker, and A. Roorda, “Retinal motion tracking in adaptive optics scanning laser ophthalmoscopy,” Proceedings of OSA Conference on Signal Recovery and Synthesis, Charlotte NC, June 2005.

[13] C. Vogel, D. Arathorn, A. Roorda, and A. Parker, “Retinal motion estimation and image dewarping in adaptive optics scanning laser ophthalmoscopy,” Optics Express, 14, 487-497 (2006)

[14] R. Snider and D. Arathorn, “Terrain discovery and navigation of a multi-articulated linear robot using map-seeking circuits”, Proceedings SPIE Defense and Security Symposium, April 2006.