

Hopfield Neural Network Based Algorithms for Image Restoration and Reconstruction—Part I: Algorithms and Simulations

Yi Sun, *Member, IEEE*

Abstract—In our previous work, the eliminating-highest-error (EHE) criterion was proposed for the modified Hopfield neural network (MHNN) for image restoration and reconstruction. The performance of the MHNN is considerably improved by the EHE criterion as shown in many simulations. In inspiration of revealing the insight of the EHE criterion, in this paper, we first present a generalized updating rule (GUR) of the MHNN for gray image recovery. The stability properties of the GUR are given. It is shown that the neural threshold set up in this GUR is necessary and sufficient for energy decrease with probability one at each update. The new fastest-energy-descent (FED) criterion is then proposed parallel to the EHE criterion. While the EHE criterion is shown to achieve the highest probability of correct transition, the FED criterion achieves the largest amount of energy descent. In image restoration, the EHE and FED criteria are equivalent. A group of new algorithms based on the EHE and FED criteria is set up. A new measure, the correct transition rate (CTR), is proposed for the performance of iterative algorithms. Simulation results for gray image restoration show that the EHE (FED) based algorithms obtained the best visual quality and highest SNR of recovered images, took much smaller number of iterations, and had higher CTR. The CTR is shown to be a rational performance measure of iterative algorithms and predict quality of recovered images.

Index Terms—Image restoration, neural network, nonlinear detection.

I. INTRODUCTION

THE HOPFIELD neural network (HNN) [1] is a useful model for image restoration and reconstruction [2]–[7]. Zhou *et al.* [2] are the first who proposed the use of the HNN in image restoration and showed the instability of the original HNN when applied to image restoration. They proposed an algorithm (called ZCVJ algorithm in this paper) to ensure the stability of the HNN. They also proposed the use of the simulated annealing (SA) algorithm that allows energy increase with a probability decreasing in time so as to converge to a better solution in stochastic sense. These two algorithms are time-consuming because energy change has to be checked step by step. The modified Hopfield neural network (MHNN) models were proposed by Paik and Katsaggelos in [3] for gray image restoration and by Sun and Yu in [4] for binary

image restoration and reconstruction. The algorithms based on the MHNN ensure network stability without checking energy change step by step. Instead, to guarantee energy decrease a threshold along with the corresponding negative threshold is set up in neural output function, which is much simpler than the check of energy change. Paik and Katsaggelos [3] proposed several separate MHNN algorithms in simultaneous, partially simultaneous, and sequential updating modes. They also proposed an algorithm with bounded time delay of visiting any neurons. However, these algorithms in [3] cannot be derived from a generalized framework. In particular, the threshold in their proposed simultaneous algorithm is much too higher than necessary as shown in this paper.

Sun and Yu [4] are the first who proposed a generalized framework for the MHNN algorithms. In [4], a class of binary MHNN algorithms with group updates was proposed. In retrospect of this class of group-updating algorithms, the thresholds set up in it turn out to be necessary and sufficient for energy descent with probability one. Many ordinary MHNN algorithms with the corresponding tightest thresholds can be found from this class of group-updating algorithms. Among them are the sequential, simultaneous, and N-group simultaneous MHNN algorithms.

Recently, Sun [8] publicized a generalized updating rule (GUR) for the binary MHNN, which can operate in *any* sequence of updating modes with guaranteed stability. Its stability and other properties are proved. The stability is guaranteed because the time-varying threshold of a neuron is adaptively determined by the sum of interconnection weights between this neuron and all other neurons updated at the same time. The ordinary MHNN algorithms (such as the sequential and simultaneous versions) are shown to be instances of the GUR. The class of group-update algorithms proposed in [4] also can be derived from this GUR.

As a reasonable extension of the binary version developed in [8], in this paper, a multi-level version of the GUR is shown to exist. Its stability and fixed-point properties are studied. In particular, for the first time we show that the threshold set up in the GUR for *any* updating mode is necessary and sufficient for the MHNN to reduce its energy with probability one at each iteration.

Although the MHNN algorithms converge to fixed points (except some partially simultaneous versions proposed in [3]), the fixed points may represent recovered images of totally different quality. Hence, starting from the same initial image, an MHNN algorithm may converge to the images of different quality if

Manuscript received June 26, 1997; revised January 19, 2000. The associate editor coordinating the review of this paper and approving it for publication was Dr. Elias S. Manolakis.

The author is with the Department of Electrical Engineering, The City College of the City University of New York, New York, NY 10031 USA (e-mail: yisun@ee-mail.engr.cuny.edu).

Publisher Item Identifier S 1053-587X(00)04937-0.

the state trajectory is different (though given sequence of updating modes). To help the MHNN choose proper neurons to update at each iteration, Sun and Yu [5] proposed the eliminating-highest-error (EHE) criterion. Simulation results in [5] show that if the EHE criterion is used, the network can converge to a more accurate solution in much fewer iterations. In [6] a combination of an analog HNN and the MHNN is applied to blind image restoration. Simulation results show that only if the EHE criterion is applied, can the networks obtain an accurate image and accurate estimate of blur parameters. In [7] it is shown that the EHE criterion can suppress streaks in the restored image, which is a problem commonly existing in the HNN based algorithms and in the conventional algorithms. Recently, the EHE criterion is successfully applied to the code-division multiple-access (CDMA) multiuser detection [13], [14] with promising results.

Motivated by revealing the mystery why the EHE criterion can help the MHNN restore high quality images in much fewer iterations than other MHNN algorithms, in this paper, we find via simple analysis that the EHE criterion can achieve the highest correct transition probability at each neural state transition. Parallel to the EHE criterion, we propose a new criterion, the fastest-energy-descent (FED) criterion, for the MHNN in image restoration and reconstruction. We show that the FED criterion achieves the largest amount of energy descent at each iteration. In image restoration, the EHE criterion is equivalent to the FED criterion. To develop the EHE and FED based algorithm, two new versions of the EHE based algorithm are proposed, and three new algorithms based on the FED criterion are developed in this paper.

In the literature of image restoration and reconstruction, the performance of an algorithm and its recovered image quality are usually evaluated through simulations. The signal-to-noise ratio (SNR) of the recovered image is measured for performance evaluation. The visual quality of the recovered image is also considered. However, the SNR and the image quality of finally recovered image are only the output of the black box—the iterative algorithm, which can not reveal the insight of the performance of the algorithm. Because of this, the comparison of SNR of restored image does not provide insightful benefit to the development of algorithms. For example, for any iterative algorithm existing in literature, it is unknown how good an iteration is. This is apparently due to the lack of a proper performance measure for a particular iteration. Since the goal of all restoration algorithms is to reduce the difference between recovered image and the original image prior to degradation, the probability that an iteration correctly reduces this difference can serve as a measure for the performance of the algorithm and as a predictor for the quality of the finally recovered image. As the statistics used in an iteration are always ready, this probability is feasible to calculate. Since this probability is determined by the operation of the algorithm at a particular iteration, it must reveal the insight of the performance of the algorithm. This probability also depends on the degradation model, statistics of noise, and other statistics, and so reveals how these parameters influence the performance of the algorithm and the recovered image quality. Moreover, this probability also directly depends on the energy function used in the algorithm, and may help us find a better energy function more

suitable for a particular algorithm and image degradation model. In the companion paper [16], we propose such a new performance measure called the correct transition probability (CTP). The CTP formulas of the EHE algorithms, PK algorithm, ZCVJ algorithm, and SA algorithm are then derived and analyzed.

In this paper, correspondingly to the CTP, we propose the correct transition rate (CTR) for the performance evaluation of an iterative algorithm in a simulation. The CTR is defined as the overall rate of correct transitions in the entire run for recovery of an image in the simulation. The CTR and CTP have statistically different meanings. The CTP is a measure for a particular iteration, is hard to estimate in simulation but suitable for theoretical derivation. In contrast, the CTR takes into account the averaged correct transition probability over all transitions and is suitable for estimation in simulation.

Via simulation for gray image restoration, in this paper, performance of six HNN based algorithms is compared. These algorithms are the SA algorithm [2], ZCVJ algorithm [2], PK algorithm [3, Algorithm 2], and three EHE based algorithms (equivalent to the FED based algorithms for image restoration). The SNR and CTR for these six algorithms in various conditions are obtained and compared. The EHE (FED) algorithms demonstrate superior performance over other algorithms in terms of SNR, CTR, and visual quality in all conditions. In the companion paper [16], we show via analysis that the EHE algorithms have much higher CTP than the other algorithms. This confirms the simulation results in this paper.

The rest of the paper is organized as follows. In Section II, a multilevel version of the GUR is presented and its properties are analyzed. In Section III, two forms of the EHE criterion are given. Two new EHE based algorithms are proposed in addition to another fully-presented EHE based algorithm. In Section IV, the FED criterion is proposed and analyzed. Three FED criterion based algorithms are developed. Computation complexity is discussed. The correct transition rate is also defined in this section. In Section V, simulation results are reported. Conclusions are made in Section VI. Proofs are included in the Appendix.

II. A GENERALIZED UPDATING RULE

A. Problem Formulation

In image restoration and reconstruction, a degraded image can be formulated by the following linear equation [9]

$$\mathbf{y} = \mathbf{A}\mathbf{x}^o + \mathbf{n} \quad (1)$$

where $\mathbf{x}^o \in \{0, 1, \dots, L-1\}^n$, $\mathbf{n}, \mathbf{y} \in \mathbb{R}^m$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$. L is the total number of image intensity levels, \mathbf{x}^o is the original image, \mathbf{n} is the noise which is supposed to be white Gaussian with mean zero and covariance matrix $\sigma^2 \mathbf{I}$ in this paper, and \mathbf{y} is the observed blurred noisy image. In image restoration, \mathbf{A} represents a spatially-shift-invariant linear system that can be usually written as a convolution over a small window. Hence, \mathbf{A} is a block Toeplitz matrix in addition to the properties that i) its diagonal elements are identical, ii) its rows (or columns) differ only by permutations, and iii) the summation of elements of any row (or column) is identical. These properties can considerably simplify computations of the HNN based algorithms. In image reconstruction, \mathbf{A} is a projection matrix with $m < n$.

Hence, \mathbf{A} is sparse with a few positive elements and other zero elements. Fundamentally, image restoration and reconstruction are the same problem in the sense that both need solve (1). Many communication problems (e.g. CDMA multiuser detection [13], [14]) also fall into the formulation (1) but with different properties of \mathbf{A} , finite support of \mathbf{x} , and interests. All algorithms discussed in this paper are applicable to (1) of any real matrix \mathbf{A} .

In the literature of image restoration and reconstruction, given an \mathbf{A} , a solution of (1) is obtained by minimizing a squared error plus a regularization term [9]

$$\mathbf{x}: \min_{\mathbf{x} \in \{0, \dots, L-1\}^n} \{ \|\mathbf{Ax} - \mathbf{y}\|^2 + \lambda \|\mathbf{Dx}\|^2 \} \quad (2)$$

where \mathbf{D} is a difference operator and $\lambda \geq 0$. The purpose to add the regularization term is to obtain a smoother solution when the noise is severe. The value of λ determines the tradeoff of the needs for smoothness and fineness of recovered image.

By defining an energy function

$$E(\mathbf{x}) = -\frac{1}{2} \mathbf{x}^T (\mathbf{W} + \lambda \mathbf{G}) \mathbf{x} + \Theta^T \mathbf{x} \quad (3)$$

where $\mathbf{W} = -\mathbf{A}^T \mathbf{A}$, $\mathbf{G} = -\mathbf{D}^T \mathbf{D}$ and $\Theta = -\mathbf{A}^T \mathbf{y}$, (2) can be rewritten as

$$\mathbf{x}: \min_{\mathbf{x} \in \{0, \dots, L-1\}^n} E(\mathbf{x}). \quad (4)$$

To solve (4) by using the HNN, let $E(\mathbf{x})$, \mathbf{x} , $\mathbf{W} + \lambda \mathbf{G}$, and Θ represent the energy function, network state, interconnection matrix, and neural bias vector of the HNN, respectively. Take the first and second derivatives of $E(\mathbf{x})$ with respect to \mathbf{x} , we have $\nabla E(\mathbf{x}) = -(\mathbf{W} + \lambda \mathbf{G})\mathbf{x} + \Theta$ and $\nabla^2 E(\mathbf{x}) = -(\mathbf{W} + \lambda \mathbf{G})$. Denote $\Delta \mathbf{x}(k) = \mathbf{x}(k+1) - \mathbf{x}(k)$, $\Delta E(k) = E[\mathbf{x}(k+1)] - E[\mathbf{x}(k)]$, and the negative gradient vector of the energy function at time k by

$$\mathbf{h}(k) = -\nabla E(\mathbf{x}(k)) = (\mathbf{W} + \lambda \mathbf{G})\mathbf{x}(k) - \Theta. \quad (5)$$

The i th element $h_i(k)$ of $\mathbf{h}(k)$ is also the i th-neuron input. Then the change of energy function because of the change of \mathbf{x} can be expressed by Taylor's expansion as

$$\Delta E(k) = -\Delta \mathbf{x}^T(k) (\mathbf{h}(k) + \frac{1}{2} (\mathbf{W} + \lambda \mathbf{G}) \Delta \mathbf{x}(k)). \quad (6)$$

The updating rule of the original HNN is $\Delta \mathbf{x}(k) = \text{sgn}[\mathbf{h}(k)]$ that ensures $-\Delta \mathbf{x}^T(k) \mathbf{h}(k) < 0$ for $\forall \Delta \mathbf{x}(k) \neq \mathbf{0}$. However, in the image restoration and reconstruction, the matrix $\mathbf{W} + \lambda \mathbf{G}$ is nonpositive definite. Ensuring $-\Delta \mathbf{x}^T(k) \mathbf{h}(k) < 0$ does not ensure $\Delta E(k) < 0$, which causes the instability of the original HNN. In order to guarantee $\Delta E(k) < 0$, the term $\Delta \mathbf{x}^T(k) (\mathbf{W} + \lambda \mathbf{G}) \Delta \mathbf{x}(k)$ in (6) has to be properly considered in the neural output functions. This requires that the neural state transition take place only if the neural input $h_i(k)$ is larger than a threshold or smaller than the negative threshold, thus yielding the MHNN.

B. A Generalized Updating Rule for Gray Image

Although the MHNN is ordinarily considered to operate in the sequential, simultaneous, and group-simultaneous updating modes, the MHNN can operate in any updating modes provided that the threshold of each neuron is properly set up to ensure stability. A generalized updating rule (GUR) of the MHNN in the binary case is presented in [8]. In this paper, we extend this GUR to the L -level case for gray image restoration and reconstruction.

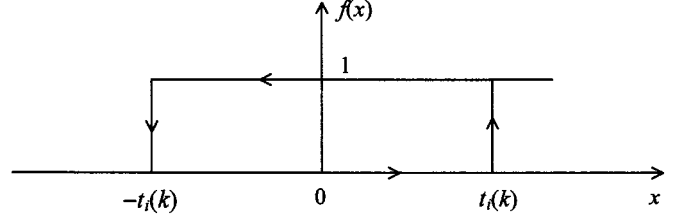


Fig. 1. The output function of the i th neuron at time k . The threshold $t_i(k)$ adaptively changes with neural index set $L(k)$ to guarantee stability of the network operating in any sequence $L(k)$ of updating modes.

GUR: Given a sequence of update modes $L(k) \subseteq \{1, \dots, n\}$ for $k \geq 0$, $\mathbf{x}(k)$ is updated by

$$x_i(k+1) = \begin{cases} x_i(k) - 1, & \text{if } i \in L(k), h_i(k) < -t_i(k) \\ & \text{and } x_i(k) > 0 \\ x_i(k) + 1, & \text{if } i \in L(k), h_i(k) > t_i(k) \\ & \text{and } x_i(k) < L - 1 \\ x_i(k), & \text{otherwise} \end{cases} \quad (7)$$

where $t_i(k)$ is the i th-neuron threshold at time k ,

$$t_i(k) = \frac{1}{2} \sum_{j \in L(k)} |w_{ij} + \lambda g_{ij}|. \quad (8)$$

The update stops at k_e when no neuron is updateable according to (7) for any $k \geq k_e$. \square

$L(k)$ for $k \geq 0$ determines the sequence of updating modes. If $i \in L(k)$, $x_i(k+1)$ will be updated and possibly be changed at time k . $L(k)$ for $k \geq 0$ also determines the threshold $t_i(k)$. As shown in (8), the threshold $t_i(k)$ depends on the interconnection weights among the neurons that are updated at the same time k (whose indices are in $L(k)$). The output function of the i th neuron at time k specified by the GUR is shown in Fig. 1. Clearly, if $t_i(k) = 0$ for $\forall i \in \{1, \dots, n\}$ and $\forall k \geq 0$, the neural output function of the MHNN becomes the sign function and the GUR becomes the updating rule of the original HNN.

We give stability properties of the GUR as follows without proof. With some modification, the reader can obtain the proof in the similar way provided in [8] for the binary case.

Theorem 1: i) For any $k \geq 0$, it is guaranteed that

$$\Delta E(k) = E[\mathbf{x}(k+1)] - E[\mathbf{x}(k)] \leq 0 \quad (9)$$

where the equality holds if and only if $\mathbf{x}(k+1) = \mathbf{x}(k)$. ii) After a finite number of updates, the GUR converges to a fixed point. iii) If the GUR converges to a fixed point \mathbf{x}^e after k_e updates (i.e., $\mathbf{x}(k) = \mathbf{x}^e$ for all $k \geq k_e$), $\bigcup_{k \geq k_e} L(k) = \{1, \dots, n\}$, and \mathbf{x}^e is not bounded, then the gradient of the energy function at \mathbf{x}^e is upper bounded by

$$\|\mathbf{h}^e\|_1 \leq \sum_{i=1}^n \min_{k \geq k_e, i \in L(k)} \{t_i(k)\} \quad (10)$$

\square

In Theorem 1, a fixed point \mathbf{x}^e is said to be bounded if there is an $i \in L(k)$ for $k \geq k_e$ such that $h_i^e(k) < -t_i(k)$ and $x_i^e(k) = 0$, or $h_i^e(k) > t_i(k)$ and $x_i^e(k) = L - 1$. Theorem 1 guarantees that the GUR monotonically reduces energy $E(\mathbf{x})$ and finally converges to a fixed point in a finite number of updates.

Many ordinary MHNN algorithms are instances of the GUR. For example, if $L(k) = \{(k \bmod N) + 1\}$, the GUR be-

comes the PK algorithm [3, Algorithm 2], a sequential MHNN algorithm. The neural threshold in (8) becomes

$$t_i^{seq} \equiv \frac{1}{2}|w_{ii} + \lambda g_{ii}| \quad (11)$$

and the corresponding upper bound of residual in (10) becomes

$$\|\mathbf{h}^e\|_1 \leq \frac{1}{2} \text{trace}(|\mathbf{W} + \lambda \mathbf{G}|) \quad (12)$$

which were already obtained by Paik and Katsaggelos in [3] as well as by Sun and Yu in [4].

By specifying $L(k) = \{1, \dots, n\}$, the GUR becomes the simultaneous MHNN algorithm. Its corresponding threshold of (8) becomes

$$t_i^{sim} \equiv \frac{1}{2} \sum_{j=1}^n |w_{ij} + \lambda g_{ij}| \quad (13)$$

and the upper bound of residual in (10) at a fixed point becomes

$$\|\mathbf{h}^e\|_1 \leq U^{sim} \equiv \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n |w_{ij} + \lambda g_{ij}|. \quad (14)$$

This simultaneous MHNN algorithm with the corresponding neural threshold and residual bound was first found by Sun and Yu in [4].

Paik and Katsaggelos also proposed a simultaneous algorithm [3, Algorithm 4] in which all neurons also update at each iteration. However, in their simultaneous algorithm, the neural threshold

$$t_i^{PK} \equiv \frac{1}{2} \|\mathbf{W} + \lambda \mathbf{G}\|_2 = \frac{1}{2} \sqrt{\sum_{i=1}^n \sum_{j=1}^n |w_{ij} + \lambda g_{ij}|^2} \quad (15)$$

is unnecessarily high and so its corresponding upper bound of residual [3]

$$\|\mathbf{h}^e\|_1 \leq U^{PK} = \frac{n}{2} \|\mathbf{W} + \lambda \mathbf{G}\|_2 \quad (16)$$

is too loose. To show this, in the Appendix, we prove the following inequalities i) if rows of $\mathbf{W} + \lambda \mathbf{G}$ differ only by permutations such as in image restoration,

$$t_i^{sim} \leq t_i^{PK} \quad (17)$$

and ii) for any real matrix $\mathbf{W} + \lambda \mathbf{G}$, including both image restoration and reconstruction

$$U^{sim} \leq U^{PK} \quad (18)$$

where the equalities in (17) and (18) hold if and only if $|w_{ij} + \lambda g_{ij}| = b$ (constant) for $\forall i, j \in \{1, \dots, n\}$. The condition $|w_{ij} + \lambda g_{ij}| = b$ implies an extremely severe degradation system or bad energy function because in this condition, say $\lambda = 0$, no information but mean of the original image is provided by \mathbf{y} . In practical image restoration and reconstruction problem, the differences between t_i^{sim} and t_i^{PK} in (17) and between U^{sim} and U^{PK} in (18) can be very large. For example, consider the conditions used in the simulation reported in Section V with the blurring function (30) of $c = 2$, difference operator (31), and $\lambda = 0.1$. We have $t_i^{sim} = 0.5 \ll t_i^{PK}(k) \cong 15$, and $U^{sim} = 8192 \ll U^{PK} \cong 284118$. They differ by a factor about 30. Obviously, The neural threshold in the simultaneous algorithm proposed by Paik and Katsaggelos in [3] and its corresponding upper bound of residual are much too looser than those of the

simultaneous MHNN algorithm derived from the GUR in this paper.

The result that the neural threshold t_i^{sim} in the simultaneous updating mode derived from the GUR is much tighter than t_i^{PK} is not by chance. For *any* updating mode, the threshold specified by (8) is the tightest as shown by the following theorem proved in the Appendix.

Theorem 2: The neural threshold $t_i(k)$'s specified by (8) are necessary and sufficient for the GUR with probability one to reduce energy E at each nonzero update for an arbitrary $\mathbf{W} + \lambda \mathbf{G} \in \mathbb{R}^{n \times n}$ and $\forall L(k) \in \{1, \dots, n\}$. \square

Theorem 2 is applicable to the class of group-updating algorithms proposed in [4] for these group-updating algorithms can be derived from the GUR.

If the GUR operates in a sequential mode, Theorem 2 is in the sense of *any* interconnection matrix as shown by the following corollary.

Corollary 1: The neural thresholds specified by (11) are necessary and sufficient for the GUR operating in a sequential model to reduce energy E with probability one at each nonzero update for any $\mathbf{W} + \lambda \mathbf{G} \in \mathbb{R}^{n \times n}$ and $\forall L(k) \in \{1, \dots, n\}$. \square

C. Wide-Sense Sequential Updating Modes

By specifying a new sequence of $L(k)$ for $k \geq 0$, one brings out a new MHNN algorithm with threshold specified by (8) and its performance measured by (10). There are infinitely many sequences of $L(k)$, $k \geq 0$, corresponding to infinitely many sequences of updating modes. Among the infinitely many updating modes, the class of wide-sense sequential updating modes is most interesting.

Definition 1: If there exists an integer $k_1 \geq 0$ such that $L(k)$ for any $k \geq k_1$ contains only one element of $\{1, \dots, n\}$, $L(k)$ for $k \geq 0$ defines a wide-sense sequential updating mode. \square

The wide-sense sequential updating modes have the following properties. Corollary 2 follows Theorem 1.

Corollary 2: If and only if the GUR operates in a wide-sense sequential updating mode, the upper bound of $\|\mathbf{h}^e\|_1$ achieves the minimum value given by (12). \square

Theorem 3: If the GUR operates in a wide-sense sequential updating mode and $\bigcup_{k \geq k_1} L(k) = \{1, \dots, n\}$, the fixed point to which the GUR converges is a local minimum point of E , and vice versa. \square

Theorem 3 in the case of the ordinary sequential updating mode is proved by Paik and Katsaggelos in [3, Theorem 2]. Since for $k \geq k_1$ the GUR operates in the ordinary sequential updating mode and $\bigcup_{k \geq k_1} L(k) = \{1, \dots, n\}$, Theorem 3 is proved. Theorem 3 is an extension of their Theorem 2 from the ordinary sequential updating mode to the wide-sense sequential updating mode.

All the EHE and FED based algorithms proposed in this paper belong to the class of wide-sense sequential updating modes.

III. EHE CRITERION AND EHE BASED ALGORITHMS

Among the wide-sense sequential updating modes, there are still infinitely many sequences $L(k)$, $k \geq 0$, each of which leads the GUR to converge to fixed points of different image quality. The EHE and FED criteria presented in this paper can help the

GUR adaptively find better sequences of updating modes based on currently recovered image so as to converge to an image of better quality.

A. EHE Criterion

Basically, the EHE criterion is applied to the sequential updating mode in which only one neuron is updated at each time. If several neurons are updated at one time, these neurons are supposed to be located at spatially different areas without correlation. In the sequential updating mode, $L(k)$ contains only one element of $\{1, \dots, n\}$. The neural threshold in the sequential updating mode is given by (11) that is independent of k .

Definition 2: In terms of the GUR, the index set of updateable neurons in the sequential updating mode is defined as

$$J(k) = \{i | h_i(k)/t_i^{seq} < -1 \quad \text{and} \quad x_i(k) > 0, \\ \text{or} \quad h_i(k)/t_i^{seq} > 1 \quad \text{and} \quad x_i(k) < L - 1\}. \quad (19)$$

□

The state transition of any neuron whose index is in $J(k)$ reduces the energy E in terms of the GUR. With some probability, the reduction of energy function reduces the distance between the original image and the network state $\|\Delta \mathbf{x}^o(k)\| \equiv \|\mathbf{x}^o - \mathbf{x}(k)\|$, which is the object for image restoration and reconstruction. This probability is usually less than one due to the facts that the noise exists and that image is blurred (i.e., the interference from other pixels in the neighborhood of the pixel of interest exists because of the blur). At any time, the state transition of any updateable neuron decreases $\|\Delta \mathbf{x}^o(k)\|$ with certain probability and meanwhile increases $\|\Delta \mathbf{x}^o(k)\|$ with the complement probability. It is desirable to change the state of a neuron whose state transition has the highest probability that this transition is correct in terms of reducing $\|\Delta \mathbf{x}^o(k)\|$. This probability is defined as the correct transition probability (CTP) in [16]. As shown by the GUR, in all MHNN based algorithms, the statistic $\mathbf{h}(k)$ determines the neural state transitions at time k . The statistic $\mathbf{h}(k)$ also provides information about the correct transition probability at time k .

Notice (1), we can rewrite (5) as

$$\mathbf{h}(k) = (\mathbf{W} + \lambda \mathbf{G})\mathbf{x}(k) + \mathbf{A}^T(\mathbf{A}\mathbf{x}^o + \mathbf{n}) \\ = -(\mathbf{W} + \lambda \mathbf{G})\Delta \mathbf{x}^o(k) + \lambda \mathbf{G}\mathbf{x}^o + \mathbf{A}^T\mathbf{n} \quad (20)$$

or equivalently, for $i = 1, \dots, n$

$$h_i(k) = -(w_{ii} + \lambda g_{ii})\Delta x_i^o(k) - \sum_{j=1, j \neq i}^n (w_{ij} + \lambda g_{ij}) \\ \cdot \Delta x_j^o(k) + \lambda \sum_{j=1}^n g_{ij}x_j^o + \sum_{j=1}^n a_{ji}n_j \quad (21)$$

where $\Delta x_i^o(k) = x_i^o - x_i(k)$ is the difference between the i th-pixel intensity of the original image and the i th-neuron state at time k . Notice that $t_i^{seq} = |w_{ii} + \lambda g_{ii}|/2 = -(w_{ii} + \lambda g_{ii})/2$. Hence

$$\frac{h_i(k)}{t_i^{seq}} = 2\Delta x_i^o(k) + 2 \sum_{j=1, j \neq i}^n \frac{w_{ij} + \lambda g_{ij}}{w_{ii} + \lambda g_{ii}} \Delta x_j^o(k) \\ - 2 \sum_{j=1}^n \frac{\lambda g_{ij}}{w_{ii} + \lambda g_{ii}} x_j^o - 2 \sum_{j=1}^n \frac{a_{ji}}{w_{ii} + \lambda g_{ii}} n_j. \quad (22)$$

Since $\Delta x_j^o(k)$ is distributed over $\{-L + 1, \dots, L - 1\}$, it is reasonable to assume $\Delta x_j^o(k)$ has a distribution symmetric to zero, and so the second term of (22) has zero mean. x_j^o is the original image that is usually smooth in a small window defined by \mathbf{G} . Since \mathbf{D} is designed such that the summation of g_{ij} is equal to zero, the third term has zero mean but may have large variance as λ is large. The fourth term of (22) is a Gaussian random variable with zero mean.

It is clear that if i) the noise is weak, ii) the original image at the neighborhood of the pixel of interest is smooth or λ is small, and iii) the following equalities are satisfied,

$$|w_{ii} + \lambda g_{ii}| > |w_{ij} + \lambda g_{ij}|, \quad i, j = 1, \dots, n, \quad j \neq i \quad (23)$$

the statistic in (22) is dominated by $\Delta x_i^o(k)$ in statistical sense. For the i th updateable neuron, if $\Delta x_i^o(k) > 0$, then $h_i(k)/t_i^{seq} > 0$ with some probability larger than 0.5. If $\Delta x_i^o(k) < 0$, then $h_i(k)/t_i^{seq} < 0$ with some probability larger than 0.5. The larger is the $|h_i(k)/t_i^{seq}|$, the larger is the probability that $\Delta x_i^o(k)$ and $h_i(k)/t_i^{seq}$ have the same sign. It is an obviously good strategy to update first the neuron that has the highest probability that $\Delta x_i^o(k)$ and $h_i(k)/t_i^{seq}$ have the same sign among all updateable neurons, thus resulting in and explaining the effectiveness of the EHE criterion [5].

EHE criterion A: If one neuron is updated at time k , i.e., $L(k)$ contains one element, this neuron should be the one whose relative absolute input $|h_i(k)/t_i^{seq}|$ is the largest among all updateable neurons. □

EHE criterion B: If N neurons are updated at time k , i.e., $L(k)$ contains N elements, these N neurons should be those whose relative absolute inputs $|h_i(k)/t_i^{seq}|$ are the N largest ones among all updateable neurons. □

For image restoration, the matrix $-(\mathbf{W} + \lambda \mathbf{G})$ corresponds to the autocorrelation function of two summed spatially-shift-invariant point spread functions. Since the autocorrelation function usually has a larger value at the origin than at any other points, (23) is true in general. When the noise is weak, it is clear that the less severely the image is blurred, the larger $|w_{ii}|$ is, and then the more effective the EHE criterion is. In image reconstruction, the projection matrix \mathbf{A} is usually a sparse matrix: a few of its elements are larger than zero and others are zeros. This implies that (23) is true. Hence, the EHE criterion can be applied to both image restoration and reconstruction.

The EHE criterion is most effective when image data is noise-free and $\lambda = 0$. When the noise is strong, condition i) is violated. To suppress noise, a larger λ must be utilized, thus further violating condition ii). In other words, when the noise is stronger, the statistic $h_i(k)/t_i^{seq}$ has a smaller chance to be dominated by $\Delta x_i^o(k)$. Hence, in general, the strong noise degrades the effectiveness of the EHE criterion.

Although a large value of λ can suppress the noise, the energy function E itself suggests that an extremely large λ imply a solution of constant image. Hence, in general, an optimal λ exists for an algorithm to compromise the smoothness and fineness of the recovered image. An optimal λ also exists for the EHE based algorithms. This is confirmed in simulation.

B. EHE Based Algorithms

In this paper, we present three EHE based algorithms. The algorithm based on EHE criterion A has the following form.

Algorithm 1:

- 0) Let $t_i = t_i^{seq}, i = 1, \dots, n$. Given $\mathbf{x}(0)$ and $k = 0$.
- 1) Find $J(k)$ according to (19). If $J(k) = \emptyset$, terminate.
- 2) Find an i in $J(k)$ such that $|h_i(k)|/t_i \geq |h_j(k)|/t_j$ for $\forall j \in J(k)$. Let $L(k) = \{i\}$.
- 3) Compute $x_i(k+1)$ for $i \in L(k)$ according to (7); update $\mathbf{h}(k)$.
- 4) $k := k+1$; Go to step 1). \square

The following two algorithms are based on EHE criterion B.

Algorithm 2:

- 0) Given $N \leq n$, $\mathbf{x}(0)$ and $k = 0$.
- 1) Let $t_i = 1/2 \sum_{j \in L_1} |w_{ij} + \lambda g_{ij}|$ for $i = 1, \dots, n$ where $L_1 = \{j \in \{1, \dots, n\} \mid |w_{ij} + \lambda g_{ij}| \text{ is one of the } N \text{ largest } |w_{il} + \lambda g_{il}|, l = 1, \dots, n\}$.
- 2) Find $J(k)$ from (19) with t_i^{seq} being replaced by t_i in $J(k)$. If $J(k) = \emptyset$ and $N = 1$, terminate.
- 3) If the total number of elements in $J(k)$ is larger than N , find $L(k) \subset J(k)$ such that $|h_i(k)|/t_i \geq |h_j(k)|/t_j$ for $\forall i \in L(k)$ and $\forall j \in J(k) \setminus L(k)$, where $L(k)$ contains N numbers. Otherwise, let $L(k) = J(k)$.
- 4) Compute $x_i(k+1)$ for $i \in L(k)$ according to (7); update $\mathbf{h}(k)$.
- 5) $k := k+1$. $N := \lfloor N/\beta \rfloor$ where $\beta > 1$ and $\lfloor x \rfloor$ is the integer part of x ; if $N < 1$, $N := 1$; Go to step 1). \square

Algorithm 3:

- 0) Let $t_i = t_i^{seq}, i = 1, \dots, n$. Given $\delta \in (0, 1]$, $\Delta\delta \in (0, 1]$ and $\mathbf{x}(0)$. Let $k = 0$.
- 1) Find $J(k)$ according to (19). If $J(k) = \emptyset$, terminate.
- 2) Find an i in $J(k)$ such that $|h_i(k)|/t_i \geq |h_j(k)|/t_j$ for $\forall j \in J(k)$. If $\delta = 1$, let $L(k) = \{i\}$; else $L(k) = \{j \in J(k) \mid |h_j(k)|/t_j \geq \delta |h_i(k)|/t_i\}$.
- 3) Compute $x_i(k+1)$ for $i \in L(k)$ according to (7); update $\mathbf{h}(k)$.
- 4) $k := k+1$; $\delta := \delta + \Delta\delta$. If $\delta > 1$, $\delta := 1$; Go to 1). \square

Before Algorithm 3 approaches the sequential updating mode, the increase of energy is possible in these two algorithms. However, we note that only those neurons whose absolute input values are largest are chosen to update according to the EHE criterion. Moreover, because the neurons that have the largest absolute input values are usually spatially separate,

the neurons that Algorithms 3 chooses to update in a few initial iterations are spatially uncorrelated. Hence, the probability of energy increase is small. This is fundamentally different from the SA algorithm that allows a large probability to increase energy. Algorithm 3 finally operates in the sequential updating mode in which energy decrease is guaranteed with the lowest threshold $t_i = t_i^{seq}$. Before the Algorithm 2 operates in the sequential updating mode, the computation of t_i at every iteration is time-consuming. To save time, the threshold can be fixed by $t_i = t_i^{seq}$. In this case, Algorithm 2 can have the similar behavior to Algorithm 3, as discussed above.

Before Algorithms 2 and 3 operate in the sequential updating mode, many neurons are updated in each iteration. Since only those neurons of largest inputs are updated at each iteration, the multi-neuron iterations do not reduce performance in Algorithms 2 and 3. Meanwhile, the multi-neuron iterations considerably reduce comparisons and so reduce computation complexity.

Since Algorithms 1, 2, and 3 all operate in the wide-sense sequential updating mode, Corollary 2 and Theorem 3 are applicable to them.

IV. FED CRITERION, FED BASED ALGORITHMS, AND CORRECT TRANSITION RATE

A. FED Criterion

Suppose the i th-neuron state changes by one at time k , i.e., $\Delta\mathbf{x}(k) = (0, \dots, 0, \Delta x_i(k), 0, \dots, 0)^T$, where $\Delta x_i(k) = -1$ or 1 . According to (6), the energy change due to this neural state transition is

$$\begin{aligned} \Delta E(k) &= -\Delta x_i(k) h_i(k) - \frac{1}{2}(w_{ii} + \lambda g_{ii}) \\ &= -\Delta x_i(k) h_i(k) + t_i^{seq}. \end{aligned} \quad (24)$$

In terms of (24), among all updateable neurons, the amount of energy change due to a neural state transition is different. Since the criterion to find a solution in (4) is to minimize the energy function, a good strategy is to update first the neuron whose state transition leads to the largest amount of energy decrease. Based on this strategy, in this paper, we propose the fastest-energy-descent (FED) criterion as follows.

FED Criterion A: If one neuron is updated at time k , this neuron should be the one whose $|h_i(k)| - t_i^{seq}$ is the largest among all updateable neurons. \square

FED Criterion B: If N neurons are updated at time k , the N neurons should be those whose $(|h_i(k)| - t_i^{seq})$'s are the N largest ones among all updateable neurons. \square

Note that according to the GRU, it is guaranteed that $\Delta x_i(k) h_i(k) > 0$ for any nonzero update. By means of (24), we can easily prove the following theorem.

Theorem 4: In the sequential updating mode, the FED criterion achieves the largest amount of energy decrease at each update. \square

Given the same amount of energy to decrease in a processing of image recovery, the FED criterion achieves the fastest convergence rate or needs the smallest number of iterations.

Although the EHE and FED criteria are differently motivated, they are equivalent in the following case.

Theorem 5: If the diagonal elements of the interconnection matrix $\mathbf{W} + \lambda\mathbf{G}$ are identical, the EHE and FED criteria are equivalent. \square

Theorem 5 is proved by noticing that the largest $|h_i(k)|/t_i^{seq}$ is equivalent to the largest $|h_i(k)| - t_i^{seq}$ if the interconnection matrix $\mathbf{W} + \lambda\mathbf{G}$ has identical diagonal elements. Since in image restoration the interconnection matrix has identical diagonal elements, the following corollary follows.

Corollary 3: For image restoration, in the sequential updating mode the EHE criterion achieves the largest amount of energy descent at each update. \square

Theorem 4 and Corollary 3 are directly applicable to FED criterion A. In FED criterion B, if those neurons whose $(|h_i(k)| - t_i^{seq})$'s are the largest among all updateable neurons are uncorrelated, Theorem 4 and Corollary 3 are also applicable to FED criterion B. In image restoration, the blurring system can usually be expressed as a convolution over a small window. In this case, the pixels located outside the window centered at the pixel of interest are uncorrelated to the pixel of interest. Hence, in most cases, Theorem 4 and Corollary 3 are also applicable to FED criterion B for image restoration.

B. FED Based Algorithms

The algorithm based on FED criterion A has the following form.

Algorithm 4:

- 0) Let $t_i = t_i^{seq}, i = 1, \dots, n$. Given $\mathbf{x}(0)$ and $k = 0$.
- 1) Find $J(k)$ according to (19). If $J(k) = \emptyset$, terminate.
- 2) Find an i in $J(k)$ such that $|h_i(k)| - t_i \geq |h_j(k)| - t_j$ for $\forall j \in J(k)$. Let $L(k) = \{i\}$.
- 3) Compute $x_i(k+1)$ for $i \in L(k)$ according to (7); update $\mathbf{h}(k)$.
- 4) $k := k+1$; Go to step 1). \square

The following two algorithms are based on FED criterion B.

Algorithm 5:

- 0) Given $N \leq n$, $\mathbf{x}(0)$ and $k = 0$.
- 1) Let $t_i = 1/2 \sum_{j \in L_1} |w_{ij} + \lambda g_{ij}|$ for $i = 1, \dots, n$ where $L_1 = \{j \in \{1, \dots, n\} | |w_{ij} + \lambda g_{ij}| \text{ is one of the } N \text{ largest } |w_{il} + \lambda g_{il}|, l = 1, \dots, n\}$.
- 2) Find $J(k)$ from (19) with t_i^{seq} being replaced by t_i in $J(k)$. If $J(k) = \emptyset$ and $N = 1$, terminate.
- 3) If the total number of elements in $J(k)$ is larger than N , find $L(k) \subset J(k)$ such that $|h_i(k)| - t_i \geq |h_j(k)| - t_j$ for $\forall i \in L(k)$ and $\forall j \in J(k) \setminus L(k)$, where $L(k)$ contains N numbers. Otherwise, let $L(k) = J(k)$.
- 4) Compute $x_i(k+1)$ for $i \in L(k)$ according to (7); update $\mathbf{h}(k)$.
- 5) $k := k+1$. $N := \lfloor N/\beta \rfloor$ where $\beta > 1$ and $\lfloor x \rfloor$ is the integer part of x ; if $N < 1$, $N := 1$; Go to step 1). \square

Algorithm 6:

- 0) Let $t_i = t_i^{seq}, i = 1, \dots, n$. Given $\delta \in [0, 1]$, $\Delta\delta \in (0, 1]$ and $\mathbf{x}(0)$. Let $k = 0$.
- 1) Find $J(k)$ according to (19). If $J(k) = \emptyset$, terminate.
- 2) Find an i in $J(k)$ such that $|h_i(k)| - t_i \geq |h_j(k)| - t_j$ for $\forall j \in J(k)$. If $\delta = 1$, let $L(k) = \{i\}$; else $L(k) = \{j \in J(k) | |h_j(k)| - t_j \geq \delta(|h_i(k)| - t_i)\}$.
- 3) Compute $x_i(k+1)$ for $i \in L(k)$ according to (7); update $\mathbf{h}(k)$.
- 4) $k := k+1$; $\delta := \delta + \Delta\delta$. If $\delta > 1$, $\delta := 1$; Go to 1). \square

Algorithm 6 finally operates in the sequential updating mode in which energy decrease is guaranteed with $t_i = t_i^{seq}$. Since Algorithms 4, 5, and 6 all operate in the wide-sense sequential updating mode, Corollary 2 and Theorem 3 are applicable to them.

C. Computation Complexity

Since the GUR updates neural states based on the current neural input vector $\mathbf{h}(k)$ and network state $\mathbf{x}(k)$, in all the EHE and FED algorithms every iteration consists of three steps of computation: 1) compare $\mathbf{h}(k)$ to find the neurons to update; 2) update $\mathbf{x}(k)$; and 3) update $\mathbf{h}(k)$. For ordinary, non-EHE, and non-FED MHN algorithms, computation of each iteration consists of only steps 2) and 3). Hence, computation of step 1) accounts for the extra operation required by the EHE and FED criteria.

In step 1), both the EHE and FED algorithms need n comparisons. A winner-take-all circuit can efficiently achieve the comparison if the algorithms are implemented by neural hardware. In addition, the EHE algorithms need n multiplications (consider division as multiplication). The additional bias— t_i in the FED algorithms does not take into account of the operation because it can be set up at the initial.

Consider step 3). Suppose that in a nonsequential updating mode, $L(k)$ has $|L(k)| > 1$ elements and so more than one neurons will be updated at time k . Since the update of $\mathbf{h}(k)$ and $\mathbf{x}(k)$ for the $|L(k)|$ neurons can be decomposed of $|L(k)|$ single-neuron updates, in what follows, we consider the sequential updating mode and $L(k) = \{i\}$. According to (5), $\mathbf{h}(k)$ can be updated due to the nonzero neural update $\Delta\mathbf{x}(k) = \mathbf{x}(k+1) - \mathbf{x}(k) \neq \mathbf{0}$ as follows,

$$\mathbf{h}(k+1) = \mathbf{h}(k) + (\mathbf{W} + \lambda\mathbf{G})\Delta\mathbf{x}(k). \quad (25)$$

In the sequential updating mode, $\Delta\mathbf{x}(k) = \Delta x_i(k)\mathbf{e}_i$ where $\Delta x_i(k) \in \{-1, 1\}$ and \mathbf{e}_i is the i th coordinate vector. Hence,

$$\mathbf{h}(k+1) = \mathbf{h}(k) + \Delta x_i(k)(\mathbf{w}_i + \lambda\mathbf{g}_i) \quad (26)$$

where $\mathbf{w}_i + \lambda\mathbf{g}_i$ is the i th column vector of matrix $\mathbf{W} + \lambda\mathbf{G}$. In image restoration and reconstruction, most elements of \mathbf{W} and \mathbf{G} are zeros. By defining an index set

$$\Psi_{wgi} = \{j | w_{ij} + \lambda g_{ij} \neq 0 \text{ for } j = 1, \dots, n\} \quad (27)$$

(26) can be expressed as

$$h_j(k+1) = h_j(k) + \Delta x_i(k)(w_{ij} + \lambda g_{ij}), \quad j \in \Psi_{wgi}. \quad (28)$$

Hence, step 3) is composed of $|\Psi_{wgi}|$ additions or subtractions depending on the sign of $\Delta x_i(k)$.

Step 2) needs one addition or subtraction.

In total, for every single-neuron update, the GUR consists of $|\Psi_{wgi}| + 1$ additions (consider a subtraction as a addition). If the EHE or FED criterion is applied, n additional comparisons are required. The EHE criterion needs other extra n multiplications.

Note that if the algorithms are implemented on neural hardware, the computation complexity is counted based on iterations (one iteration can have many single-neuron updates) instead of single-neuron update. This is because neural hardware is supposed to provide parallel computation and so no matter how many neurons are updated in one iteration, the computation time is the same as that for one signal-neuron update. As demonstrated in the simulation, the EHE (FED) algorithms can considerably reduce the number of iterations while converging to the images of the same quality. The reason is that the EHE and FED criteria select those neurons to update at each iteration, whose updates have better chance to be correct. Hence, when implemented on neural hardware, the EHE (FED) algorithms can also have superior performance over other algorithms in computation time. We note that although the other MHNN algorithms can also operate in partially simultaneous updating mode as shown by the GUR, the image quality becomes worse because of the increased number of neurons updated at each iteration.

Finally, the computation complexity of the EHE and FED algorithms can be simplified in practice. In image restoration, if we assume that the image can be extended periodically and \mathbf{A} and \mathbf{D} are convolutions over small windows, the thresholds t_i of all n neurons are identical. $|h_i|/t_i$ as well as $|h_i(k)| - t_i$ can be replaced by $|h_i|$ for specification of $L(k)$ in the EHE (FED) algorithms. Hence, the extra n multiplications in the EHE algorithms are eliminated. For the same reason, the computation of t_i is simplified at the initial.

Because any row (or column) of \mathbf{A} and \mathbf{D} , and so of \mathbf{W} and \mathbf{G} can be expressed by their first row (or column), when all these algorithms are implemented on a nonneural hardware or computer, the required memory for storage of these matrices is only $1/n$ of that defined by the size of these matrices.

D. Correct Transition Rate

As analyzed in Section III, the EHE criterion achieves the highest probability that a neural transition is correct in terms of decreasing distance $\|\mathbf{x}^o - \mathbf{x}(k)\|$ among all updateable neurons. In this paper and the companion [16], we call this probability the correct transition probability (CTP). We propose in [16] the CTP as a new performance measure for iterative algorithms for image restoration and reconstruction. The CTP is the probability that with given statistics at a particular iteration time k the neural state transition is correct. In computer simulation, the CTP is hard to estimate. Consider the probability averaged over all CTP's in all iterations in a long run for image recovery.

This averaged CTP has a different statistical meaning from any CTP at a particular iteration because the statistical properties of the statistics in the recovery processing vary iteration by iteration. However, the averaged CTP is suitable for estimation in computer simulation. In this paper, the estimate of the averaged CTP is called the correct transition rate (CTR) that is defined as follows.

Consider an algorithm that takes a total of k_e iterations to recover an image. In each iteration, a number of neurons are updated. Among the updated neurons, some neural states are changed. These state changes are counted in the total number of transitions K_t . If the state transition of a neuron results in the decrease of $|x_i^o - x_i(k)|$, the difference between the i th-neuron state and the i th pixel of the original image, the transition is said to be a correct transition and is counted in the total number of the correct transitions K_c . Then the CTR is defined as

$$\text{CTR} = \frac{K_c}{K_t}. \quad (29)$$

Obviously, the CTR approximates the probability that an algorithm correctly identifies wrong neuronal states based on the current network state in a long processing for the recovery of an image. The higher the CTR is, the better the performance is. If $\text{CTR} = 1$, the algorithm can recover an error-free image with the least transitions and iterations. On the other hand, the closer the CTR is to 0.5, the more iterations are necessary for an algorithm to recover an image because the algorithm decreases $\|\mathbf{x}^o - \mathbf{x}(k)\|$ in nearly half number of all transitions and increases $\|\mathbf{x}^o - \mathbf{x}(k)\|$ in the other half. If $\text{CTR} < 0.5$, the recovered image is worse than the initial image, i.e., $\|\mathbf{x}^o - \mathbf{x}(k)\| > \|\mathbf{x}^o - \mathbf{x}(0)\|$, because there are more transitions to increase than to decrease $\|\mathbf{x}^o - \mathbf{x}(k)\|$.

V. SIMULATION RESULTS

Performance of six HNN based algorithms is compared by simulation for gray image restoration. These algorithms are the SA algorithm, ZCVJ algorithm, PK algorithm, and the EHE based Algorithms 1, 2, and 3 (equivalent to Algorithms 4, 5, and 6, respectively, for image restoration). The SA algorithm, ZCVJ algorithm, PK algorithm and Algorithm 1 operate in the ordinary sequential updating mode, and Algorithms 2 and 3 operate in the wide-sense sequential updating modes. Hence, all six algorithms operate in the wide-sense sequential updating modes. In terms of Theorem 3, the images restored by the PK algorithm and Algorithms 1–3 are local minimum points of the energy function. Since the ZCVJ algorithm is equivalent to the PK algorithm (but with considerably more complex computation in check of energy change in the ZCVJ algorithm), the ZCVJ algorithm also converges to the local minimum point of energy function. The SA algorithm may converge to the global minimum in statistical sense. In the SA algorithm, ZCVJ algorithm, and PK algorithm, $L(k)$'s for $k \geq 0$ are pre-specified such that the neurons (image pixels) are visited column by column and row by row in order. On the other hand, in Algorithms 1, 2, and 3, $L(k)$'s for $k \geq 0$ are adaptively determined according to the EHE criterion based on the current network state at each step.

In all simulations, as shown in Fig. 5(a), the Lena image of size 128×128 is used as the original image \mathbf{x}^o . The blurring matrix \mathbf{A} represents a convolution of uniform blur over a window of size $(2c - 1) \times (2c - 1)$, i.e.,

$$f_A(i, j) = \frac{1}{(2c-1)^2} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}. \quad (30)$$

The larger the c is, the more severely the image is blurred. $c = 2$ is considered in all simulations. The difference matrix \mathbf{D} denotes a convolution

$$f_D(i, j) = \frac{1}{100} \begin{bmatrix} 0.7 & 1 & 0.7 \\ 1 & -6.8 & 1 \\ 0.7 & 1 & 0.7 \end{bmatrix}. \quad (31)$$

The signal-to-noise ratio (SNR) of the blurred noisy image is defined as $\text{SNR} = 10 \log_{10}(\sigma_S^2/\sigma^2)$ where σ_S^2 is the pixel variance estimated from the noise-free blurred image $\mathbf{A}\mathbf{x}^o$ and σ^2 is the pixel variance of the noise added to the blurred image. Given the SNR, we can know σ^2 . In addition to the CTR, the improved SNR is also considered in measure of the restored images, which is defined as

$$\text{dSNR} \equiv 10 \log_{10} \left(\frac{\|\mathbf{y} - \mathbf{x}^o\|^2}{\|\mathbf{x}^e - \mathbf{x}^o\|^2} \right)$$

where \mathbf{x}^e is the finally restored image. In all simulations, the image quantified from the blurred noisy image \mathbf{y} is used as initial image $\mathbf{x}(0)$. Hence, at the initial, $\text{dSNR} \cong 0$ dB.

In all simulations, the initial temperature in the SA algorithm is 64 and decreases by a factor 1.02 at every cycle of n single-neuron iterations. After the temperature becomes $T = 0$, the SA algorithm becomes the ZCVJ algorithm so that the energy descent is guaranteed. In Algorithm 2, the initial $N = 1024$, and $\beta = 1.01$ are set. In Algorithm 3, initial $\delta = 0.9$, and $\Delta\delta = 0.0002$ are set.

The dSNR and CTR as well as the visual quality of restored images in all simulations indicate that the EHE criterion can improve the performance of the MHNN algorithm. In return, this implies that the CTR is a good measure of the image restoration capability and a good predictor of finally-restored image quality.

Simulation A: In this simulation, we intend to compare the EHE based algorithms with other algorithms in various conditions. We also intend to observe how dSNR, CTR and residual energy of the restored image change with SNR and λ . Since the SA, PK, and ZCVJ algorithms are similar in the sense that they do not apply the EHE criterion (the SA algorithm is different in statistical sense), we consider the PK algorithm a representative of them in comparison. We do not consider the SA algorithm in this comparison because it is too much time-consuming. Since Algorithm 3 converges faster among the three EHE algorithms, we consider it a representative of the EHE based algorithms in the comparison.

The simulation results are shown in Figs. 2–4 for the dSNR, CTR, and residual energy, respectively. As shown in Fig. 2, the

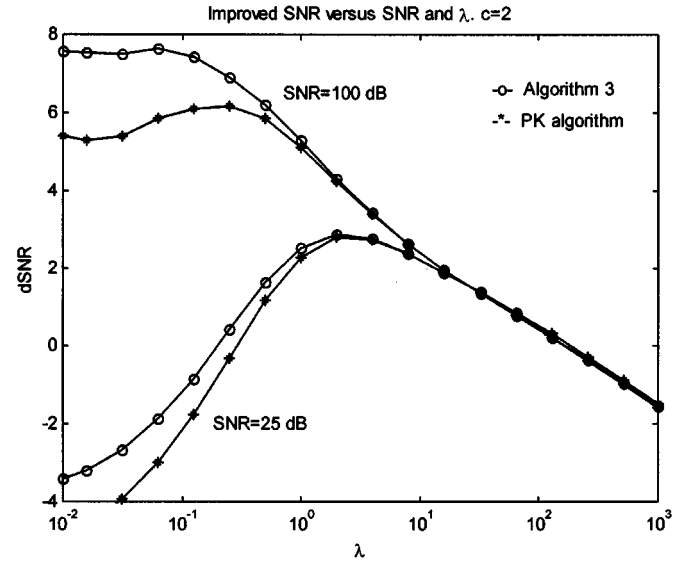


Fig. 2. Improved SNR of restored images versus SNR of blurred noisy image and regularization.

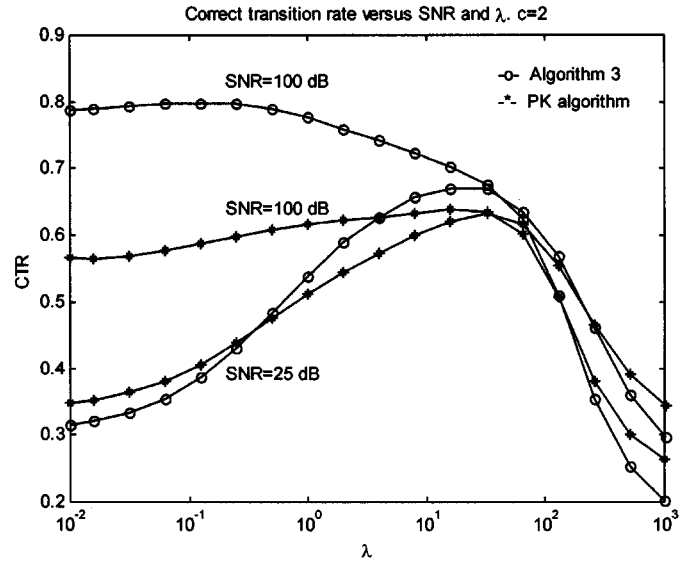


Fig. 3. The correct transition rate versus SNR of blurred noisy image and regularization.

improved SNR of the image restored by the EHE based algorithm is always higher than that by the PK algorithm for all SNR and λ . The largest improvement of dSNR of the EHE algorithm upon the PK algorithm is about 2 dB and occurs when the SNR is high and $\lambda = 0$. This improved performance decreases with λ . When the SNR is high, the EHE algorithm also has the best performance when $\lambda = 0$. This is because when the noise is weak, the regularization term does not serve to suppress noise rather than to interfere to the signal under detection, thus deteriorating the restoration of the original image. On the other hand, when the SNR is low, the regularization term helps suppress the influence of noise and benefits the restoration of the original image. If λ is too large, the energy function is dominated by the regularization term and forces any algorithm to lose restoration capability as shown by the energy function. Hence, when the noise is strong, optimal values of λ exist for both the EHE and

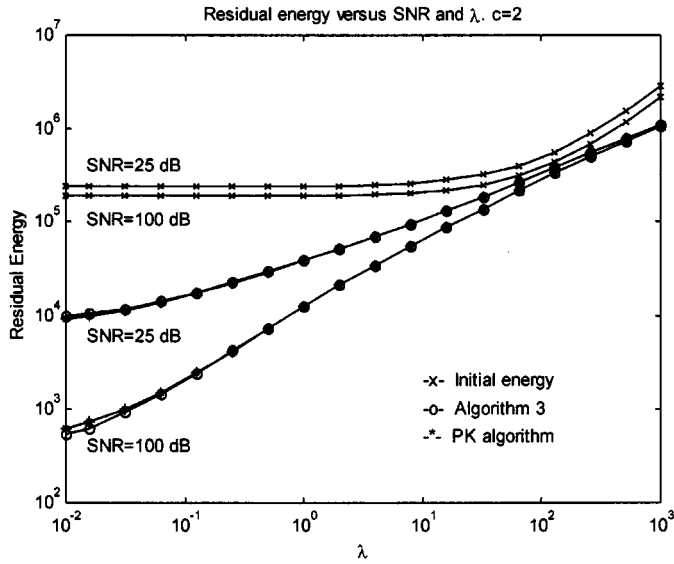


Fig. 4. The residual energy of the restored images versus SNR of blurred noisy image and regularization.

PK algorithms. The visual quality of restored image also suggests the existence of the optimal value of λ for low SNR. In the case of low SNR, if a small λ is used, the restored images (not shown in this paper) are rough and noisy. On the other hand, if a very large λ is used, the restored image is blurred though smooth. Hence, the visual quality of restored images requires a tradeoff between the smoothness and fineness, thus suggesting the existence of the optimal value of λ . It is also observed from Fig. 2 that the gap of the dSNR created by the difference of SNR decreases with λ .

As shown in Fig. 3, the CTR's of these two algorithms behave similarly to the dSNR's. When the SNR is high, the CTR of the EHE algorithm is much higher (the largest percent difference is about 20%) than that of the PK algorithm in a large range of λ and drops down to below the CTR of the PK algorithm when λ is too large. When the SNR is low, the CTR of the EHE algorithm is higher than that of the PK algorithm if λ is properly chosen, and is lower if λ is too small or too large. In all cases, the CTR of the EHE algorithm is higher than that of the PK algorithm if both CTR's are higher than 0.5, and is lower if both are lower than 0.5. The cross point is about 0.5 in all conditions. Compared with the dSNR in Fig. 2, we can see that the cross point of CTR = 0.5 corresponds to the crucial point of dSNR = 0 dB. If the CTR of either algorithm is below the crucial point 0.5, the dSNR also drops below the critical point 0 dB in all cases. The visual quality of restored images (not given in this case in the paper) also becomes worse than that of the blurred noisy image y when the CTR and dSNR become below their crucial points. This confirmation of the CTR, dSNR and image visual quality implies that if its CTR is lower than the crucial point 0.5, an algorithm loses its capability of image restoration. In other words, CTR is a rational performance measure about the restoration capability of an iterative algorithm. In addition, it is a good predictor for the quality of the finally restored image. All of these also show that the EHE criterion can improve the performance of an MHNN algorithm and the quality of restored image.

TABLE I
PERFORMANCE COMPARISON OF SIX HOPFIELD NEURAL NETWORK BASED ALGORITHMS

Algorithm	dSNR (dB)	Iter., k_e	K_t	CTR(%)	Figure
SA	7.19	11,026,432	6,605,501	50.50	Fig. 6 (a)
ZCVJ	5.38	1,622,016	337,415	56.88	Fig. 6 (b)
PK	5.41	337,897	337,897	56.89	Fig. 6 (c)
Alg. 1	6.84	140,568	140,568	75.80	Fig. 6 (d)
Alg. 2	6.91	97,398	140,710	75.80	Fig. 6 (e)
Alg. 3	6.85	33,938	141,342	75.72	Fig. 6 (f)

Fig. 4 shows the residual energy of the restored images versus the SNR and λ . To show meaningful values of energy, in the figure we consider the energy defined by

$$E(x) + \|y\|^2/2 = [\|Ax - y\|^2 + \lambda\|Dx\|^2]/2$$

where $x = x(0)$ for the initial energy and $x = x(k_e) = x^e$ for the residual energy with $x(k_e) = x^e$ being the restored image. As shown in Fig. 4, both algorithms are capable of reducing energy in all conditions. The larger is the SNR and the smaller is λ , the larger is the amount of reduced energy. In the simulated cases, the largest factor of energy reduction is about 523 when SNR = 100 dB and $\lambda = 0.1$. When λ is large, both algorithm are incapable to reduce energy effectively. This means that when λ is large, the image quantified from the blurred noise image (as done in the simulation) is a good or better solution. The difference of residual energy between the images obtained by these two algorithms exist when λ is small though this difference looks small in the scale of the initial energy. When SNR = 100 dB and λ is small, the EHE algorithm achieves the lower residual energy than the PK algorithm. However, when the noise is strong and λ is small, the residual energy of the EHE algorithm becomes a little larger than that of the PK algorithm. This confirms Fig. 3 in that the EHE criterion becomes less effective when the noise is strong. Nevertheless, as shown by the dSNR in Fig. 2 and the visual quality of restored images (e.g., partly shown in Fig. 6), in all situations the EHE algorithm obtains better images than the PK algorithm does.

Simulation B: In this simulation, all six MHNN algorithms are compared in the conditions of SNR = 40 dB and $\lambda = 0.1$. The difference between the initial image $x(0)$ and the original lena image is $\|x(0) - x^o\|_1 = 149\,236$, which requires 149 236 correct, and no wrong single-neuron transitions to restore the perfect lena image. Simulation results are given in Table I. The blurred image is shown in Fig. 5(b). The restored images are shown in Fig. 6(a)–(f). The SA algorithm uses a very large number of iterations to obtain a good image with the highest dSNR. This confirms its CTR close to 0.5. The ZCVJ algorithm uses a large number of iterations to obtain a worse image of the lowest dSNR. Its CTR is a little higher. The PK algorithm has the same performance as that of the ZCVJ algorithm in dSNR, CTR, and image visual quality except that it takes a much smaller number of iterations to achieves this same performance. In terms of their low correct transition rates, a large portion of neural state transitions taken by the SA, ZCVJ and PK algorithms are impractical and waste computation time.

All three EHE based algorithms demonstrate better performance than the other three algorithms. In very small total numbers of neural transitions K_t (close to $\|x(0) - x^o\|_1 = 149\,236$),

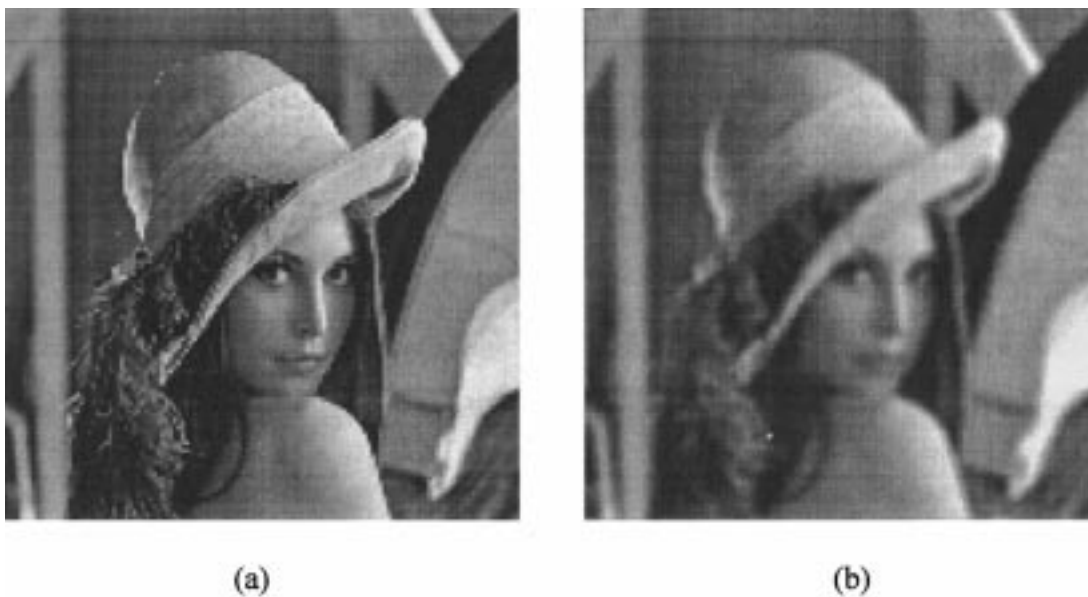


Fig. 5. (a) Original Lena image. (b) Blurred image with $c = 2$ and 40 dB additive white Gaussian noise.

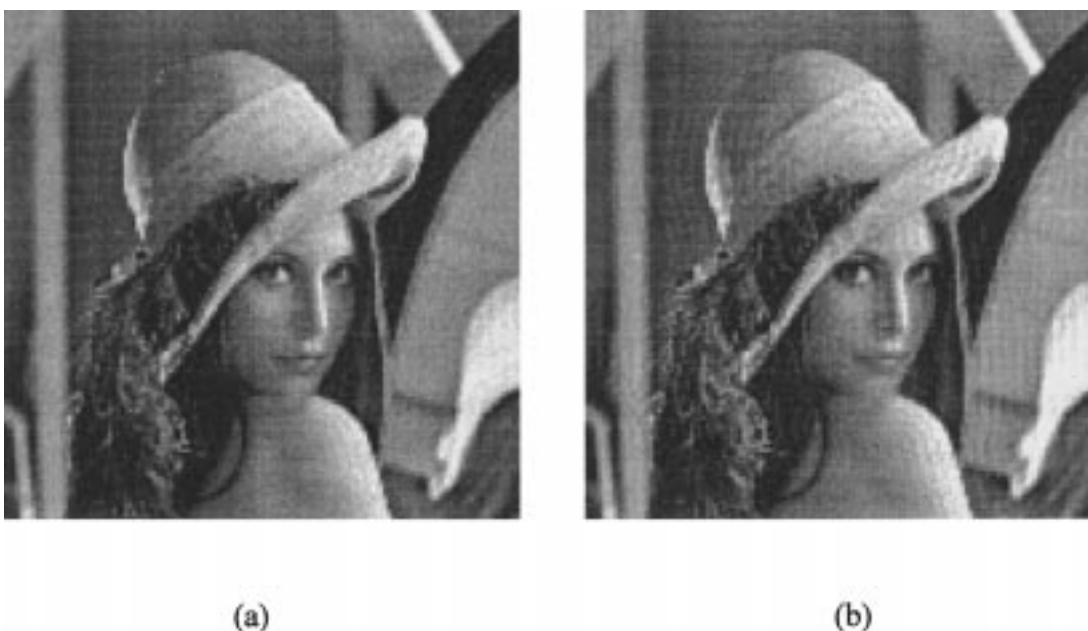


Fig. 6. Restored images. (a) By SA algorithm and (b) by ZCVJ algorithm.

they obtain more accurate images of higher dSNR's. The visual quality of their restored images is also highest. The numbers of iterations k_e taken by the EHE algorithms are significantly smaller. These confirm the analysis in previous sections that the EHE criterion (equivalent to the FED criterion for image restoration) achieves the largest amount of energy descent at each neural transition. The CTR's of these three EHE based algorithms are about 20 percent higher than those of the other three algorithms. Because of the aid of the EHE criterion, with higher probabilities the EHE based algorithms can correctly identify the wrong neural states based on the current network state.

The simulation also shows that the CTR is a good performance measure. As shown in Table I, the larger is the CTR, the

larger is the dSNR and the smaller is the total number of iterations. Furthermore, a high CTR corresponds to a high visual quality of restored image as shown in Fig. 6.

As shown in Fig. 6, the images restored by the ZCVJ and PK algorithms severely suffer from the streaks, which is a problem existing in many other conventional algorithms. The image restored by the SA algorithm has obviously-suppressed, slight streaks, which is due to its robustness in statistical sense. The streaks existing in the images restored by the SA, ZCVJ and PK algorithms can be explained as follows. Since the pixels are updated row by row and column by column in all these three algorithms, the error at one pixel may be propagated to and accumulated on the near pixels updated in following iterations because pixels are correlated locally. When the accumulated error



(c)



(d)



(e)



(f)

Fig. 6. (Continued.) Restored images. (c) By PK algorithm, (d) by Algorithm 1, (e) by Algorithm 2, and (f) by Algorithm 3.

is too large, it is suddenly reduced so as to satisfy the limitation of error. Then the error begins to accumulate from a small value to a large value again. This procedure repeats and yields the wave-like streaks on the image. In contrast, the EHE criterion based algorithms can effectively avoid the production of the streaks. If the number of neurons whose absolute inputs are largest (and so are chosen by the EHE criterion) in one or more consecutive iterations is small, these chosen neurons are usually spatially isolated. Hence, the updates of these neurons do not accumulate and propagate errors, thus avoiding the occurrence of streaks in a finally restored image.

Comparing Algorithms 1, 2, and 3, we observe that Algorithms 2 and 3 take much smaller numbers of iterations to obtain the images of almost the same quality of that obtained by the Algorithm 1. Algorithms 2 and 3 reduce the total number of

iterations because before operating in the sequential updating mode, they update a large number of neurons at one iteration. The neurons that have the largest absolute inputs and are chosen by the EHE criterion are spatially isolated with a large probability. Hence, the updates of such neurons at the same iteration do not increase the chance to make a wrong update while significantly reducing the number of iterations. Note that if the algorithms are implemented on neural hardware, the computation time accounts for only the number of iterations and is independent of the number of updated neurons at each iteration because the neural hardware is supposed to provide parallel computation. Hence, the EHE based Algorithms 2 and 3 can considerably reduce computation time of neural hardware.

We also carried out the experiments on binary image reconstruction from a few projections with and without limited view

angles. In this case, the EHE and FED criteria are not equivalent. The results also demonstrate the superiority of the EHE and FED based algorithms over others.

VI. CONCLUSIONS

In this paper, a generalized updating rule suitable for gray image restoration and reconstruction is presented. This GUR guarantees the energy decrease when operating in *any* sequence of updating modes. For the first time, it is shown that the neural threshold established in the GUR is necessary and sufficient with probability one to reduce energy.

The EHE criterion and three EHE based algorithms (two of them are new) for image restoration and reconstruction are presented. The novel FED criterion and three FED based algorithms are newly proposed. The EHE criterion achieves the highest correct transition probability and the FED criterion achieves the largest amount of energy descent at each iteration. For image restoration, the EHE and FED criteria are equivalent.

The novel correct transition rate is proposed for performance measure of iterative algorithms and for prediction of quality of finally restored image. The closer to one is the correct transition rate, the fewer are the iterations and the more accurate is the restored image. If the correct transition rate is lower than 0.5, the algorithm has no capability of recovering an image.

Simulation results show that the EHE (FED) algorithms converge to images that are more accurate in much fewer iterations than other HNN algorithms. The restored images by the EHE (FED) algorithms have much better visual quality and much higher improved SNR. The EHE (FED) based algorithms can effectively prevent errors from propagation and accumulation, thus avoiding appearance of streaks in finally restored image. When the noise is weak, the optimal value of λ for the EHE (FED) algorithms is zero. When the noise is strong, optimal nonzero values of λ exist for the EHE (FED) algorithms as well as the non-EHE algorithms. Simulation results also suggest that the correct transition rate be a rational performance measure of iterative algorithms and a good predictor of the finally-restored image quality.

The EHE and FED criteria increase a little computation complexity with additional comparisons. If implemented on neural hardware, the EHE and FED criteria can considerably reduce computation time.

This whole work can be applied to other applications that have the same problem formulation as the digital image restoration and reconstruction, e.g., CDMA multiuser detection.

APPENDIX

Proof of (17) and (18): For any positive numbers x_1, \dots, x_n , we have [17, pp. 61]

$$\frac{1}{n} \sum_{i=1}^n x_i \leq \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \quad (\text{A1})$$

where the equality holds if and only if $x_1 = x_2 = \dots = x_n$. When rows of $\mathbf{W} + \lambda \mathbf{G}$ differ only by permutations, t_i^{sim} is independent of i . By applying (A1)

$$\begin{aligned} t_i^{sim} &= \frac{1}{2} \sum_{j=1}^n |w_{ij} + \lambda g_{ij}| \\ &= \frac{1}{2n} \sum_{i=1}^n \sum_{j=1}^n |w_{ij} + \lambda g_{ij}| \\ &\leq \frac{n}{2} \sqrt{\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |w_{ij} + \lambda g_{ij}|^2} \\ &= \frac{1}{2} \sqrt{\sum_{i=1}^n \sum_{j=1}^n |w_{ij} + \lambda g_{ij}|^2} \\ &= t_i^{PK} \end{aligned} \quad (\text{A2})$$

where the equality holds if and only if $|w_{ij} + \lambda g_{ij}| = b$. For any real matrix $\mathbf{W} + \lambda \mathbf{G}$

$$\begin{aligned} U^{sim} &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n |w_{ij} + \lambda g_{ij}| \\ &\leq \frac{n^2}{2} \sqrt{\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |w_{ij} + \lambda g_{ij}|^2} \\ &= \frac{n}{2} \sqrt{\sum_{i=1}^n \sum_{j=1}^n |w_{ij} + \lambda g_{ij}|^2} \\ &= U^{PK} \end{aligned} \quad (\text{A3})$$

where the equality is true if and only if $|w_{ij} + \lambda g_{ij}| = b$. (Q.E.D.)

Proof of Theorem 2: The sufficiency is guaranteed by Theorem 1. In what follows, we prove the necessity.

Given $L(k)$ and $\mathbf{x}(k)$ at time k where $\mathbf{x}(k)$ is not bounded. We consider an interconnection matrix $\mathbf{W} + \lambda \mathbf{G}$ such that $w_{ij} + \lambda g_{ij} \leq 0, \forall i, j \in L(k)$. Note that such an interconnection matrix is common in image restoration and reconstruction. In image restoration, the elements of \mathbf{A} are usually nonnegative, and in image reconstruction the elements of \mathbf{A} are always nonnegative. If λ is small, the elements of $\mathbf{W} + \lambda \mathbf{G}$ are nonpositive.

Consider $\forall \delta > 0$ such that $t_i(k) - \delta \geq 0$. In what follows, we show that if the neural threshold $t_i(k)$ of (8) in the GUR is replaced by a smaller threshold $t_i(k) - \delta$, the probability that at time k the energy change is smaller than zero is smaller than one, i.e., $\Pr(\Delta E(k) < 0) < 1$.

We define a set Ω of \mathbb{R}^n by $\Omega = \{\mathbf{z} \in \mathbb{R}^n | t_i(k) - \delta < z_i \leq t_i(k), i \in L(k)\}$. $\mathbf{h}(k)$ in (5) can be rewritten as

$$\mathbf{h}(k) = (\mathbf{W} + \lambda \mathbf{G})\mathbf{x}(k) - \mathbf{A}^T \mathbf{x}^o + \mathbf{A}^T \mathbf{n}. \quad (\text{A4})$$

Since \mathbf{n} is white Gaussian noise with zero mean and covariance matrix $\sigma^2 \mathbf{I}$, if $\sigma \neq 0$, the probability that $\mathbf{h}(k)$ is located in Ω is greater than zero, i.e.

$$\Pr(\mathbf{h}(k) \in \Omega) > 0. \quad (\text{A5})$$

Consider $\mathbf{h}(k) \in \Omega$. Suppose that the threshold $t_i(k)$ of (8) in the GUR is replaced by the new threshold $t_i(k) - \delta$ (which is

strictly smaller than $t_i(k)$), in terms of the GUR, all $x_i(k)$ for $i \in L(k)$ must be updated such that $\Delta x_i(k) = x_i(k+1) - x_i(k) = 1$ because $h_i(k) > t_i(k) - \delta$, $\forall i \in L(k)$. Such a $\Delta \mathbf{x}(k)$ is a nonzero update of the GUR.

Denote the i th coordinate vector and the i th column vectors of \mathbf{W} and \mathbf{G} by \mathbf{e}_i , \mathbf{w}_i , and \mathbf{g}_i , respectively. Notice that $\Delta x_i(k) = 0$ for any $i \notin L(k)$. In this case, the energy change according to (6) is

$$\begin{aligned}
 \Delta E(k) &= - \sum_{i \in L(k)} \Delta x_i(k) \mathbf{e}_i^T \\
 &\quad \cdot \left(\mathbf{h}(k) + \frac{1}{2} (\mathbf{W} + \lambda \mathbf{G}) \sum_{j \in L(k)} \Delta x_j(k) \mathbf{e}_j \right) \\
 &= - \sum_{i \in L(k)} \mathbf{e}_i^T \left(\mathbf{h}(k) + \frac{1}{2} \sum_{j \in L(k)} (\mathbf{w}_j + \lambda \mathbf{g}_j) \right) \\
 &= - \sum_{i \in L(k)} \left(h_i(k) + \frac{1}{2} \sum_{j \in L(k)} (w_{ij} + \lambda g_{ij}) \right) \\
 &= \sum_{i \in L(k)} \left(-h_i(k) - \frac{1}{2} \sum_{j \in L(k)} (w_{ij} + \lambda g_{ij}) \right) \\
 &= \sum_{i \in L(k)} \left(-h_i(k) + \frac{1}{2} \sum_{j \in L(k)} |w_{ij} + \lambda g_{ij}| \right) \\
 &= \sum_{i \in L(k)} (-h_i(k) + t_i(k)) \\
 &= \sum_{i \in L(k)} (-h_i(k) + t_i(k)) \geq 0 \tag{A6}
 \end{aligned}$$

where the last inequality holds due to $\mathbf{h}(k) \in \Omega$. This proves that $\Pr(\Delta E(k) \geq 0 | \mathbf{h}(k) \in \Omega) = 1$. Hence

$$\begin{aligned}
 \Pr(\Delta E(k) \geq 0) &= \Pr(\Delta E(k) \geq 0 | \mathbf{h}(k) \in \Omega) \Pr(\mathbf{h}(k) \in \Omega) \\
 &\quad + \Pr(\Delta E(k) \geq 0 | \mathbf{h}(k) \notin \Omega) \Pr(\mathbf{h}(k) \notin \Omega) \\
 &\geq \Pr(\Delta E(k) \geq 0 | \mathbf{h}(k) \in \Omega) \Pr(\mathbf{h}(k) \in \Omega) \\
 &= \Pr(\mathbf{h}(k) \in \Omega) > 0 \tag{A7}
 \end{aligned}$$

which implies

$$\Pr(\Delta E(k) < 0) = 1 - \Pr(\Delta E(k) \geq 0) < 1. \tag{A8}$$

(Q.E.D.)

ACKNOWLEDGMENT

The author thanks Prof. S.-Y. Yu and Prof. J.-G. Li of Shanghai Jiao Tong University for their helpful discussions at the initial stage of this work. The author also thanks the anonymous reviewers for their careful review and constructive comments.

REFERENCES

- [1] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Nat. Acad. Sci.*, vol. 79, pp. 2554–2558, Apr. 1982.
- [2] Y.-T. Zhou, R. Chellappa, A. Vaid, and B. K. Jenkins, "Image restoration using a neural network," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 7, pp. 1141–1151, July 1988.
- [3] J. K. Paik and A. K. Katsaggelos, "Image restoration using a modified Hopfield neural network," *IEEE Trans. Image Processing*, vol. 1, pp. 49–63, Jan. 1992.
- [4] Y. Sun and S.-Y. Yu, "A modified Hopfield neural network used in bilevel image restoration and reconstruction," in *Proc. Int. Symp. Information Theory Application*, vol. 3, Singapore, Nov. 16–20, 1992, pp. 1412–1414.
- [5] —, "An eliminating highest error criterion in Hopfield neural network for bilevel image restoration," in *Proc. Int. Symp. Inform. Theory Applicat.*, vol. 3, Singapore, Nov. 16–20, 1992, pp. 1409–1411.
- [6] H.-J. Liu and Y. Sun, "Blind bilevel image restoration using Hopfield neural networks," in *Proc. IEEE Int. Conf. Neural Networks*, San Francisco, CA, Mar. 28–Apr. 1, 1993, pp. 1656–1661.
- [7] Y. Sun, J.-G. Li, and S.-Y. Yu, "Improvement on performance of modified Hopfield neural network for image restoration," *IEEE Trans. Image Processing*, vol. 4, pp. 688–692, May 1995.
- [8] Y. Sun, "A generalized updating rule for modified Hopfield neural network for quadratic optimization," *Neurocomput.*, vol. 19, pp. 133–143, 1998.
- [9] G. Demoment, "Image reconstruction and restoration: Overview of common estimation structures and problems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 2024–2036, Dec. 1989.
- [10] J. Bruck and J. W. Goodman, "A generalized convergence theorem for neural networks," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1089–1092, Sept. 1988.
- [11] J. Bruck, "On the convergence properties of the Hopfield model," *Proc. IEEE*, vol. 78, pp. 1579–1585, Oct. 1990.
- [12] E. Goles-Chacc, F. Fogelman-Soulie, and D. Pellegrin, "Decreasing energy functions as a tool for studying threshold networks," *Disc. Appl. Math.*, vol. 12, pp. 261–277, 1985.
- [13] Y. Sun, "Search algorithms based on eliminating-highest-error and fastest-metric-descent criteria for bit-synchronous CDMA multiuser detection," in *Proc. IEEE Int. Conf. on Comm., ICC'98*, Atlanta, GA, June 7–11, 1998, pp. 390–394.
- [14] —, "Eliminating-highest-error and fastest-metric-descent criteria and iterative algorithms for bit-synchronous CDMA multiuser detection," in *Proc. IEEE Int. Conf. Comm.*, Atlanta, GA, June 7–11, 1998, pp. 1576–1580.
- [15] M. A. T. Figueiredo and J. M. N. Leitao, "Sequential and parallel image restoration: Neural network implementations," *IEEE Trans. Image Processing*, vol. 3, pp. 789–801, Nov. 1994.
- [16] Y. Sun, "Hopfield neural network based algorithms for image restoration and reconstruction—Part II: Performance analysis," *IEEE Trans. Signal Processing*, vol. 48, pp. 2119–2131, July 2000.
- [17] R. G. Bartle, *The Elements of Real Analysis*, 2nd ed. New York: Wiley, 1976.



Yi Sun (M'92) received the B.S. and M.S. degrees in electrical engineering from the Shanghai Jiao Tong University, Shanghai, China, in 1982 and 1985, respectively, and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, in 1997.

From 1985 to 1993, he was a Lecturer with the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University. In the Summer of 1993, he was a Visiting Scientist with the Department of Mechanical Engineering, Concordia University, Montreal, P.Q., Canada. From March to September 1997, he was a Postdoctoral Research Fellow with the Radiology Department, University of Utah, Salt Lake City, where he studied MRI imaging. From October 1997 to August 1998, he was a Postdoctoral Research Associate working on wireless communications in the Department of Electrical and Systems Engineering, University of Connecticut, Storrs. Since September 1998, he has been an Assistant Professor with the Department of Electrical Engineering, City College, City University of New York. His research interests are in the areas of wireless communications (with focus on CDMA multiuser detection, slotted CDMA networks, channel equalization and sequence detection, and multicarrier systems), image and signal processing, medical imaging, and neural networks.