L^p Approximation of Sigma–Pi Neural Networks

Yue-hu Luo and Shi-yi Shen

***Abstract*—A feedforward Sigma–Pi neural networks with a single hidden layer of m neural is given by**

$$\sum_{j=1}^{m} c_j g\left(\prod_{k=1}^{n} \frac{x_k - \theta_k^j}{\lambda_k^j}\right)$$

where c_j, θ_k^j, $\lambda_k \in \mathbb{R}$. In this paper, we investigate the approximation of arbitrary functions f: $\mathbb{R}^n \to \mathbb{R}$ by a Sigma-Pi neural networks in the L^p norm. For an L^p locally integrable function $g(t)$ can approximation any given function, if and only if $g(t)$ can not be written in the form $\sum_{j=1}^{n}\sum_{k=0}^{m} a_{jk}(\ln|t|)^{j-1}t^k$.

***Index Terms*—L^p approximation, neural network, Sigma–Pi function.**

I. Introduction

One of the most interesting questions connection neural networks is the approximation capability of neural networks. There have been many papers related to this topic. For the reference one can read [1]–[13] and the references given there. In this paper, the p-approximation capability of Sigma–Pi neural networks is investigated. In mathematical terminology, it is to find some conditions such that under which all the linear combinations of elements $g\left(\prod_{k=1}^{n} x_k - \theta_k/\alpha_k\right)$ are dense in $L^p(K)$ for any compact set K in $\mathbb{R}^n$, where $x = (x_1, x_2, \ldots, x_n)$ is a variant, $\theta_k \in \mathbb{R}$ are constants. Let $g(t)$ be a function on $\mathbb{R}$. $g(t)$ is said to be a SP-function if all functions of the following form:

$$\sum_{j=1}^{m} c_j g\left(\prod_{k=1}^{n} \frac{x_k - \theta_k^j}{\lambda_k^j}\right), \qquad m = 1, 2, \ldots \tag{1}$$

are dense in $C(K)$ for any compact set K in $\mathbb{R}^n$. $g(t)$ is said to be an L^pSP-function if all functions of the following form:

$$\sum_{k=1}^{m} c_j g(\lambda_k \cdot x - \theta_k), \qquad \lambda_k \in \mathbb{R}^n, \theta_k \in \mathbb{R} \tag{2}$$

are dense in $C(K)$ for any compact set K in $\mathbb{R}^n$. $g(t)$ is said to be an L^pTW-function if all functions of the form (2) are dense in $L^p(K)$ for any compact set K in $\mathbb{R}^n$.

It has been proved in [9] and [10] that if $g(t)$ is continuous on $\mathbb{R}$ then $g(t)$ is a SP-function if and only if $g(t)$ can not be written in the following form:

$$a_0 + \sum_{j=1}^{n} (\ln|t|)^{j-1} \sum_{k=1}^{m} a_{jk} t^k. \tag{3}$$

Recently, the present authors in [15] have proved that the previous result still holds under the condition that $g(t)$ is locally Riemman integrable, i.e., if $g(t)$ is locally Riemman integrable on $\mathbb{R}$ then $g(t)$ is an SP-function if and only if $g(t)$ can not be written in the form (3). Therefore, if the previous conditions are satisfied, then $g(t)$ is an L^pSP-function. It is well known that an L^p Locally integrable function on $\mathbb{R}$ may neither be continuous, nor be locally Riemman integrable. It is natural to ask what is that characteristic condition for an L^p locally integrable function to be an L^pSP-function. The purpose of this paper is to answer this question and give a necessary and sufficient condition to solve the problem. The remainder of this paper is organized as follows. Our main results are presented in Section II. In order to prove the main results, several lemmas are given in Section III. Finally, the proof of the main results is presented in Section IV.

II. Main Results

The main results in this paper are as follows.

Theorem 1: Let $g(t)$ be L^p locally integrable on $\mathbb{R}$ with $1 \le p < \infty$. Then $g(t)$ is an L^pSP-function if and only if $g(t)$ can not be written in the following form on $\mathbb{R}$ (a.e):

$$a_0 + \sum_{j=1}^{n}\sum_{k=0}^{m} (\ln|t|)^{j-1} a_{jk} t^k \tag{4}$$

where a_{jk} are constants.

The following Corollary 2 is an immediate consequence of Theorem 1.

Corollary 2: Let $g(t)$ be L^p locally integrable on $\mathbb{R}$ with $1 \le p < \infty$. Then $g(t)$ is an L^pTW-function if and only if $g(t)$ is not a polynomial on $\mathbb{R}$ (a.e).

Remark 3: Hornik in [4] and M. Leshno *et al.* in [3] proved that $g(t)$ is a L^pTW-function if $g(t)$ is essentially locally bounded and nonpolynomial on $\mathbb{R}$. Cheng in [12] proved that $g(t)$ is an L^pTW-function if $g(t)$ is L^p locally integrable and belongs to $S'(\mathbb{R})$ (i.e., the integral $\int_{\mathbb{R}} f(t)g(t)\,dt$ is existent for any rapidly decreasing infinitely differentiable function $f(t)$ on $\mathbb{R}$). Corollary 2 is a generalization of these results.

III. Some Notations and Lemmas

For $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, an n-multi-index α, i.e., $\alpha \in \mathbb{R}^n$ and each coordinate of α is a nonnegative integer, let $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$, $\lambda x = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n)$, $|\alpha| = \sum_{k=1}^{n} \alpha_k$, $\lambda^{-1} = (\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1})$ if each $\lambda_k \neq 0$, and $\pi(\cdot)$ be a function on $K \subset \mathbb{R}^n$ defined by

$$\pi(x) = \prod_{k=1}^{n} x_k, \qquad \forall\, x = (x_1, x_2, \ldots, x_n) \in K. \tag{5}$$

Manuscript received October 8, 1998; revised June 3, 1999.

Y.-h. Luo is with the Department of Mathematics, Nanjin University of Science and Technology, Nanjin, China, 210094.

S.-y. Shen is with the Department of Mathemathics, Nankai University, Tianjin, China, 300071.

Publisher Item Identifier S 1045-9227(00)09849-0.

For a function g on $\mathbb{R}$ and a function f on $\mathbb{R}^n$, let supp f be the support of f, i.e., supp $f = \overline{\{x \in \mathbb{R}^n | f(x) \neq 0\}}$. Let

$$\Sigma_0(f, K) = \left\{ \sum_{k=1}^{m} c_k g \left(y_k \prod_{k=1}^{n} (x_k - \theta_k) \right) | y_k, c_k, \lambda_k \in \mathbb{R}, \right. \\ \left. 1 \le k \le n, m \ge 1 \right\} \quad (6)$$

$$\Sigma_0(f, K) = \left\{ \sum_{k=1}^{m} c_k f(\lambda_k^{-1} x + \theta_k) \,|\theta_k, \lambda_k \in \mathbb{R}^n, c_k \in \mathbb{R}, \right. \\ \left. 1 \le k \le n, m \ge 1 \right\}^{*)} \quad (7)$$

where $x = (x_1, x_2, \ldots x_n) \in K$ is a variant.[1] And let $\hat{g}$ be defined by

$$\hat{g}\colon \hat{g}(x) = g(\pi(x)), \qquad \forall\, x \in \mathbb{R}^n. \quad (8)$$

For a function ϕ on $\mathbb{R}$ and a function ψ on $\mathbb{R}^n$, $g \diamond \phi$ and $f * \psi$ are defined by

$$g \diamond \phi(t) = \int_{\mathbb{R}} g(ts)\phi(s)\, ds, \qquad \forall\, t \in \mathbb{R}$$

and

$$f * \psi(x) = \int_{\mathbb{R}^n} f(x-y)\psi(y)\, dy, \qquad \forall\, x \in \mathbb{R}^n \quad (9)$$

respectively. Let $C^\infty(\mathbb{R}^n)$ be the set consisting of all infinitely continuously differentiable functions on $\mathbb{R}^n$. For a set A in $\mathbb{R}^n$, let

$$\begin{aligned} C^\infty(A) &= \{f | f \in C^\infty(\mathbb{R}^n), \text{ supp } f \subset A\}, \\ C_0(A) &= \{f | f \in C(\mathbb{R}^n), \text{ supp} f \text{ is compact and} \\ &\qquad \text{supp } f \subset A\}, \\ C_0^\alpha(A) &= \{f | f \in C(\mathbb{R}^n), \text{ supp} f \text{ is compact} \\ &\qquad \text{supp } f \subset A \text{ and } D^\beta f \in C(R^n), \forall\, \beta \le \alpha\}, \\ C_0^\infty(A) &= \{f | f \in C^\infty(\mathbb{R}^n), \text{ supp } f \text{ is compact and} \\ &\qquad \text{supp } f \subset A\}, \\ L_{loc}^p(\mathbb{R}^n) &= \{f | f \text{ is } L^p \text{ locally integrable on } \mathbb{R}^n\}. \end{aligned}$$

If f is α-times differential at x, $D^\alpha f(x)$ is the partial derivative of $f(x)$ at x, i.e.,

$$D^\alpha f(x) = \frac{\partial^{|\alpha|} f(x)}{\partial {x_1}^{\alpha_1} \partial {x_2}^{\alpha_2} \ldots \partial {x_n}^{\alpha_n}}.$$

If f is locally integrable on $\mathbb{R}^n$, then $D^\alpha f$ is the general derivative, i.e., $D^\alpha f$ is a linear functional on $C_0^\infty(\mathbb{R}^n)$ defined by

$$D^\alpha f(\phi) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x) D^\alpha \phi(x)\, dx, \quad \forall\, \phi \in C_0^\infty(\mathbb{R}^n).$$

In order to prove Theorem 1, we need the following Lemmas. Lemma 3–Lemma 6 can be proved by the argument used in [3], [4], [9], [10], [15]. Lemma 1 is Corollary 4 in [9], Lemma 2 is Theorem 10 in [9] (or Corollary in [10]).

[1] $\pi(x) \neq 0$ is satisfied

Lemma 1 [9]: Let K be a compact set in $\mathbb{R}^n$ and $f \in C(\mathbb{R}^n)$. If $D^\alpha f \neq 0$, then x^α belongs to the closure of $\Sigma_0(f, K)$ in $C(K)$.

Lemma 2 [9], [10]: Let $g(t) \in C(R)$. Then the following statements are equivalent:

1) the closure of $\Sigma(g, K)$ in $C(K)$ is equal to $C(K)$ for any compact set K in $\mathbb{R}^n$;
2) g cannot be written in the form (3).

Lemma 3: Suppose that $g(t) \in C^\infty(\mathbb{R} \backslash \{0\})$ and $0 \notin (a, b)$. Then $g(t)$ can not be written in the form (4) on the interval (a, b), if and only if for any n-multiindex α there exists an $x_\alpha \in \mathbb{R}^n$ such that

$$D^\alpha \hat{g}(x_\alpha) \neq 0, \qquad \pi(x_\alpha) \in (a, b). \quad (10)$$

Proof: Necessity. Otherwise, there exists an n-multiindex α such that

$$D^\alpha \hat{g}(x) = 0, \qquad \forall\, \pi(x) \in (a, b) \quad (11)$$

which implies that

$$D^{(m+1)e} \hat{g}(x) = 0, \qquad \forall\, \pi(x) \in (a, b) \quad (12)$$

for any integer m with $(m+1)e \ge \alpha$, where $e = (1, 1, \ldots, 1)$. By some computation, we see that (12) is equivalent to

$$\sum_{j=1}^{(n-1)(m+1)+1} q_j (\pi(x))^{j-1} g^{(m+j)}(\pi(x)) = 0 \\ \forall\, \pi(x) \in (a, b) \quad (13)$$

where q_j's are constants depending on m. In other words, $g(t)$ is a solution of the following differential equation of order $n(m+1)$:

$$\sum_{j=1}^{(n-1)(m+1)+1} q_j t^{j-1} g^{(m+j)}(t) = 0, \qquad \forall\, \pi(x) \in (a, b). \quad (14)$$

It is easy to verify that $D^{(m+1)e} (\ln \pi(x))^{j-1} (\pi(x))^k = 0$ for any j, k with $1 \le j \le n$ and $0 \le k \le m$, which implies that $t^j \ln^k |t|$, $0 \le j \le n-1$ and $0 \le k \le m$ are solutions of (14). Thus any solution of (14) may be written in the form (4) because $t^j \ln^k |t|$, $0 \le j \le n-1$, and $0 \le k \le m$ are linearly independent. This completes the proof of the necessity.

Sufficiency. It is easy to see that $g(t)$ satisfies (12) if $g(t)$ can be written in the form (4) on (a, b), which completes the proof of sufficiency. ∎

Lemma 4: Let $g(t) \in L_{loc}^p$, $0 \notin (a, b)$ and $0 < \delta < 1$. If $g(t)$ can not be written in the form (4) on (a, b), then there exists a $\phi_0 \in C_0^\infty([1-\delta, 1+\delta])$ such that $(g \diamond \phi_0)(t)$ can not be written in the form (4) on (a, b).

Proof: Otherwise, $(g \diamond \phi)(t)$ may be written in the form (4) on (a, b) for any $\phi \in C_0^\infty([1-\delta, 1+\delta])$. Since $g \diamond \phi \in C^\infty(\mathbb{R} \setminus \{0\})$, applying Lemma 1 to $g \diamond \phi$, we see that for any $\phi \in C_0^\infty([1-\delta, 1+\delta])$ there exists an n-multiindex α_ϕ such that

$$D^{\alpha_\phi} \widehat{g \diamond \phi}(x) = 0, \qquad x \in \mathbb{R}^n, \pi(x_\alpha) \in (a, b)$$

which implies that

$$C_0^\infty([1-\delta, 1+\delta]) = \bigcup_\alpha W_\alpha \tag{15}$$

where $W_\alpha = \{\phi | \phi \in C_0^\infty([1-\delta, 1+\delta]), D^\alpha \widehat{g \diamond \phi}(x) = 0, \forall x \in \mathbb{R}^n, \pi(x) \in (a, b)\}$. Since $C_0^\infty([1-\delta, 1+\delta])$ is a Frechét space (see, e.g., [14, example 1.46]), it follows from (15) and the Bair's category theory that there exits at least a W_{α_0} such that W_{α_0} contains a nonempty open set, which implies that $W_{\alpha_0} = C_0^\infty([1-\delta, 1+\delta])$. Letting $m = m_0$ be an integer with $me \geq \alpha_0$ and applying Lemma 3 to each element in W_{α_0}, we see that $g \diamond \phi$ can be written in the form (4) with $m = m_0$ on (a, b) for any $\phi \in C_0([1-\delta, 1+\delta])$.

Take $\phi_n \in C_0^\infty([1-\delta, 1+\delta])$ such that $\{g \diamond \phi_n\}_1^\infty$ converges to g in $L^p(a, b)$. Such a sequence of functions is clearly existent. Let X_0 be the subspace spanned by $\{t^k \ln^j |t|, 0 \leq k \leq m_0, 0 \leq j \leq n-1\}$. It is obvious that X_0 is of finite dimension, hence X_0 is closed. According to the previous conclusions proved, $g \diamond \phi_n$ belongs to the subspace X_0. Since $||g \diamond \phi_n - g||_p \to 0 (n \to \infty)$, the closeness of X_0 implies that $g \in X_0$, i.e., $g(t)$ can be written in the form (4) on (a, b), which is a contradiction! The proof is then complete. ■

Lemma 5:

1) Let $g \in L_{loc}^p(\mathbb{R})$, $\phi \in C_0((0, 2))$, K be a compact set in $\mathbb{R}^n$ and $\pi(x) \neq 0, \forall x \in K$. Then $\widehat{g \diamond \phi} \in \overline{\Sigma(g, K)}$ the closure of $\Sigma(g, K)$ in $L^p(K)$;
2) Let $f \in L_{loc}^p(\mathbb{R}^n)$, $\phi \in C_0(\mathbb{R}^n)$ and K be a compact set in $\mathbb{R}^n$. Then $f * \phi \in \overline{\Sigma_0(f, K)}$ the closure of $\Sigma_0(f, K)$ in $L^p(K)$.

Proof (1): From the assumption, we may assume that $\operatorname{supp} \phi \subset [\delta_0, 2-\delta_0]$ with $\delta \in (0, 1)$. From the integrability of $\widehat{g}$, we have

$$\lim_{\delta \to 0^+} \sup_{\substack{s, s' \in [\delta_0, 2-\delta_0] \\ |s-s'| \leq \delta}} \int_K |g(\pi(x)s) - g(\pi(x)s')|^p \, dx = 0. \tag{16}$$

Let $M = \max_s |\phi(s)|$. For any positive integer m, let Δ_m: $\delta_0 = a_{0,m} < a_{1,m} < \ldots < a_{m',m} = 2-\delta_0$ be a partition of $[\delta, 2-\delta_0]$ with $|\Delta_m| = \max_k(a_{k,m} - a_{k-1,m}) < 1/m$. It is clear that $\sum_{k=1}^{m'} \int_{a_{k-1,m}}^{a_{k,m}} g(\pi(x)a_{k.m})\phi(s)\, ds$ belongs to $\Sigma(g, K)$. Thus, it follows that $\widehat{g \diamond \phi} \in \overline{\Sigma(g, K)}$ from (16) and the following inequality:

$$\begin{aligned}
&\left(\int_K \left| \int_{\delta_0}^{2-\delta_0} g(\pi(x)s)\phi(s)\, ds - \sum_{k=1}^{m'} \int_{a_{k-1,m}}^{a_{k,m}} \right.\right. \\
&\qquad \left.\left. \cdot\, g(\pi(x)a_{k.m})\phi(s)\, ds \right|^p dx \right)^{1/p} \\
&\leq M \left(\int_K \left| \sum_{k=1}^{m'} \int_{a_{k-1,m}}^{a_{k,m}} |g(\pi(x)s)\phi(s) \right.\right. \\
&\qquad \left.\left. -\; g(\pi(x)a_{k,m})|^p \, ds)^{1/p} \left(\int_{a_{k-1,m}}^{a_{k,m}} ds \right)^{1-1/p} \right|^p \right)^{1/p} \\
&\leq 2M(1-\delta_0) \sup_{\substack{s, s' \in [\delta_0, 2-\delta_0] \\ |s-s'| \leq 1/m}} \\
&\qquad \left(\int_K |g(\pi(x)s) - g(\pi(x)s')|^p \right)^{1/p}.
\end{aligned} \tag{17}$$

This complete the proof of the conclusion (1). The proof of the conclusion (2) is similar to that of the conclusion (1). And the detail is omitted. ■

Lemma 6: Let $f \in L_{loc}^p(\mathbb{R}^n)$ and k be a nonnegative integer, $e = (1, 1, \ldots, 1)$. Then the following statements are equivalent:

1) $\overline{\Sigma_0(f, K)} = L^p(K)$ for any compact set K in $\mathbb{R}^n$;
2) for any n-multi-index $\alpha \geq ke$, there exists a $\psi \in C_0^\alpha(\mathbb{R}^n)$ such that
$$\int_{\mathbb{R}^n} f(x) D^\alpha \psi(x)\, dx \neq 0; \tag{18}$$
3) for any n-multi-index $\alpha \geq ke$, there exists a $\psi \in C_0^\alpha(\mathbb{R}^n)$ such that (18) is satisfied.

Proof: (2) $\Longrightarrow$ (3) is obvious.

(1) $\Longrightarrow$ (2): Otherwise, there exists an n-multi-index α such that $D^\alpha f = 0$. Let $K \subset \mathbb{R}^n$ be a compact with $Int(K) \neq 0$. Then, for any $\psi \in C_0^\infty(K)$ and $\theta, \lambda \in \mathbb{R}^n$ with $\pi(\lambda) \neq 0$, we have

$$\begin{aligned}
&\int_K f(\lambda^{-1} x + \theta) D^\alpha \psi(x)\, dx \\
&\quad = \pi(\lambda) \int_{\mathbb{R}^n} f(x) D^\alpha \psi(\lambda x - \lambda\theta)\, dx = 0
\end{aligned} \tag{19}$$

which implies that

$$\begin{aligned}
&\int_{\mathbb{R}^n} h(x) D^\alpha \psi(x)\, dx = 0, \qquad \forall \psi \in C_0^\infty(\mathbb{R}^n), \\
&h(x) \in \overline{\Sigma_0(f, K)}.
\end{aligned} \tag{20}$$

Taking into account the assumption (1), (20) implies that $\int_K x^\alpha D^\alpha \psi(x)\, dx = 0, \forall \psi \in C_0^\infty(K)$, which is clearly impossible since $Int(K) \neq 0$. This completes the proof of (1) $\Longrightarrow$ (2).

(3) $\Longrightarrow$ (1): For any n-multi-index β, let α be an n-multi-index with $\alpha \geq ke$ and $\alpha \geq \beta$. According to the assumption, there exists a $\psi \in C_0^\alpha(\mathbb{R}^n)$ such that (18) is satisfied. Let ψ_- be the function defined by $\psi_-(x) = \psi(-x), \forall x \in \mathbb{R}^n$. Then $f * \psi_-$ belongs to $\overline{\Sigma_0(f, K)}$ according to Lemma 5. It is easy to see that (18) implies that $D^\alpha(f * \psi_-)(0) \neq 0$. Thus, by Lemma 5 and Lemma 1, $x^\beta \in \overline{\Sigma_0(f * \psi_-, K)} \subset \overline{\Sigma_0(f, K)}$. Hence, $\Sigma_0(f, K)$ is dense in $L^p(K)$. The proof of (3) $\Longrightarrow$ (1) is complete. ■

IV. The Proof of the Theorem

The Proof of Theorem 1: Necessity: Otherwise, $g(t)$ can be written in the form (4). In this case, it can be verified that $g(t)$ may be written in the form following form:

$$\hat{g}(x) = \sum_{j=1}^n \sum_{k=0}^m c_{jk} x_j^k g_{jk}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \tag{21}$$

where c_{jk} is a constant and each $g_{jk}(x_1, \ldots, x_j, x_{j+1}, \ldots, x_n)$ is an L^p locally integrable function on R^{n-1}. [In fact, it is easy to see that $(\pi(x))^k(\ln \pi(x))^{j-1}$ can be written in the form (21) with $g(t) = t^k(\ln t)^{j-1}$ for any k, j with $0 \le k \le m$ and $1 \le j \le n$, which yield (21).] By (21), a straightforward calculation shows that the general derivative $D^\alpha \hat{g}(x)$ of $\hat{g}(x)$ satisfies the following equation:

$$D^\alpha \hat{g}(x) = 0, \qquad \text{for any } \alpha \ge (m+1)e.$$

Hence, by Lemma 4, there exists a compact set K in $\mathbb{R}^n$ such that $\overline{\Sigma_0(f, K)} \ne L^p(K)$, a contradiction. This completes the proof of the necessity.

Sufficiency: Without loss of generality, we may assume that $\pi(x) \ne 0, \forall x \in K$. Otherwise, we may choose a $\theta \in \mathbb{R}^n$ such that $K' = K + \theta$ possess such a property.

If $g(t)$ can not be written in the form (4) on some interval (a, b) with $0 \notin (a, b)$, then Lemma 4 implies that there exists a $\phi \in C_0^\infty((0, 2))$ such that $g \diamond \phi$ can not be written in the form (4) on (a, b). It is easy to verify that $g \diamond \phi \in L^p_{loc}(\mathbb{R}^n)$. Thus, by Lemma 3 and Lemma 6, $\Sigma_0(g \diamond \phi, K)$ is dense in $L^p(K)$. If $g(t)$ may be written in the form (4) on each interval (a, b) with $0 \notin (a, b)$, $g(t)$ may be written in the form (4) on $(0, \infty)$ and on $(-\infty, 0)$, respectively. Thus, $g(t)$ may be written in the following form:

$$g(t) = \sum_{j=1}^{n}\sum_{k=0}^{m} h_{jk}(t)(\ln |t|)^{j-1}t^k \tag{22}$$

where $h_{jk}(t) = a_{jk}$, if $t > 0$, $h_{jk}(t) = b_{jk}$, if $t < 0$, a_{jk} and b_{jk} are constants.

Cases I: $a_{j0} = b_{j0}$ for any j with $1 \le j \le n$. Then $g_1(t) \triangleq g(t) - \sum_{j=1}^n a_{j0}(t)(\ln |t|)^{j-1}$ is continuous on $\mathbb{R}$ and can not written in the form (3). Thus it from Lemma 3 and Lemma 6 that $\Sigma(g, K)$ is dense in $L^p(K)$.

Cases II: There exists an integer j with $1 \le j \le n$ such that $a_{j0} \ne b_{j0}$. Let j_0 be the largest j with $1 \le j \le n$ and $a_{j0} \ne b_{j0}$. For any $\alpha \ge e$, take $\phi \in C_0^\infty([-1, 1]^n)$ such that $D^{\alpha-e}\phi(0) \ne 0$. Let $Q_I = I_1 \times I_2 \times \ldots \times I_n$, where I_k is [0, 1] or $[-1, 0]$. It is easy to see that

$$\int_{\mathbb{R}^n} \hat{h}_{jk}(x) D^\alpha \left(\phi\left(\frac{x}{\delta}\right)\right) dx = \delta^{n-|\alpha|}\int_{\mathbb{R}^n} \hat{h}_{jk}(\delta x) D^\alpha \phi(x)\, dx \tag{23}$$

$$\int_{\mathbb{R}^n} (\ln |\pi(x)|)^{j-1} D^\alpha \phi(x)\, dx = 0, \qquad j = 1, 2, \ldots, n \tag{24}$$

$$\int_{Q_I} D^\alpha \phi(x)\, dx = (-1)^{\tau(I)} D^{\alpha-e}\phi(0) \ne 0 \tag{25}$$

where $\tau(I)$ is the number of elements in the set $\{k\colon I_k = [0, 1]\}$. Let $\delta > 0$ satisfy $\delta \in (0, e^{-1})$. If $j_0 < j \le n$, then $a_{j0} = b_{j0}$. Thus, from (24), we have

$$\int_{\mathbb{R}^n} \hat{h}_{jk}(x) D^\alpha \left(\phi\left(\frac{x}{\delta}\right)\right) dx = 0. \tag{26}$$

If $1 \le j < j_0$, we have, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^n} \hat{h}_{jk}(x) D^\alpha \left(\phi\left(\frac{x}{\delta}\right)\right) dx\right| &= 2^j \delta^{n-|\alpha|} \max\{|a_{j0}|, |b_{j0}|\} \times |\ln \delta^n|^{j_0-1} \\ &\quad \times \max_{1\le i\le j_0} \int_{\mathbb{R}^n} |\ln \pi(x)|^i |D^\alpha \phi(x)|\, dx.\end{aligned} \tag{27}$$

If $j = j_0$, from (25), we have

$$\begin{aligned}&\int_{Q_I} \hat{h}_{j0}(x) D^\alpha \left(\phi\left(\frac{x}{\delta}\right)\right) dx \\ &= c_{0,I}(-1)^{\tau(I)}(\ln \delta^n)^{j_0-1}\delta^{n-|\alpha|} + \delta^{n-|\alpha|}\sum_{k=1}^{j_0-1}\int_{Q_I} \\ &\quad \cdot \frac{(j_0-1)!}{k!(j_0-1-k)!}(\ln \delta^n)^{j_0-1-k}(\ln \pi(x))^k D^\alpha \phi(x)\, dx\end{aligned}$$

where

$$c_{0,I} = \left(a_{j_0 0}\frac{1+(-1)^{n-\tau(I)}}{2} + b_{j_0 0}\frac{1-(-1)^{n-\tau(I)}}{2}\right)$$

which implies that

$$\begin{aligned}&\left|\int_{\mathbb{R}^n} \hat{h}_{j0}(x) D^\alpha \left(\phi\left(\frac{x}{\delta}\right)\right) dx\right| \\ &\ge \delta^{n-|\alpha|}\Bigg(2^{n-1}|a_{j0} - b_{j0}||\ln \delta^n|^{j_0}|D^{\alpha-e}\phi(0)| \\ &\quad - \max\{|a_{j_0 0}|, |b_{j_0 0}|\} \times 2^{j_0}|\ln \delta^n|^{j_0-1}\max_{1\le i\le j_0} \\ &\quad \cdot \int_{\mathbb{R}^n} |\ln \pi(x)|^i |D^\alpha \phi(x)|\, dx\Bigg).\end{aligned} \tag{28}$$

Moreover it can be verified that, for $1 \le j \le n$, $1 \le k \le m$

$$\begin{aligned}&\left|\int_{\mathbb{R}^n} \hat{h}_{jk}(x) D^\alpha \phi\left(\frac{x}{\delta}\right) dx\right| \\ &\le \delta^{2n-|\alpha|}\max_{1\le j\le n,\, 1\le k\le m}\{|a_{jk}|, |b_{jk}|\}\,|n \ln \delta|^{n-1} \\ &\quad \times 2^n \max\{|D^\alpha \phi(x)|\colon x \in [-1, 1]^n\}.\end{aligned} \tag{29}$$

For any $\alpha \ge e$, from (26)–(29), there exits a $\delta_0 \in (0, e^{-1})$ such that

$$\begin{aligned}&\left|\int_{\mathbb{R}^n} \hat{h}_{j_0 0}(x) D^\alpha \left(\phi\left(\frac{x}{\delta_0}\right)\right) dx\right| \\ &> \sum_{j=1,\, j\ne j_0}^{n}\sum_{k=0}^{m}\left|\int_{\mathbb{R}^n} \hat{h}_{jk}(x) D^\alpha \left(\phi\left(\frac{x}{\delta_0}\right)\right) dx\right|\end{aligned}$$

which yields

$$\int_{\mathbb{R}^n} \hat{g}(x) D^\alpha \left(\phi\left(\frac{x}{\delta_0}\right)\right) dx \ne 0. \tag{30}$$

Hence, $\Sigma_0(\hat{g}, K)$ is dense in $L^p(K)$ according to Lemma 6, which implies $\Sigma_0(g, K)$ is dense in $L^p(K)$. The proof of the sufficiency is then complete. ■

V. CONCLUSION

In this paper, the problem on the approximation of several variables by Sigma–Pi neural networks is investigated. A necessary and sufficient condition for an L^p locally integrable function to be qualified as an active function in Sigma–Pi neural networks is obtained.

ACKNOWLEDGMENT

The authors wish to express their gratefulness to the reviewers for their valuable comments and suggestion on revising this paper.

REFERENCES

[1] Y. Ito, "Approximation of continuous function on $\mathbb{R}^d$ by linear combinations of shifted rotations of a sigmiodal function without scaling," *Neural Networks*, vol. 5, no. 5, pp. 105–115, 1992.

[2] H. N. Mhaskar, "Approximation by superposition of sigmoidal and radial functions," *Advances Appl. Math.*, vol. 13, pp. 350–373, 1992.

[3] M. Leshno *et al.*, "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function," *Neural Networks*, vol. 4, no. 6, pp. 861–867, 1993.

[4] K. Hornik, "Some new results on neural-network approximation," *Neural Networks*, vol. 6, no. 6, pp. 1096–1072, 1993.

[5] B. Lenze, "Note on a density question for neural works," *Numer. Func. Optim.*, vol. 15, no. 7 & 8, pp. 909–913, 1994.

[6] T. P. Chen, H. Chen, and R.-w. Liu, "Approximation capability in $C(\mathbb{R}^n)$ by multilayer feedforward network and problems," *IEEE Trans. Neural Networks*, vol. 6, pp. 25–31, Jan. 1995.

[7] T. P. Chen and H. Chen, "Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems," *IEEE Trans. Neural Networks*, vol. 6, pp. 911–917, July 1995.

[8] J. G. Attali and G. Pagès, "Approximations by a multilayer perceptron: A new approach," *Neural Networks*, vol. 10, no. 6, pp. 1069–1081, 1997.

[9] A. Pinkus, "TDI-subspace of $C(\mathbb{R}^d)$ and some density problems from neural networks," *J. Approx. Theory*, vol. 85, pp. 269–287, 1996.

[10] T. P. Chen and X. W. Wu, "Characteristics of activation function in Sigma–pi neural networks," *Fudan Daxie Xie Bao*, vol. 36, no. 6, pp. 639–644, 1997.

[11] T. P. Chen and H. Chen, "University approximation capability to function of EBF neural networks with arbitrary activation functions," *Circuits Syst. Signal Processing*, vol. 15, no. 5, pp. 671–683, 1996.

[12] T. P. Chen, "Approximation problems in systems from neural networks," *Sci. China*, vol. 36, no. 4, pp. 414–421, 1994.

[13] C. M. Li and T. P. Chen, "Approximation problems in Sigma–Pi neural networks," *Chinese Sci. Bull.*, vol. 41, no. 8, pp. 1072–1074, 1996.

[14] W. Rudin, *Functional Analysis*. New York: McGraw-Hill, 1973.

[15] Y. H. Luo and S. Y. Shen, "Approximation problems of Sigma–Pi neural networks," *Appl. Math. J. Chinese Uni. Ser. A*, vol. 15, no. 1, pp. 107–112, 2000.