Where $t_{i,m} \in \mathbb{R}$ is on the line segment between $\omega_i^m.\xi^j$ and $\omega_i^{m+1}.\xi^j$, by combination (7), (11), and (12) and (39), we have

$$E(\omega^{(m+1)J}) - E(\omega^{mJ}) \le -\frac{1}{\eta_m}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m \omega_i^{mJ}\|^2 + \frac{\lambda}{2J}\sum_{k=1}^{N}\sum_{\ell=1}^{J}\left(\frac{f'(\omega_k^{mJ})}{f(\omega_k^{mJ})^{1/2}} + F''(t_{n,k,m})d_k^{m,j}\right)(\omega_k^{(m+1)J} - \omega_k^{mJ})$$
$$+\delta_1 + \delta_2 + \delta_3 \quad (40)$$

Where

$$\delta_1 = \sum_{\ell=1}^{J} g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\sum_{k=1}^{N}\left(\prod_{\substack{i=1\\i\neq n}}^{N}(\omega_i^{mJ}.\xi^j)(\omega_k^{(m+1)J} - \omega_k^{mJ})\xi^j\right)$$

$$\delta_2 = \sum_{\ell=1}^{J}\frac{1}{2}g_j''(t_{i,m}')\left(\prod_{i=1}^{N}(\omega_i^{(m+1)J}.\xi^j) - \prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)^2$$

$$\delta_3 = \frac{1}{2}\sum_{\ell=1}^{J} g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\sum_{\substack{k_1,k_2=1\\k_1\neq k_2}}^{N}\left(\prod_{\substack{i=1\\i\neq k_1,k_2}}^{N} t_{i,m,j}\right)[(\omega_{k_1}^{(m+1)J} - \omega_{k_1}^{mJ})][(\omega_{k_2}^{(m+1)J} - \omega_{k_2}^{mJ})]\,\xi^j$$

Where $t_{i,m}'$ *and* $t_{n,k,m,j}$ lies in between $\omega_i^{mJ}.x^j$ and $\omega_i^{(m+1)J}.x^j$, and from (23), (24) and (45), $M = \frac{\sqrt{6}}{\sqrt{a^3}}$, *and* $F(x) \equiv (f(x))^{\frac{1}{2}}$. Note that

$$F'(x) = \frac{f'(x)}{2\sqrt{f(x)}}$$
$$F''(x) = \frac{2f''(x).f(x) - [f'(x)]^2}{4[f(x)]^{\frac{3}{2}}}$$
$$\le \frac{f''(x)}{2\sqrt{f(x)}} \le \frac{\sqrt{6}}{2\sqrt{a^3}}$$

By *(25), (30) and Lemma 4.3 for* $1 \le j \le J, 1 \le k \le N, m = 0,1,2,\dots$, *and Cauchy- Schwartz Theorem, we have*

$$\frac{\lambda}{2J}F''(t_{n,k,m})\left|\left(\sum_{i=1}^{N}\sum_{j=1}^{J}(d_k^{m,j}).(\omega_k^{(m+1)J} - \omega_k^{mJ})\right)\right|$$
$$\le \lambda M\sum_{i=1}^{N}\sum_{j=1}^{J}|d_k^{m,j}|\circ\|\omega_k^{(m+1)J} - \omega_k^{mJ}\|$$
$$\le \frac{\lambda M}{J}C_4^2\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2$$
$$\le C_{10}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \quad (41)$$

Where $C_{10} = \lambda M(1 + C_3\eta_0)^2/J$ and $t_{n,k,m}$ lies in between $\omega_i^{mJ}.x^j$ and $\omega_i^{(m+1)J}.x^j$

By Assumption (A1), (A2), (12) and (25), we have

$$|\delta_1| \le \frac{1}{\eta_m}\left|\sum_{i=1}^{N}\sum_{j=1}^{J}\left(g_j'\left(\prod_{i=1}^{N}(\omega_i^{mJ}.\xi^j)\right)\prod_{\substack{i=1\\i\neq k}}^{N}(\omega_i^{mJ}.\xi^j)\xi^j + \frac{\lambda}{2J}\frac{f'(\omega_k^{mJ})}{f(\omega_k^{mJ})^{1/2}}\right)\right|\|\omega_k^{(m+1)J} - \omega_k^{mJ}\|$$
$$\le C_{11}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \quad (42)$$

Where $C_{11} = C_4NJ$.

By Assumption (A1), (21), (24), and (26) for $m = 0,1,2,\dots$, we have

$$|\delta_2| \le \frac{1}{2}CC_5^2(1 + NC_3\eta_m)^2\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2$$
$$\le C_{12}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \quad (43)$$

Where $C_{12} = \frac{1}{2}CC_5^2(1 + NC_3\eta_m)^2$.

Using Assumption (A1), (A2), (25) and Cauchy- Schwartz Theorem, we get

$$|\delta_2| \le \frac{1}{2}C^{N-1}\left|\sum_{j=1}^{J}\sum_{\substack{k_1,k_2=1\\k_1\neq k_2}}^{N}\left((\omega_{k_1}^{(m+1)J} - \omega_{k_1}^{mJ})(\omega_{k_2}^{(m+1)J} - \omega_{k_2}^{mJ})\,\xi^j\right)\right|$$
$$\le \frac{1}{2}JC^{N+1}\sum_{\ell=1}^{J}\sum_{\substack{k_1,k_2=1\\k_1\neq k_2}}^{N}|d_{k_1}^{m,j}|.|d_{k_2}^{m,j}|$$
$$\le \frac{1}{2}JC^{(N+1)}(N-1)C_4^2\sum_{k=1}^{N}\left(\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|\right)^2$$
$$\le C_{13}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2, \; m = 0,1,2,\dots \quad (44)$$

Where $C_{13} = JC^{(N+1)}(N-1)C_4^2/2$.

Substituting (41) - (44) into (40), then, we have

$$E(\omega^{(m+1)J}) - E(\omega^{mJ}) \le \left(-\frac{1}{\eta_m} + C_{10} + C_{11} + C_{12} + C_{13}\right)\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2$$
$$\le -\left(\frac{1}{\eta_m} - C_1\right)\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2$$
$$\le 0. \quad (45)$$

This completes the proof to statement $(i)$ of theorem 3.2.

**Proof to $(ii)$ of theorem 3.2.**

From the conclusion of $(i)$, we know that the nonnegative sequence $\{E(W^m)\}$ is monotone. However, it is also bounded below. Hence there must exist $E^* \ge 0$ such that $\lim_{k\to\infty} E(W^m) = E^*$. The proof to $(ii)$ it thus completed.

**Proof to $(iii)$ of theorem 3.2.**

It is follows from Assumption (A4) that $\beta > 0$. Taking $\beta = \frac{1}{\eta_m} - C_1$ and using (45), we suppose that $M$ is positive integer, we have

$$E(W^{(M+1J)}) \le E(W^{MJ}) - \beta\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{MJ}\|^2 \le \cdots$$
$$\le E(W^0) - \beta\sum_{m=0}^{M}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2.$$

Since $E(W^{m+1}) \ge 0$, we have

$$\beta\sum_{m=0}^{M}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 \le E(\omega^0) \le \infty.$$

Let $M \to \infty$, then

$$\sum_{m=0}^{\infty}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 < \frac{1}{\beta}E(W^0) < \infty.$$

Thus results in

$$\lim_{m\to\infty}\sum_{i=1}^{N}\sum_{j=1}^{J}\|\Delta_j^m\omega_i^{mJ}\|^2 = 0.$$

From (10) - (12) and (A1) it is easily get

$$\lim_{m\to\infty}\|\Delta_j^m\omega_i^{mJ}\| = 0, \lim_{m\to\infty}\|E_{\omega_k}(\omega^{mJ})\| = 0 \quad (46)$$

The proof to $(iii)$ is thus completed.

**Proof to $(iv)$ of theorem 3.2.**

Note that the error function $E(W)$ defined in (7) is continuous and differentiable. According to (46), (A5) and

Lemma 4.2, we can easily get the desired results, i.e., there exists a point $\omega^* \in \Omega_0$ such that

$$\lim_{m\to\infty}\left(\omega_i^{m_J}\right) = \omega_i^*$$

This completes the proof to $(iv)$

## 3. Conclusions

In this paper, we investigate a Batch Gradient Method with Smoothing $L_{1/2}$ Regularization for Pi-sigma Neural Networks. The Smoothing $L_{1/2}$ Regularization is a term proportional to the magnitude of the weights. We prove under moderate conditions that the weights of the networks are keeping bounded in the learning process. The both weak and convergence results require the boundedness of the weights is precondition.

## 4. Acknowledgment

We gratefully acknowledge Dalanj University and Dalian University of Technology for supporting this research. Special thanks to Prof. Dr. Wei Wu and Dr. Yan Liu for their kind helps during the period of the study.

## References

[1] M. Sinha, K. Kumar, and P. K. Kalra, "Some new neural network architectures with improved learning schemes", Soft computing, 4, pp. 214 – 223, 2000.
[2] P. Estevez, Y. Okabe, "Training the piecewise linear-high order neural network through error back propagation", International Joint Conference on Neural Networks, 1, pp. 711 -716, 1991.
[3] Y. Shin, J. Ghosh, "The pi-sigma network: an efficient higher-order neural network for pattern classification and function approximation", International Joint Conference on Neural Networks, 1, pp. 13 -18, 1991.
[4] A. J. Hussain, P. Liatsis, "A new recurrent polynomial neural network for predictive image coding", Image Processing and Its Applications, 1, pp. 82- 86, 1999.
[5] L. Villalobos, F. L. Merat, "Learning capability assessment and feature space optimization for higher-order neural networks", IEEE Transaction on Neural Networks,, 6, pp. 267 – 272, 1995.
[6] Y. Shin, J. Ghosh, D. Samani, "Computationally efficient invariant pattern classification with higher-order pi-sigma networks", Intelligent Engineering Systems through Artificial Neural Networks, 2, pp. 379 – 384, 1992.
[7] J. L. Jing, F. Xu, R. S. Piao, "Application of pi-sigma neural network to real- time classification of seafloor sediments", Applied Acoustics, 24, pp. 346 – 350, 2005.
[8] J. A. Hussaina, P. Liatsisb, "Recurrent pi- sigma networks DPCM image coding", Neurocomputing, 55, pp. 363- 382, 2002.
[9] M. Sinha, K. Kumar, P. K. Kalra, "Some new neural network architures with improved learning schemes", Soft Convergence, 4, pp. 214 – 223, 2004.
[10] S. Haykin, "Neural Networks: A Comprehensive Foundation", 2nd edition. Beijing: *Tsinghua University Press* and Prentice Hall, 2001.
[11] M. Ishikawa, "Structural learning with forgetting", Neural Networks, 9(3), pp. 509 – 521, 1996.
[12] V. Setiono, "A penalty – function approach for pruning feed forward neural networks", Neural Networks, 9, pp. 185 -204, 1997.
[13] E. D. Karnin, "A simple Procedure for pruning back – propagation trained neural networks", IEEE Trans. Neural Network, 1(2), pp. 239 – 242, 1990.
[14] Wu Wei, "Hongmei Shao and Zhengxue Li, Convergence of batch BP algorithm with penalty for FNN training" Lecture Notes in Computer Science, 4232, pp. 562-569, 2006.
[15] Shao Hongmei, Zheng Gaofeng, "Boundedness and convergence of online gradient method with penalty and momentum", Neurocomputing, 74, pp. 765-770, 2011.
[16] Shao Hongmei, Wu Wei, Li Feng, "Convergence of online gradient method with a penalty term for feed forward neural networks with stochastic inputs". Numerical Mathematics, 1(14), pp. 87-96, 2005.
[17] T. Evgeniou, M. Pontil, T. Poggio, "Regularization networks and support vector machine"s, Adv. Comput. Math., 13, pp. 1 – 50, 2000.
[18] Wu Wei, Fan Qinwei, Jacek M. Zurada, Jian Wang, Dakun Yang, Yan Liu, "Batch gradient method with smoothing $L_{1/2}$ regularization for training of feed forward neural networks", Neural Networks, 50, pp. 72-78, 2014.
[19] Fan Qinwei, Jacek M. Zurada and Wei Wu, "Convergence of online gradient method for feed forward neural networks with smoothing $L_{1/2}$ regularization penalty", Neurocmputing, 131, pp. 208 – 216, 2014.
[20] P. Billingsley, "Probability and Measure", New York: John Wiley, 1979.
[21] X. D. Kang, Y. Xiong, C. Zhang, W. Wu, "Convergence of online gradient algorithm with stochastic inputs for Pi-sigma neural networks", Proc. IEEE Symp. Foundations of Comput. Intelligence, IEEE Press, pp. 564 – 569, 2007.
[22] W. Wu, G. R. Feng, X. Li, "Training multilayer perceptrons via minimization of sum of ridge functions", Advances in Computational Mathematics, 17, pp. 335 – 347, 2002.
[23] Y. X. Yuan, W. Y. Sun, "Optimization Theory and Methods", Science Press, Beijing, 2011.
[24] Shao Hongmei, Zheng Gaofeng, "Boundedness and convergence of online gradient method with penalty and momentum", Neurocomputing, 74, pp. 765 – 770, 2011.
[25] Z. X. Li, W. Wu, Y. L. Tian, "Convergence of an online gradient method for feed forward neural networks with stochastic inputs", Journal of Computational and Applied Mathematics, 163(1), pp. 165 –176, 2004

## Authors Profiles

**Khidir Shaib Mohamed** (PhD student in Computational Mathematics) received the B.S. in Mathematics from Dalanj University – Dalanj – Sudan (2006) and M.S. in Applied Mathematics from Jilin University – Changchun – China (2011). He works as a lecturer of mathematics at College of Science – Dalanj University since (2011). Now he is a PhD student in Computational Mathematics at School of University of Technology, Dalian.

**Yousif Shoaib Mohammed** (Assistant Professor of Computational Physics) received the B.S. in Physics from Khartoum University – Oudurman – Sudan (1994) and High Diploma in Solar Physics from Sudan University of Science and Technology – Khartoum – Sudan (1997) and M.S. in Computational Physics (Solid – Magnetism) from Jordan University – Amman – Jordan and PhD in Computational Physics (Solid – Magnetism – Semi Conductors) from Jilin University – Changchun – China (2010) and worked as Researcher at Africa City of Technology – Khartoum – Sudan since 2012. He worked at Dalanj University since 1994 up to 2013 then from 2013 up to now at Qassim University – Kingdom of Saudi Arabia.

**Abd Elmoniem Ahmed Elzain** is Associate Professor in Physics and Researcher received the B.Sc. in Physics from Kassala University – Kassala - Sudan (1996) and High Diploma of Physics from Gezira University – Madani – Sudan (1997) and M.Sc. in Physics from Yarmouk University – Erbid – Jordan (2000) and PhD in Applied Radiation Physics from Kassala University – Kassala – Sudan (2006). He worked at Kassala University since 1996 up to 2010 then from 2010 up to now at Qassim University – Kingdom of Saudi Arabia.

**Mohamed El-Hafiz M. Noor** is Assistant Professor in Mathematice, received the B.Sc. in Mathematics from Wadi Elneel University – Atbara - Sudan (1996) and M.Sc. of Mathematics from Sudan University of Science and Technology – Khartoum – Sudan (1999) and PhD in Mathematics from Alneelain University – Khartoum – Sudan (2010). He worked at Zalingi University since 1997 up to 2013 then from 2012 up to now at Qassim University – Kingdom of Saudi Arabia.

**Ellnoor Abaker Abdrhman Noh** (Associate Professor of Physical Chemistry) received the B.S. in Physics from Khartoum University – Oudurman – Sudan (1991) and M.S. in Physical Chemistry (Corrosion) from Yarmouk University – Erbid – Jordan (1998) and PhD in Physical Chemistry (Computational) from North East Normal University – Changchun – China (2005). He worked at Dalanj University since 1992 up to 2011 then from 2011 up to now at Albaha University – Kingdom of Saudi Arabia.