# On the Use of Hyperspheres in Artificial Immune Systems as Antibody Recognition Regions

Thomas Stibor[1] and Jonathan Timmis[2] and Claudia Eckert[1]

[1] Department of Computer Science
Darmstadt University of Technology
{*stibor,eckert*}*@sec.informatik.tu-darmstadt.de*
[2] Departments of Electronics and Computer Science
University of York, Heslington, York
*jtimmis@cs.york.ac.uk*

**Abstract.** Using hyperspheres as antibody recognition regions is an established abstraction which was initially proposed by theoretical immunologists for use in the modeling of antibody-antigen interactions. This abstraction is also employed in the development of many artificial immune system algorithms. Here, we show several undesirable properties of hyperspheres, especially when operating in high dimensions and discuss the problems of hyperspheres as recognition regions and how they have affected overall performance of certain algorithms in the context of real-valued negative selection.

## 1 Introduction

Work in theoretical immunology has developed various representations for the interactions between antibody and antigen, and affinity metrics for modeling these such interactions. These antibody-antigen binding models were proposed for describing antibody cross-reactivity and multi-specificity [1] or for estimating the antibody repertoire size [2]. This work has provided much of the foundations for the development of artificial immune system (AIS) [3].

AIS is a paradigm inspired by the immune system and is used for solving computational and information processing problems. AIS exploit principles and methods developed by theoretical and experimental immunology, and abstract certain properties which can be implemented in computational systems [3]. In this paper, the abstraction we will consider is the hypersphere. This abstraction of hyperspheres has been used in many artificial immune system algorithms which have been applied to many areas such as anomaly detection, pattern recognition and clustering problems [4,5,6,7,8,9]. In this paper we describe mathematical properties of hyperspheres, which manifest themselves in high-dimensional space, and we provide suggestions on the applicability of hyperspheres as recognition units. Moreover we discuss the applicability of hyperspheres in the context of real-valued negative selection and explain reported poor classification results shown in [6].

The paper is organized as follows : In section 2 the real-valued shape-space is outlined and the most commonly used Euclidean distance is presented. Section 3 describes the abstraction of an antibody as a hypersphere. In section 3.1 the known hypersphere volume formula and the construction idea of that formula is shown and properties of that formula are presented in section 4. Next, the maximum volume of hyperspheres with respect to the dimension and the radius is presented in section 4.1, and we highlight unexpected properties of hyperspheres in high dimensions. In section 4.2, based on the mathematical observations, implications on the use of hyperspheres as antibody recognition regions are provided. We then present an algorithm for estimating, as opposed to exactly calculating, the total space of overlapping hyperspheres (section 5). Finally, results in sections 3.1, 4 and 5 are applied to explain in section 6 the poor classification results shown in [6].

## 2 Real-Valued Shape-Space and Euclidean Distance

The notion of *shape-space* was introduced by Perelson and Oster [1] and allows a quantitative affinity description between antibodies and antigens. More precisely, a shape-space is a metric space with an associated distance (affinity) function. The real-valued shape-space is the $n$-dimensional Euclidean space $\mathbb{R}^n$, where every element is represented as a $n$-dimensional point or simply as a vector represented by a list of $n$ real numbers. The Euclidean distance[3] $d$ is the (standard) distance between any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and is defined as :

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2} \tag{1}$$

Moreover, the Euclidean distance $d$ satisfies the metric properties :

$$\begin{aligned} \text{non-negativity}: &\quad d(\mathbf{x}, \mathbf{y}) \geq 0 \\ \text{reflexivity}: &\quad d(\mathbf{x}, \mathbf{y}) = 0 \text{ iff } \mathbf{x} = \mathbf{y} \\ \text{symmetry}: &\quad d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x}) \\ \text{triangle inequality}: &\quad d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \geq d(\mathbf{x}, \mathbf{z}) \\ & \\ & \text{for all vectors} \quad \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n \end{aligned}$$

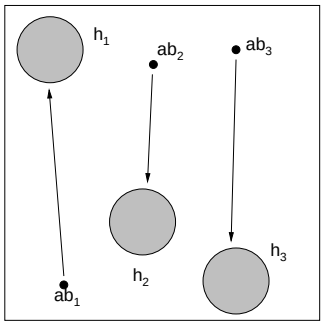
and therefore is frequently applied as a distance measurement in AIS algorithms.

## 3 Hyperspheres as Antibody Recognition Regions

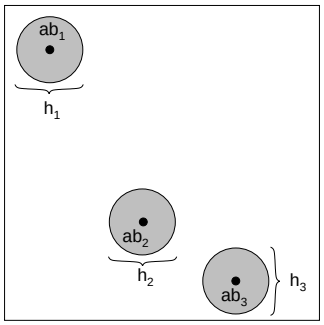
In the original work by Perelson and Oster [1], real-valued shape-space is introduced for estimating the probability that a randomly encountered antigen is recognized by at least one of the antibodies. An antibody is specified by $n$ parameters, e.g. the length, width, charge, etc. and can be described as a $n$-dimensional

[3] also termed Euclidean norm

point in the shape-space $\mathbb{R}^n$. Furthermore, an antibody recognizes not only one specific antigen, but several similar antigens which have a certain specificity — this property is called cross-reactivity[4]. In [1] each antibody is represented as a $n$-dimensional point and its (cross-reactivity) recognition space is modeled as a hypersphere — called an antibody recognition region. Antigens which lie within the hypersphere are recognized by the associated antibody. From an immunological point of view, antibodies recognize antigens which have a complementary binding site instead of similar binding regions (see Fig. 1(a)). This inspired Hart et al. [10] to develop a simulation to investigate empirically complementary binding properties in a immune network, with regard to emerging recognition regions. Hart et al. reported that the resultant immune network depended very much on the affinity metric employed (see [10] for further details).




(a) Complementary antibodies recognition regions. Antibodies $ab_i$ recognize all antigens which lie within the complementary hyperspheres $h_i$




(b) Non-complementary antibodies recognition regions. Antibodies $ab_i$ recognize all antigens which lie within the hyperspheres $h_i$

**Fig. 1.** Real-valued Shape-Space with complementer and non-complementer antibody recognition regions (modeling cross-reactivity)

For solving information processing problems, like pattern recognition, anomaly detection and clustering problems, the complementary recognition approach is possibly less appropriate, as it less obvious how one might employ such an idea. For such problems, it is useful to recognize points which are similar instead of complementary and therefore, similarity antibody-antigen recognition approaches are typically applied (see Fig. 1(b)). More precisely, an antibody can be represented as a hypersphere with center $\mathbf{ab} \in \mathbb{R}^n$ and a radius $r \in \mathbb{R}$. An antigen $\mathbf{ag} \in \mathbb{R}^n$ is recognized by an antibody $\mathbf{ab}$, when it lies within the hypersphere, i.e. $d(\mathbf{ab}, \mathbf{ag}) \leq r$.

[4] a well described explanation of the difference between cross-reactivity and multi-specificity is provided in [1], page 661

### 3.1 Volume of Hyperspheres

The volume of a $n$-dimensional hypersphere with radius $r$ can be calculated as follows :

$$V(n,r) = r^n \cdot \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$$

where

$\Gamma(n+1) = n! \quad \text{for} \quad n \in \mathbb{N} \quad \text{and}$
$\Gamma(n+\frac{1}{2}) = \frac{1 \cdot 3 \cdot 5 \cdot 7 \cdot \ldots \cdot (2n-1)}{2^n}\sqrt{\pi} \quad$ for half-integer arguments.

We briefly show the construction idea[5] behind the the volume calculation of hyperspheres. For a in-depth description see [11], where the complete construction and a proof is shown.

The volume $V(n)$ of a $n$-dimensional *unit sphere* can be constructed inductively

$$\begin{aligned}
V(2) &= \pi \\
V(3) &= \frac{4}{3}\pi \\
&\vdots \\
V(n) &= \begin{cases} \frac{\pi^{n/2}}{(n/2)!} & , n \text{ even} \\ \frac{2^n \pi^{(n-1)/2}\,((n-1)/2)!}{n!} & , n \text{ odd} \end{cases}
\end{aligned}$$

Given a 2-dimensional unit circle

$$C^2 = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$$

The volume $V(C^2)$ can be calculated as a summation of infinitely thin "stripes".

$$\begin{aligned}
V(C^2) &= 2 \cdot \int_{-1}^{1} \sqrt{1 - x_2^2}\, dx_2 \\
&= 2 \cdot \int_{0}^{\pi} \sqrt{1 - \cos^2(t)}\, \sin(t)\, dt \\
&= 2 \cdot \int_{0}^{\pi} \sin^2(t)\, dt \\
&= \int_{0}^{\pi} dt = \pi
\end{aligned}$$

[5] taken from [11]

$V(C^2) \to V(C^3)$

$$\begin{aligned} V(C^3) &= \int_{-1}^{1} \pi \left( \sqrt{1 - x_3^2} \right)^2 \, dx_3 \\ &= \pi \int_{-1}^{1} (1 - x_3^2) \, dx_3 \\ &= \frac{4}{3}\pi \end{aligned}$$

$\vdots$

$V(C^{n-1}) \to V(C^n)$

$$\begin{aligned} V(C^n) &= V(C^{n-1}) \cdot \int_{-1}^{1} (1 - x_n^2)^{(n-1)/2} \, dx_n \\ &= \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} \end{aligned}$$

**Proposition 1.** *The volume of a n-dimensional hypersphere with radius r is*

$$V(n, r) = r^n \cdot \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2} + 1\right)} \tag{2}$$

*Proof.* see [11]

## 4 Curse of Dimensionality

The phenomenon "curse of dimensionality" was first mentioned by Bellman [13] during his study of optimizing a function of a few dozen variables in an exhaustive search space. For example, given a function defined on a unitary hypercube of dimension $n$, in each dimension 10 discrete points are considered for evaluating the function. In dimension $n = 2$, this results in 100 evaluations, whereas in dimension $n = 10$, $10^{10}$ function evaluations are required. In general, an exponential number of $(1/\epsilon)^n$ function evaluations are required to obtain an optimization error of $\epsilon$ and therefore is computationally infeasible, even for a moderate $n$.

This simple example shows how problems like function optimization, which are computationally feasible in lower dimensions, transform to computationally infeasible problems in higher dimensions. A similar phenomenon (but not from the perspective of computational complexity) can be observed with hyperspheres in high-dimensional spaces, where they loose their familiar properties. In high-dimensions $\mathbb{R}^n$, i.e. $n > 3$, hyperspheres have undesirable properties. These properties (the following corollaries) can be derived directly from proposition (1).

**Corollary 1.** *The volume of hyperspheres converges to* 0 *for* $n \to \infty$.

$$\lim_{n \to \infty} V(n, r) = 0$$

*Proof.*

$$\lim_{n\to\infty}\left(r^n\cdot\frac{\pi^{n/2}}{\underbrace{\Gamma\left(\frac{n}{2}+1\right)}_{\approx\sqrt{2\pi}e^{-n}\,n^{n+\frac{1}{2}}}}\right)$$

$$=\frac{1}{\sqrt{2\pi}}\lim_{n\to\infty}\left(\frac{(\overbrace{r\,e\,\sqrt{\pi}}^{c})^n}{n^{n+\frac{1}{2}}}\right)=\frac{1}{\sqrt{2\pi}}\lim_{n\to\infty}\left(\frac{c^n}{n^{n+\frac{1}{2}}}\right)=0$$

□

**Corollary 2.** *The fraction of the volume which lies at values of the radius between $r-\epsilon$ and $r$, where $0<\epsilon<r$ is*

$$V_{fraction}(n,r,\epsilon)=1-\left(1-\frac{\epsilon}{r}\right)^n$$

*Proof.*

$$1-\frac{V(n,r-\epsilon)}{V(n,r)}=1-\left(\frac{\frac{(r-\epsilon)^n\cdot\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}}{\frac{r^n\cdot\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}}\right)=1-\left(1-\frac{\epsilon}{r}\right)^n$$

□

Corollary (1) implies that the higher the dimension the smaller the volume of a hypersphere for a fixed radii. This property is investigated in more detail, in the following section.
Corollary (2) reveals that in high-dimensional spaces, points which are uniformly randomly distributed inside the hypersphere, are predominately concentrated in a thin shell close to the surface or, in other words, at very high dimensions the entire volume of a hypersphere is concentrated immediately below the surface.

*Example 1.* Given a hypersphere with radius $r=1$, $\epsilon=0.1$ and $n=50$ and $k$ points which are uniformly randomly distributed inside the hypersphere, approximately $1-\left(1-\frac{0.1}{1}\right)^{50}\approx 99,5\,\%$ of the $k$ points lie within the thin $\epsilon$-shell close to the surface.

### 4.1 Volume Extrema

By keeping the radius fixed and differentiating the volume $V(n,r)$ with respect to $n$, one obtains the dimension[6] where the volume is maximal :

$$\frac{\partial}{\partial n}\left(\frac{r^n\cdot\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}\right)=\frac{r^n\ln(r)\,\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}+\frac{r^n\pi^{n/2}\ln(\pi)}{2\,\Gamma\left(\frac{n}{2}+1\right)}-\frac{r^n\pi^{n/2}\,\Psi\left(\frac{n}{2}+1\right)}{2\,\Gamma\left(\frac{n}{2}+1\right)}\quad(3)$$

[6] The dimension is obviously a nonnegative integer, however we consider term (3) analytically as a real-valued function

$$\text{where} \quad \Psi(n) = \frac{\partial}{\partial n} \ln \Gamma(n)$$

Vice versa, keeping the dimension fixed and differentiate term (2) with respect to $r$, it is not solvable in roots, i.e. no extrema exists :

$$\frac{\partial}{\partial r} \left( \frac{r^n \cdot \pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} \right) = \frac{r^n \, n \, \pi^{n/2}}{r \, \Gamma\left(\frac{n}{2}+1\right)} \tag{4}$$

For instance a hypersphere with radius $r = 1$ reaches its maximum volume in dimension 5 and looses volume in lower and higher dimensions. In figure 2 this property is visualized for different radius lengths $r = \{0.9, 1.0, 1.1, 1.2\}$. One can see that for each radius length in dimension from $n = 0$ to $n = 25$, the associated hypersphere reaches a maximal volume in a certain dimension and looses volume asymptotically in higher and lower dimensions.

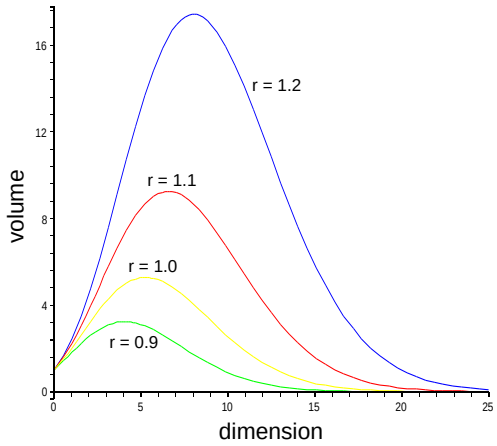



**Fig. 2.** Hypersphere volume plot for radius lengths $r = \{0.9, 1.0, 1.1, 1.2\}$ and dimension $n = 0, \ldots, 25$. Obviously, $n$ is a nonnegative integer, but the graph is drawn treating $n$ as continuously varying.

Table 1 presents the dimension where a hypersphere reaches its maximum volume for different radius lengths. Surprisingly, for radius lengths $r = 0.05$ and $r = 0.1$ the maximum volume lies in negative real-valued numbers. Hence, a volume maximization for such small radius lengths is not feasible, as the dimension is a nonnegative integer.

**Table 1.** Dimension where a hypersphere reaches the maximum volume for radius lengths $r = \{0.05, 0.1, 0.2, \ldots, 1.0\}$. Results are obtained by considering term (3) as a real-valued function.

| Radius $r$ | 0.05 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Dimension $\lfloor n \rfloor$ | -9.17 $\cdot 10^7$ | -88.94 | 1.59 | 1.12 | 1.0 | 1.03 | 1.20 | 1.53 | 2.14 | 3.23 | 5.27 |

### 4.2 Using Hyperspheres as Antibody Recognition Regions in Artificial Immune Systems

The results and observations presented in sections 3.1, 4 and 4.1 indicate that high-dimensional real-valued shape-spaces strongly bias the volume (recognition space) of hyperspheres. A hypersphere, for example with radius $r = 1$ has a high volume in relation to its radius length, up to dimension 15 (see Fig. 2). In higher dimensions ($n > 15$), for $r = 1$ the volume is nearly 0. This means that the recognition space — or in the context of antibody recognition region (covered space) — is nearly 0. In contrast, a radius that is too large ($r > 2$) in high dimensional spaces ($n > 10$) will imply an exponential volume. This exponential volume behavior, in combination with an unprecise volume estimation of overlapping hyperspheres, is the reason for the poor classification results reported in the paper [6] and is discussed in the subsequent sections.

## 5 Estimating Volume of Overlapping Hyperspheres

In section 3.1 a formula for calculating the exact volume of a hypersphere given by the dimension and the radius was shown. However, many proposed artificial immune system algorithms for solving pattern recognition, anomaly detection and clustering problems using not only *one* but multiple overlapping hyperspheres for classifying points [4,5,6,7,8,9]. Calculating analytically the total volume of overlapping hyperspheres is a very difficult task. Just the simple 2-dimensional case of three overlapping circles with different radii is a mathematical challenge. In the following section we describe a method to estimate the volume of (overlapping) hyperspheres.

### 5.1 Monte Carlo Integration

The Monte Carlo Integration is a method to integrate a function over a complicated domain, where analytical expressions are very difficult to apply – e.g. the calculation of the volume of overlapping hyperspheres in higher dimensions. Given integrals of the form $I = \int_{\mathcal{X}} h(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$, where $h(\mathbf{x})$ and $f(\mathbf{x})$ are functions for which $h(\mathbf{x})f(\mathbf{x})$ is integrable over the space $\mathcal{X}$, and $f(\mathbf{x})$ is a non-negative valued, integrable function satisfying $\int_{\mathcal{X}} f(\mathbf{x}) d\mathbf{x} = 1$. The Monte Carlo integration picks $N$ random points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$, over $\mathcal{X}$ and approximates the integral as

$$I \approx \frac{1}{N} \sum_{n=1}^{N} h(\mathbf{x}_n) \tag{5}$$

The absolute error of this method is *independent* of the dimension of the space $\mathcal{X}$ and decreases as $1/\sqrt{N}$. By applying this integration method, two fundamental questions arise :

- How many observations should one collect to ensure a specified statistical accuracy ?
- Given $N$ observations from a Monte Carlo Experiment, how accurate is the estimated solution ?

Both question are answered and discussed in [15]. Using the Chebyshev's inequality and specifying a *confidence level* $1-\delta$, one can determine the smallest sample size $N$ that guarantees an integration error no larger than $\epsilon$. In [15] this specification is called the $(\epsilon, \delta)$ *absolute error criterion* and leads to the worst-case sample size

$$N := \lceil 1/4\delta\epsilon^2 \rceil \tag{6}$$

### 5.2 Monte Carlo Hyperspheres Volume Integration

Using equations (5) and (6) a simple algorithm can be developed which estimates the total space (volume) covered by the hyperspheres inside the unitary hypercube $[0,1]^n$.

**Algorithm 1**: Monte Carlo Hyperspheres Volume Integration

**input** : $H$ = set of hyperspheres, $\epsilon$ = absolute error of the estimated volume, $\delta$ = confidence level
**output**: total volume of $H$

```
1  begin
2      inside ⟵ 0
       // calculate required worst-case
       // sample size N
3      N ⟵ ⌈1/4δε²⌉
4      for i ← 1 to N do
5          x ⟵ random point from [0,1]^n
6          foreach h ∈ H do
7              if dist(c_h, x) ≤ r_h then
                   // c_h is center of h, r_h is radius of h
8                  inside ⟵ inside + 1
9                  goto 5:
10     return (inside/N)
11 end
```

## 6 Limitation of Real-Valued Negative Selection in Higher Dimensions

In [6] an immune inspired real-valued negative selection algorithm was compared to different statistical anomaly detection techniques[7] for a high-dimensional anomaly detection problem. The investigations observed that the poorest classification results were real-valued negative selection, when compared to the statistical anomaly detection techniques on a 41-dimensional problem set (see [6] for further details). In this section, we attempt to explain this observation.

### 6.1 Real-Valued Negative Selection

The real-valued negative selection is an immune-inspired algorithm applied for anomaly detection. Roughly speaking, immune negative selection is a process which eliminates self-reactive lymphocytes and ensures that *only* those lymphocytes enter the blood stream that do not recognize self-cells[8]. As a consequence, lymphocytes which survive the negative selection process, are capable of recognizing nearly all foreign substances (like viruses, bacteria, etc.) which do not belong to the body. Abstracting this principle and modeling immune components according to the AIS framework [3] one obtains a technique for anomaly detection :

- Input : $S$ = set of points $\in [0,1]^n$ gathered from normal behavior of a system.
- Output : $D$ = set of hyperspheres, which recognizing a proportion $c_0$ of the total space $[0,1]^n$, except the normal points.
- Detector generation : While covered proportion $c_0$ is not reached, generate hyperspheres.
- Classification : If unseen point lies within a hypersphere, it does not belong to the normal behavior of the system and is classified as an anomaly.

A formal algorithmic description of real-valued negative selection is provided in [6].

### 6.2 Poor Classification Results

In [6] the real-valued negative selection technique (see section 6.1) was benchmarked by means of ROC analysis on a high-dimensional anomaly detection problem. The authors reported a detection rate of approximately $1\,\% - 2\,\%$ and a false alarm rate of $0\,\%$ when applying the real-valued negative selection algorithm. The false alarm rate of $0\,\%$ can be explained by learning $100\,\%$ of the training data and benchmarking with the training and testing data — similar false alarm rates results on other benchmarked data sets are reported in [5,16]. Benchmarking with $100\,\%$ training and testing data should be avoided, as in

[7] Parzen-Window, one class SVM
[8] Cells which belongs to the body

general it results in a high overfitted learning model and no representative (classification) results on the generalization performance will be obtained.

Moreover, the authors in [6] reported steady space coverage problems: these can be explained also by lack of precision when estimating the volume integration. Using term (6), which gives the worst-case sample size when given $\epsilon, \delta$, and applying the inequality

$$N + 1 > \frac{1}{4\delta\epsilon^2} \Longleftrightarrow \epsilon > \left(\frac{1}{4\delta(N+1)}\right)^{1/2} \tag{7}$$

one can easily see why the authors in [6] reported such steady space coverage problems for the estimated hyperspheres coverage of $c_0 = 80\,\%$. For the parameter $c_0$ which was originally proposed in [5] one obtains according to [5,6] a sample size of $N = 1/(1 - c_0) = 5$. Evaluating term (7) with a given confidence level of $90\,\%$, one obtains an integration error $\epsilon$ of greater than $65\,\%$. Inequality (7) can be used to explain the reported steady space coverage problems, however it does not explain thoroughly the poor classification results described in [6] — this is now explained by means of the results shown in sections 4 and 5.

Investigating the 41-dimensional data set [17], one can statistically verify[9], that the whole normalized non-anomalous class is concentrated at one place inside the unitary hypercube $\mathcal{U} = [0, 1]^{41}$. In [18] this characteristic is called "empty space phenomenon" and arises in any data set that does *not* grow exponentially with the dimension of the space.

In [6] the authors additionally reported, that the real-valued negative selection algorithm terminated when (on average) 1.4 detectors were generated. By generating only one detector (hypersphere) with, for example, a radius $r = 3$ and a detector center which does not necessarily lie inside $\mathcal{U}$, the volume of that hypersphere amounts $5.11 \cdot 10^{10}$. The unitary hypercube $\mathcal{U} = [0, 1]^{41}$ has a total volume of 1, however most of the volume of a hypercube is concentrated in the large corners, which themselves become very long "spikes" [12]. This can be verified by comparing the ratio of the distance $\sqrt{n}$ from the center of the hypercube to one of the edges to the perpendicular distance $a/2$ to one of the edges (see Fig. 3).

$$\frac{\left(\sum_{i=1}^{n}(\frac{a}{2})^2\right)^{1/2}}{\frac{a}{2}} = \frac{\left(n\,\frac{a^2}{4}\right)^{1/2}}{\frac{a}{2}} = \sqrt{n} \qquad \text{where } n \text{ is the dimension} \tag{8}$$

For $n \to \infty$, the term (8) goes to $\infty$ and therefore the volume is concentrated in very long "spikes" of $\mathcal{U}$.

As a consequence, the hypersphere covers some of those (high-volume) spikes which are lying within the $V_{fraction}$ proportion of the hypersphere. Hence, the real-valued negative selection algorithm terminates with only a very small number of large radii detectors (hyperspheres) which are covering a limited number of spikes. As a result a large proportion of the volume of the hypercube does

[9] by means of covariance matrix

not lie within the hyperspheres — it lies in the remaining (high-volume) spikes, though the hypersphere volume is far higher than the hypercube volume.

These observations in combination with the unprecise volume integration of overlapping hyperspheres results in the poor classification results reported in [6].

From our point of view, the real-valued negative selection would appear to be a technique that is not well suited for high-dimensional data sets, i.e. data dimensions far higher than 41 — a well established benchmark in the field of pattern classification is for instance the problem of handwritten digit recognition, the dimensionality of this problem domain is 256 [19,20]. We propose this is in part because it makes more sense to formulate a classification model with regard to the given training elements, instead of complementary space. The complementary (anomalous) space is exponentially large when compared to the "normal" space in high dimensions. The real-valued negative selection technique attempts to cover this high-dimensional space with hyperspheres, but as we have shown, these have adverse properties in such high-dimensional spaces.

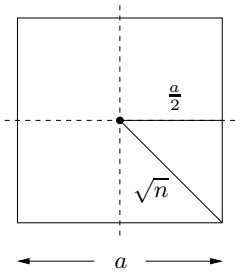



**Fig. 3.** Distance ratio $\frac{\sqrt{n}}{a/2}$ between a line from center to a corner and a perpendicular line from center to an edge

In [18] Verleysen discusses in detail, this curse of dimensionality problem, with respect to artificial neural networks. He suggests in general to change the distance measure function for high-dimensional problems, for instance by applying a higher-order norm ($h > 2$)

$$d_h(\mathbf{x}, \mathbf{y}) = \sqrt[h]{\mid x_1 - y_1 \mid^h + \ldots + \mid x_n - y_n \mid^h} \tag{9}$$

instead of the standard Euclidean norm. In the context of *inductive biases*[10], Freitas and Timmis [21] discussed different affinity measures in artificial immune systems. They illustrated the advantages and disadvantages of the 1-norm and 2-norm (see term (9)) and showed how one of these norm when compared to the other norm can lead to an overemphasizing of the distance. As a final summarizing sentence, the authors suggested that when developing an AIS, one should make a careful choice of the norm, as the norm should take into account

[10] effectiveness in problem domains

the characteristics (in our case the dimension) of the data being analyzed. Unfortunately, there seems to be no theoretical results, for correctly choosing the value $h$ with regard to the data dimension [18].

## 7 Conclusion

The immune system is an impressive recognition system with many appealing properties for the construction of artificial immune system algorithms. Abstracting antibodies as hyperspheres and applying the Euclidean distance metric for quantifying binding strengths, is an established method for modeling and simulating the immune systems.

For developing competitive immune-inspired algorithms the antibody-antigen representation and affinity metric is a crucial parameter. We have found that applying the abstraction of these hyperspheres for immune-inspired algorithms can lead to poor results, especially for high-dimensional classification problems.

In this paper, we have shown that these hypersphere have undesirable properties in high dimensions — the volume tends to zero and nearly all uniformly randomly distributed points are close to the hypersphere surface. We have presented these hypersphere properties and have provided an explanation for poor classification results reported in [6]. In addition, we have now explained the limitations of the real-valued negative selection for high-dimensional classification problems, when employing hyperspheres. There is no reason to suggest that the hypersphere properties we have discussed in this paper, are not valid observations for all high-dimensional classification problems where hyperspheres are applied as recognition regions. Therefore, as a result, these adverse hypersphere properties could bias all (artificial immune system) algorithms, which employ hyperspheres as recognition units.

## References


1. Perelson, A.S., Oster, G.F.: Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-nonself discrimination. In: J. Theor. Biol. Volume 81. (1979) 645–670
2. Percus, J.K., Percus, O.E., Perelson, A.S.: Predicting the size of the t-cell receptor and antibody combining region from consideration of efficient self-nonself discrimination. Proceedings of National Academy of Sciences USA **90** (1993) 1691–1695
3. de Castro, L.N., Timmis, J.: Artificial Immune Systems: A New Computational Intelligence Approach. Springer-Verlag (2002)
4. Gonzalez, F., Dasgupta, D., Nino, L.F.: A randomized real-valued negative selection algorithm. In: Proceedings of the 2nd International Conference on Artificial Immune Systems – ICARIS. Volume 2787 of Lecture Notes in Computer Science., Edinburgh, UK, Springer-Verlag (2003) 261–272
5. Ji, Z., Dasgupta, D.: Real-valued negative selection algorithm with variable-sized detectors. In: Genetic and Evolutionary Computation – GECCO, Part I. Volume 3102 of Lecture Notes in Computer Science., Seattle, WA, USA, Springer-Verlag (2004) 287–298

6. Stibor, T., Timmis, J., Eckert, C.: A comparative study of real-valued negative selection to statistical anomaly detection techniques. In: Proceedings of 4th International Conference on Artificial Immune Systems – ICARIS. Volume 3627 of Lecture Notes in Computer Science., Springer-Verlag (2005) 262–275
7. Watkins, A., Boggess, L.: A new classifier based on resource limited artificial immune systems. In: Proceedings of the 2002 Congress on Evolutionary Computation CEC2002, IEEE Press (2002) 1546–1551
8. Bezerra, G.B., Barra, T.V., de Castro, L.N., Zuben, F.J.V.: Adaptive radius immune algorithm for data clustering. In: Proceedings of 4th International Conference on Artificial Immune Systems – ICARIS. Volume 3627 of Lecture Notes in Computer Science., Springer-Verlag (2005) 290–303
9. Bentley, P.J., Greensmith, J., Ujjin, S.: Two ways to grow tissue for artificial immune systems. In: Proceedings of 4th International Conference on Artificial Immune Systems – ICARIS. Volume 3627 of Lecture Notes in Computer Science., Springer-Verlag (2005) 139–152
10. Hart, E., Ross, P.: Studies on the implications of shape-space models for idiotypic networks. In: Proceedings of 3th International Conference on Artificial Immune Systems – ICARIS. Volume 3239 of Lecture Notes in Computer Science., Springer-Verlag (2004) 413–426
11. Leppmeier, M.: Kugelpackungen von Kepler bis heute. Vieweg Verlag (1997)
12. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press (1995)
13. Bellman, R.: Adaptive Control Processes: A Guided Tour. Princeton University Press (1961)
14. Mosegaard, K., Sambridge, M.: Monte Carlo analysis of inverse problems. Inverse Problems **18** (2002) 29–54
15. Fishman, G.S.: Monte Carlo Concepts, Algorithms, and Applications. Springer (1995)
16. Stibor, T., Mohr, P.H., Timmis, J., Eckert, C.: Is negative selection appropriate for anomaly detection ? In: Proceedings of Genetic and Evolutionary Computation Conference – GECCO-2005, ACM Press (2005) 321–328
17. Hettich, S. and Bay, S. D.: KDD Cup 1999 Data (1999) http://kdd.ics.uci.edu.
18. Verleysen, M.: Learning high-dimensional data. Limitations and Future Trends in Neural Computation **186** (2003) 141–162
19. Vapnik, V.N.: The Nature of Statistical Learning Theory. Second edn. Springer-Verlag (1999)
20. Schölkopf, B., Platt, J.C., Shawe-Taylor, Shawe-Taylor, Smola, A.J., Williamson, R.C.: Estimating the support of a high-dimensional distribution. Technical Report MSR-TR-99-87, Microsoft Research (MSR) (1999)
21. Freitas, A., Timmis, J.: Revisiting the Foundations of Artificial Immune Systems: A Problem Oriented Perspective. In: Proceedings of the 2nd International Conference on Artificial Immune Systems – ICARIS. Volume 2787 of Lecture Notes in Computer Science., Springer-Verlag (2003) 229–241