JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Study of Lattice Vectors in Self-Organizing Maps

[1]Femy N S, [2]Dr.Sasi Gopalan

[1]Research Scholar, [2]Professor
[1,2]Department of Mathematics,
[1,2]Cochin University of Science and Technology, Kochi-22, Kerala, India

Abstract: Self-Organizing Maps is an unsupervised artificial neural network that is trained to produce a low-dimensional representation of high-dimensional data. In this work, a novel approach is made by introducing lattice vectors to the input space of self-organizing maps. The algebraic and geometric properties of lattice vectors in the linear initialization method of weight initialization in self-organizing maps are studied. It is also found that each data vector in the input space of self-organizing maps will lie in the fundamental region associated with the basis vectors of the lattice. The experimental analysis is done by taking dataset from UCI machine learning repository and using MATLAB R2021a.

Index Terms **– Determinant, Fundamental parallelepiped, Lattice vectors, Projection, Self-Organizing Maps.**

I. Introduction

Self-Organizing Maps (SOM) is introduced by Teuvo Kohonen in 1980's. It plays an important role in the visualization of high-dimensional data in a low-dimensional space. Many works have been done in the field of supervised and unsupervised machine learning using lattice theory. The work of Laurentiu Iancu [9], uses lattice algebra approach to neural computation. The study of Bhavana and Sarma [10], used the concept lattice in dimensionality reduction by matrix factorization in data mining. Noah Stephens [15], works on the Gaussian measure over lattices. Kelechi Chuwkunonyerem emerole [6], works on optimizing Gaussian measure of Lattices using dimensionality reduction. Mohammad-Reza Sadeghi et. al [17], work on feed-forward neural network lattice decoding algorithm in deep learning. The term 'lattice' has two meanings-one is related to the theory of partial ordering on sets and the other is related to discrete subgroups of $\mathbb{R}^n$ [8]. Lattices have many significant applications in coding theory and cryptography. In this work, a novel approach to an unsupervised machine learning technique called Self-Organizing Maps using lattice vectors is presented by considering lattices as additive subgroups of $\mathbb{R}^n$.

The paper is structured as follows. Section 1 gives the introduction. In Section 2 preliminaries of Self-Organizing Maps, architecture of SOM, and basic mathematical definitions are discussed. Section 3 presents mathematical analysis of SOM using lattice vectors. In Section 4 experimental analysis of theorems in Section 3 is done using MATLAB R2021a. Finally, the paper ends with a conclusion in Section 5.

II. Preliminaries

2.1 Self-Organizing Maps

In SOM the input layer consists of neurons representing the features of the input vectors in $\mathbb{R}^n$. The output layer consists of neurons representing the clusters, and associated with each neuron in the output layer there is a weight vector in $\mathbb{R}^n$. The minimum distance between the input vector and the weight vector in the output layer is evaluated using the Euclidean distance [11]. The input vector is mapped into the neuron with minimum distance. The weight initialization is done using linear initialization in which weight vectors are spanned by the eigenvectors corresponding to the largest two eigenvalues of the input data [1], [2], [16]. Let $\{u_1, u_2\}$ be two such eigenvectors formed. Throughout this work $\{u_1, u_2\}$ are linearly independent vectors in $\mathbb{R}^n$.

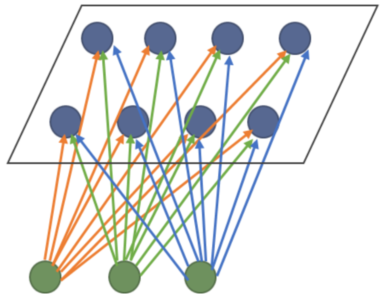

Figure 1: Architecture of SOM

Definition 2.1[5],[8]: Let $\{\overrightarrow{b_1}, \overrightarrow{b_2}, \dots, \overrightarrow{b_m}\}$ be linearly independent set of column vectors in $\mathbb{R}^n, n \geq m$. The lattice generated by $\{\overrightarrow{b_1}, \overrightarrow{b_2}, \dots, \overrightarrow{b_m}\}$ is the set $L = \{\sum_{i=1}^{m} l_i \overrightarrow{b_i} : l_i \in \mathbb{Z}\}$ of integer linear combinations of $\overrightarrow{b_i}$. The vectors $\{\overrightarrow{b_1}, \overrightarrow{b_2}, \dots, \overrightarrow{b_m}\}$ are called lattice basis.

Definition 2.2[8]: A basis B of a lattice L is a $n \times m$ matrix formed by taking the columns to be basis vectors $\overrightarrow{b_i}, i = 1,2, \dots, m$. Thus $L = \{B\vec{x} : \vec{x} \in \mathbb{Z}^m\}$.

Definition 2.3[14][18]: A lattice $L \subseteq \mathbb{R}^n$ is called discrete if $\exists\ \varepsilon > 0$ such that $\forall\ x \neq y \in L, \|x - y\| \geq \varepsilon$.

Definition 2.4[14][18]: A subset $L \subseteq \mathbb{R}^n$ is called additive subgroup of $\mathbb{R}^n$ if $\forall\ x, y \in L, x - y \in L$.

Definition 2.5[12]: A set $F \subseteq \mathbb{R}^n$ is called a fundamental region of a lattice $L \subseteq \mathbb{R}^n$ if the following conditions are satisfied:

1. $\mathbb{R}^n = \bigcup_{\vec{v} \in L} (\vec{v} + F)$
2. For every $\overrightarrow{v_1}, \overrightarrow{v_2} \in L$ with $\overrightarrow{v_1} \neq \overrightarrow{v_2}, (\overrightarrow{v_1} + F) \cap (\overrightarrow{v_2} + F) = \phi$.

The cosets of lattice L in $\mathbb{R}^n$ forms the fundamental region. An important fundamental region is a fundamental parallelepiped.

Definition 2.6[7][12]: Given two linearly independent vectors $u_1, u_2 \in \mathbb{R}^n$ their fundamental parallelepiped is defined as

$P(u_1, u_2) = \{\sum_{i=1}^{2} x_i \overrightarrow{u_i} : x_i \in \mathbb{R}, 0 \leq x_i < 1\}$. $P(u_1, u_2)$ is half-open.

Example 2.7: The lattice $\mathcal{L}$ with basis vectors $\overrightarrow{u_1} = (1,0)$ and $\overrightarrow{u_2} = (0,1)$ in $\mathbb{R}^2$ and its associated fundamental parallelepiped is shown in Figure 2. The lattice $\mathcal{L}$ with basis in a two dimensional space can be easily visualized but for greater than two it is difficult to visualize. So a fundamental parallelepiped associated with the basis vectors $\overrightarrow{u_1}$ and $\overrightarrow{u_2}$ in $\mathbb{R}^2$ is shown in Figure 2.

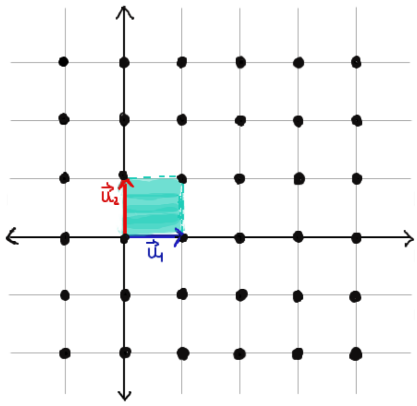

Figure 2

Theorem 2.8[8]: Let B be a $n \times m$ basis matrix for a lattice L where $n > m$. Then there is a linear map $T: \mathbb{R}^n \to \mathbb{R}^m$ such that $T(L)$ is a rank m lattice and $\|T(\vec{u})\| = \|\vec{u}\|$ for all $\vec{u} \in L$. If the linear map is represented by a $m \times n$ matrix T so that $T(\vec{u}) = T\vec{u}$ then a basis matrix for the image of L under the projection T is the $m \times m$ matrix TB which is invertible.

Definition 2.9[8]: The determinant of a lattice L with basis B is $|\det(PB)|$, where P is the projection of Theorem 2.8.

III. MATHEMATICAL ANALYSIS

Definition 3.1: In linear initialization method of weight initialization in SOM, let $\{u_1, u_2\}$ be the eigenvectors corresponding to the largest two eigenvalues of the input vectors in $\mathbb{R}^n$. The lattice generated by $\{u_1, u_2\}$ is the set $\mathcal{L} = \{B\vec{x} : \vec{x} \in \mathbb{Z}^2\}$.

Lemma 3.1: In linear initialization method of SOM the vectors in the lattice $\mathcal{L}$ generated by the basis $\{u_1, u_2\} \subseteq \mathbb{R}^n$ form an abelian group under addition in $\mathbb{R}^n$.

Proof:

1. For $\overrightarrow{v_1}, \overrightarrow{v_2} \in \mathcal{L}$, $\overrightarrow{v_1} + \overrightarrow{v_2} \in \mathcal{L}$ (Closure Property)
2. $(\overrightarrow{v_1} + \overrightarrow{v_2}) + \overrightarrow{v_3} = \overrightarrow{v_1} + (\overrightarrow{v_2} + \overrightarrow{v_3})$ (Associative Property)
3. For every $\vec{v} \in \mathcal{L}\ \exists\ \vec{0} \in \mathcal{L}$ such that $\vec{v} + \vec{0} = \vec{v} = \vec{0} + \vec{v}$. (Additive Identity)
4. For every $\vec{v} \in \mathcal{L}\ \exists - \vec{v} \in \mathcal{L}$ such that $\vec{v} + -\vec{v} = \vec{0} = -\vec{v} + \vec{v}$. (Additive inverse)

Hence the lattice $\mathcal{L}$ forms an abelian group under addition in $\mathbb{R}^n$[4].

Theorem 3.2: In linear initialization method of SOM the lattice $\mathcal{L}$ formed by the basis $\{u_1, u_2\}$ in $\mathbb{R}^n$ is a discrete additive subgroup of $\mathbb{R}^n$.

Proof: Assume that $\mathcal{L}$ is a lattice.
For $\vec{v_1}, \vec{v_2} \in \mathcal{L}$, $\vec{v_1} - \vec{v_2} \in \mathcal{L}$. Hence $\mathcal{L}$ is an additive subgroup of $\mathbb{R}^n$.
The length of any lattice vector must be greater than the length of a shortest lattice vector.
Therefore, from the lower bound on a shortest lattice vector, $\|\vec{v_1} - \vec{v_2}\| \geq \lambda_1(\mathcal{L})$.
Let $\varepsilon = \lambda_1(\mathcal{L})$.
Hence $\|\vec{v_1} - \vec{v_2}\| \geq \varepsilon$.
Therefore, $\mathcal{L}$ is a discrete additive subgroup of $\mathbb{R}^n$.

Theorem 3.3: In linear initialization method of weight initialization in SOM, let B be a $n \times 2$ basis matrix for a lattice $\mathcal{L}$ which is a subset of $\mathbb{R}^n$ where $n > 2$. Then there is a linear map $P: \mathbb{R}^n \to \mathbb{R}^2$ such that $P(\mathcal{L})$ is a rank 2 lattice and $\|P(\vec{v})\| = \|\vec{v}\|$ for all $\vec{v} \in \mathcal{L}$. If the linear map is represented by a $2 \times n$ matrix P so that $P(\vec{v}) = P\vec{v}$ then a basis matrix for the image of $\mathcal{L}$ under the projection P is the 2×2 matrix $PB = I$, the identity matrix, which is invertible.

Proof: Let B be the $n \times 2$ basis matrix with columns $\vec{u_i}, i = 1,2$.
Define $V = \text{Span}\{\vec{u_1}, \vec{u_2}\}$
By Gram-Schmidt orthogonalization process choose a basis $\vec{v_1}, \vec{v_2}$ for V that is orthonormal with respect to the innerproduct in $\mathbb{R}^n$.
Choose $\vec{v_1} = \vec{u_1}$

$$\vec{v_2} = \vec{u_2} - \frac{\langle \vec{u_2}, \vec{v_1} \rangle}{\langle \vec{v_1}, \vec{v_1} \rangle} \vec{v_1}$$

Define a linear map $P: V \to \mathbb{R}^2$ by $P(\vec{v_i}) = \vec{e_i}$
For $\vec{v} = \sum_{i=1}^{2} x_i \vec{v_i} \in V$ we have

$$\begin{aligned}
\|\vec{v}\| &= \sqrt{\langle \vec{v}, \vec{v} \rangle} \\
&= \sqrt{\langle \textstyle\sum_{i=1}^{2} x_i \vec{v_i}, \sum_{i=1}^{2} x_i \vec{v_i} \rangle} \\
&= \sqrt{\langle x_1 \vec{v_1} + x_2 \vec{v_2}, x_1 \vec{v_1} + x_2 \vec{v_2} \rangle} \\
&= \sqrt{{x_1}^2 \langle \vec{v_1}, \vec{v_1} \rangle + x_1 x_2 \langle \vec{v_1}, \vec{v_2} \rangle + x_2 x_1 \langle \vec{v_2}, \vec{v_1} \rangle + {x_2}^2 \langle \vec{v_2}, \vec{v_2} \rangle} \\
&= \sqrt{{x_1}^2 + {x_2}^2} \text{ since } \langle \vec{v_1}, \vec{v_2} \rangle = \langle \vec{v_2}, \vec{v_1} \rangle = 0 \text{ and } \langle \vec{v_1}, \vec{v_1} \rangle = \langle \vec{v_2}, \vec{v_2} \rangle = 1. \\
&= \|P\vec{v}\|
\end{aligned}$$

Since the vectors $\vec{u_i}$ form a basis for V, the vectors $P\vec{u_i}$ are linearly independent. Hence PB is an invertible matrix and $P(\mathcal{L})$ is a lattice of rank 2.

Theorem 3.4: In linear initialization method of SOM, for every input data vector $\vec{x} \in \mathbb{R}^n$, there exist a unique lattice vector $\vec{v} \in \mathcal{L}$, such that $\vec{x} \in \left(\vec{v} + P(u_1, u_2)\right)$, where $P(u_1, u_2)$ is the fundamental parallelepiped associated to u_1 and u_2.

Proof: Let $P(u_1, u_2)$ be the fundamental parallelepiped associated with u_1 and u_2.
Let $\vec{v} \in \mathcal{L}$, where $\mathcal{L}$ is the lattice in $\mathbb{R}^n$.
Then $\vec{v} + P(u_1, u_2)$ forms a partition of the whole space $\mathbb{R}^n$ [12].
This is because as $P(u_1, u_2)$ is a fundamental region, $\mathbb{R}^n = \bigcup_{\vec{v} \in \mathcal{L}} \left(\vec{v} + P(u_1, u_2)\right)$ and for every $\vec{v_1}, \vec{v_2} \in \mathcal{L}$ with $\vec{v_1} \neq \vec{v_2}$, $\left(\vec{v_1} + P(u_1, u_2)\right) \cap \left(\vec{v_2} + P(u_1, u_2)\right) = \phi$.
Hence for every $\vec{x} \in \mathbb{R}^n$, there exists a unique point $\vec{v} \in \mathcal{L}$, such that $\vec{x} \in \left(\vec{v} + P(u_1, u_2)\right)$.

Theorem 3.5: In linear initialization of SOM, let $\mathcal{L}$ be a lattice of rank 2, and $u_1, u_2 \in \mathcal{L}$ be two linearly independent lattice vectors then $\{u_1, u_2\}$ form a basis of $\mathcal{L}$ if and only if $P(u_1, u_2) \cap \mathcal{L} = \{0\}$.

Proof: Assume first $\vec{u_1}$ and $\vec{u_2}$ form a basis of lattice $\mathcal{L}$. Then by definition, $\mathcal{L}$ is the set of all integer linear combinations of $\vec{u_1}$ and $\vec{u_2}$.
Since $P(\vec{u_1}, \vec{u_2})$ is defined as the set of linear combinations of $\vec{u_1}$ and $\vec{u_2}$ with coefficients in $[0,1)$, the intersection of two sets is $\{0\}$.
Conversely, assume $P(\vec{u_1}, \vec{u_2}) \cap \mathcal{L} = \{0\}$. Since $\mathcal{L}$ is a rank 2 lattice and $\vec{u_1}$ and $\vec{u_2}$ are linearly independent, we can write any lattice vector $\vec{x} \in \mathcal{L}$ as $\sum x_i \vec{u_i}$ for some $x_i \in \mathbb{R}$.
Since by definition a lattice is closed under addition, the vector $x^1 = \sum (x_i - [\![x_i]\!]) \vec{u_i}$ is also in $\mathcal{L}$, where $[\![x_i]\!]$ denotes the greatest integer less than or equal to x_i.
By our assumption, $x^1 = 0$.
Hence $\sum x_i \vec{u_i} = \sum [\![x_i]\!] \vec{u_i}$.
This implies that all $x_i's$ are integers and hence $\vec{x}$ is an integer linear combination of $\vec{u_1}$ and $\vec{u_2}$.
Therefore $\{\vec{u_1}, \vec{u_2}\}$ form a basis of $\mathcal{L}$.

Definition 3.6: The determinant of a lattice $\mathcal{L}$ with basis B is $|\det(PB)|$, where P is the projection of Theorem 3.3.

Remark: The determinant of a lattice is inversely proportional to the density of lattice vectors. Hence smaller the value of determinant the denser will be the lattice. If we take a large ball $B(0, r)$ with center at the origin and radius r in the span of $L(B)$ then the number of lattice vectors inside $B(0, r)$ approaches to $\frac{vol(B(0,r))}{\det(L(B))}$ as the size of the ball goes to infinity. This gives an approximate number of lattice vectors inside the ball $B(0, r)$ in SOM.

IV. EXPERIMENTAL ANALYSIS

Analysis of Theorem 3.3 is done with datasets from UCI machine learning repository and using MATLAB R2021a [3], [13]. The datasets used are IrisInputs which is a 4×150 matrix of four attributes (petal length, petal width, sepal length, and sepal width) of 150 flowers. The covariance matrix for each dataset is evaluated and from that the corresponding eigenvectors and eigenvalues are found. The eigenvectors corresponding to the largest two eigenvalues are taken as $\vec{u_1}$ and $\vec{u_2}$. Table 1 shows the covariance matrix for IrisInputs dataset. Table 2 gives the eigenvectors and eigenvalues of the covariance matrix. In Table 3 the largest two eigenvectors $\vec{u_1}$ and $\vec{u_2}$, the basis matrix B formed by $\vec{u_1}$ and $\vec{u_2}$, the matrix P, which is the projection matrix are given. Table 4 gives two vectors in $\text{Span}\{\vec{u_1}, \vec{u_2}\}$ and shows that $\|P\vec{v}\| = \|\vec{v}\|$ for each of these vectors. In Table 5 the basis matrix B for lattice $\mathcal{L}$, the projection matrix P, and its product matrix PB are found, showing that the product matrix PB is an Identity matrix and is invertible.

Table 1: Covariance matrix

Data(X)	A=Cov(X)			
IrisInputs	0.6857	-0.0393	1.2737	0.5169
	-0.0393	0.1880	-0.3217	-0.1180
	1.2737	-0.3217	3.1132	1.2964
	0.5169	-0.1180	1.2964	0.5824

Table 2: Eigenvectors and eigenvalues of the covariance matrix

Data(X)	Eigenvector(A)				Eigenvalue(A)			
IrisInputs	-0.3173	0.5810	0.6565	0.3616	0.0237	0	0	0
	0.3241	-0.5964	0.7297	-0.0823	0	0.0785	0	0
	0.4797	-0.0725	-0.1758	0.8566	0	0	0.2422	0
	-0.7511	-0.5491	-0.0747	0.3588	0	0	0	4.2248

Table 3: The largest two eigenvectors and the basis matrix B

Data(X)	$\overrightarrow{u_1}$	$\overrightarrow{u_2}$	$B = [\overrightarrow{u_1}, \overrightarrow{u_2}]$	$P = B^T$
IrisInputs	0.3616	0.6565	$\begin{bmatrix} 0.3616 & 0.6565 \\ -0.0823 & 0.7297 \\ 0.8566 & -0.1758 \\ 0.3588 & -0.0747 \end{bmatrix}$	$\begin{bmatrix} 0.3616 & -0.0823 & 0.8566 & 0.3588 \\ 0.6565 & 0.7297 & -0.1758 & -0.0747 \end{bmatrix}$
	-0.0823	0.7297		
	0.8566	-0.1758		
	0.3588	-0.0747		

Table 4: Two vectors $\overrightarrow{v_1}, \overrightarrow{v_2}$ in Span$\{\overrightarrow{u_1}, \overrightarrow{u_2}\}$ showing $\|P\vec{v}\| = \|\vec{v}\|$

Data(X)	$\overrightarrow{v_1} = 2\overrightarrow{u_1} + \overrightarrow{u_2}$	$\overrightarrow{v_2} = 3\overrightarrow{u_1} - \overrightarrow{u_2}$	$P\overrightarrow{v_1}$	$P\overrightarrow{v_2}$	$\|\overrightarrow{v_1}\|$	$\|\overrightarrow{v_2}\|$	$\|P\overrightarrow{v_1}\|$	$\|P\overrightarrow{v_2}\|$
IrisInputs	1.3797	0.4282	$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 3 \\ -1 \end{bmatrix}$	2.2361	3.1623	2.2361	3.1623
	0.5652	-0.9765						
	1.5374	2.7455						
	0.6430	1.1512						

Table 5: The basis matrix B of lattice $\mathcal{L}$, the projection matrix P and its product matrix PB

Data(X)	P	B	PB
IrisInputs	$\begin{bmatrix} 0.3616 & -0.0823 & 0.8566 & 0.3588 \\ 0.6565 & 0.7297 & -0.1758 & -0.0747 \end{bmatrix}$	$\begin{bmatrix} 0.3616 & 0.6565 \\ -0.0823 & 0.7297 \\ 0.8566 & -0.1758 \\ 0.3588 & -0.0747 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

V. CONCLUSION

In this work, a new approach to SOM is introduced by using lattice vectors. In the linear initialization method of weight initialization in SOM, the algebraic and geometric properties of lattice vectors are studied. It is shown that for each data vector in the input space there exists a unique lattice vector in the fundamental region associated with the basis vectors of the lattice. Experimental analysis is done using MATLAB R2021a by taking datasets from the UCI machine learning repository. Future studies can be conducted on random sample initialization of weight initialization in self-organizing maps.

REFERENCES

[1] Akinduko AA, Mirkes EM, Gorban AN. 2016. SOM: Stochastic initialization versus principal components. Information Sciences; 364:213-21.

[2] Attik M, Bougrain L, Alexandre F. 2005. Self-organizing map initialization. In International Conference on Artificial Neural Networks (pp. 357-362). Springer, Berlin, Heidelberg.

[3] Beale MH, Hagan MT, Demuth HB. Neural network toolbox. User's Guide, MathWorks. 2010;2:77-81.

[4] Birkhoff G. 1938. Lattices and their applications. Bulletin of the American Mathematical Society. 1938 Dec;44(12. P1):793-800.

[5] Cassels JW. 2012. An introduction to the geometry of numbers. Springer Science & Business Media.

[6] Emerole KC, Boussakta S. 2020. Optimizing Gaussian Measure of Lattices using Dimensionality Reduction. In ICC 2020-2020 IEEE International Conference on Communications (ICC) (pp. 1-7). IEEE.

[7] Fischer RF. 2005. Precoding and signal shaping for digital transmission. John Wiley & Sons.

[8] Galbraith SD. 2012. Mathematics of public key cryptography. Cambridge University Press.

[9] Iancu L. 2005. Lattice algebra approach to neural computation. University of Florida.

[10] Jamalpur B, Sarma SS. 2015. Dimensionality reduction by matrix factorization using concept lattice in data mining.

[11] Kohonen T. 2012. Self-organizing maps. Springer Science & Business Media.

[12] Kschischang FR, Feng C. 2014. An introduction to lattices and their applications in communications.

[13] MathWorks, Self-Organizing Maps, https://in.mathworks.com/help/deeplearning/ug/cluster-with-self-organizing-map-neural-network.html

[14] Micciancio D, Goldwasser S. 2002. Complexity of lattice problems: a cryptographic perspective. Springer Science & Business Media.

[15] Noah Stephens-Davidowitz, 2017. On the Gaussian Measure over Lattices, PhD thesis, New York University.

[16] Ritter H, Martinetz T, Schulten K. 1992. Neural computation and self-organizing maps: an introduction. Reading, MA: Addison-Wesley.

[17] Sadeghi MR, Amirzade F, Panario D, Sakzad A. 2018. A Neural Network Lattice Decoding Algorithm. In 2018 IEEE Information Theory Workshop (ITW) (pp. 1-5). IEEE.

[18] V.Vaikuntanathan. 2011. Lattices in Computer Science, Class notes at the University of Toronto.