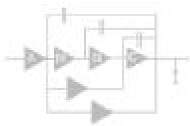
Spectral Associative Memories



Ron Spencer, Ph.D.

Texas A&M University
Analog & Mixed Signal Center,
Electrical Engineering Dept.

Preview

- Objectives/Scope
- Content addressable memory
- Hopfield formalism
- Probabilistic formulation of Hopfield nets w/quantum mechanical analogies
- Spectral associative memory (SAM)
- Bit error rates of Bipolar-SAM (SNR, decoding time, & oversampling)
- CMOS Implementation of Bipolar-SAM

Objectives/Scope

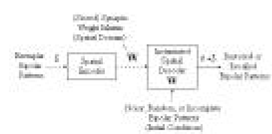
What are the objectives of this work?

- To understand the similarities and differences between associative memory and quantum computation.
- To explore quantum implementations that exploit quantum parallelism (*spurious-free recall*?).
- To compare spectral associative memory with other telecommunication architectures (e.g. TURBO & LDPC decoders) in terms of *coding gain*, *noise immunity*, *feasibility* and *ease of implementation*.
- Cross-fertilization of telecommunications <=> neural networks <=> quantum mechanics.

Two types of addressibility:

- "Address-addressibility" - retrieval by **address**
  e.g.: RAM, HDD, CDROM
- "Content-addressibility" - retrieval by **content**

  - requires *long-range connectivity* & "*mass-action*"

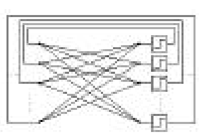


W is analogous to an observable in quantum mechanics

v is analogous to an observed state

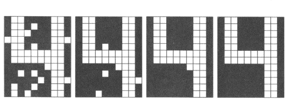
- Long-range, off-diagonal (non-local) connectivity leads to "mass-action".

$$\mathbf{W} = \begin{bmatrix} 0 & w_{12} & w_{13} & w_{14} & w_{15} \\ w_{21} & 0 & w_{23} & w_{24} & w_{25} \\ w_{31} & w_{32} & 0 & w_{34} & w_{35} \\ w_{41} & w_{42} & w_{43} & 0 & w_{45} \\ w_{51} & w_{52} & w_{53} & w_{54} & 0 \end{bmatrix}$$



- Extrinsic redundancy leads to error correction and noise immunity, but for a price! Spatial dimension.

The network is initialized with a noisy, unknown pattern which converges to one of the store attractors that is "nearest" in terms of Hamming distance.



Pattern coding: For a given bipolar pattern, **s**, we wish to store in the associative memory network, a network attractor may be constructed in the form of a "*weight matrix*":

$$\mathbf{W} = \frac{1}{p}\sum_{m=1}^{p} \mathbf{s}^{(m)}\mathbf{s}^{(m)T} - \mathbf{I} \qquad \mathbf{W} = \begin{bmatrix} 0 & w_{12} & w_{13} & w_{14} & w_{15} \\ w_{21} & 0 & w_{23} & w_{24} & w_{25} \\ w_{31} & w_{32} & 0 & w_{34} & w_{35} \\ w_{41} & w_{42} & w_{43} & 0 & w_{45} \\ w_{51} & w_{52} & w_{53} & w_{54} & 0 \end{bmatrix}$$

where $\mathbf{ss}^T$ is the autoassociative *outer product* of pattern vector **s**.

Example: Let $\mathbf{s} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ then $\mathbf{ss}^T = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$

and $\mathbf{W} = \mathbf{ss}^T - \mathbf{I} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$

Observation: Encoding one pattern automatically encodes the complement by default.

Pattern recall, or *error correction* from extrinsic redundancy:

1) asynchronous matrix-vector multiplications

2) thresholding

Example: Consider the following initial pattern: v ? ?1? ?1? ?? ?1?

Iteration 1: Wv|₀ ? ?0 1 ?1??1?? ?1 0 ?1??1 ? ?? ?1 ?1 0 ??1 ? ? v ? ?1? ?? ?1? (no change)

Iteration 2: Wv|₁ ? ?0 1 ?1??1 ? ?1 0 ?1??1? ?? ?1 ?1 0 ??1 ? ? v ? ?1? ?? ?1? (no change)

Iteration 3: Wv|₂ ? ?0 1 ?1??1 ? ?1 0 ?1??1 ? ?? ?1 ?1 0 ??1? ? v ? ?1 ? ?1 ? ?? 1?

Iteration 4: Wv|₃ ? ?0 1 ?1??1? ?1 0 ?1??1 ? ?? ?1 ?1 0 ??? 1? ? v ? ?1 ? ?1 ? ?? 1? (no change, steady -state, same as encoded pattern)

Steady state has been reached after 3 iterations.

Hopfield Storage & Recall Example

Two vectors to be encoded:

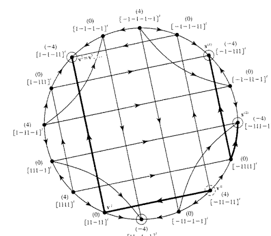
$\mathbf{s}^{(1)}$? ??1 ?1 1 1?

$\mathbf{s}^{(2)}$? ??1 1 1 ?1?

Superposed attractors:

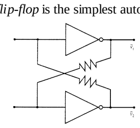
W ? ?0 0 ?2 0 ? ?0 0 0 ?2? ?? 2 0 0 0 ? ?0 ?2 0 0 ?

Flattened trajectory map & energies :



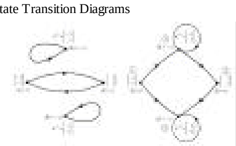
Flip-Flop (*2-D Hopfield*)

The *flip-flop* is the simplest autoassociative memory (n=2).



Bistable energy landscape:

State Transition Diagrams



Synchronous Asynchronous

(Images from Zurada's book)

Noise Removal vs. Content-Addressibility

Single-pattern associative memory serves only to remove noise. When only one pattern is stored, content-addressibility is not possible and the recalled pattern *does not depend on the initial conditions.*

Content-addressability is only possible when more than one pattern is stored in the weight matrix. When multiple patterns are stored, the recalled pattern *depends on the initial conditions.*

Multiple patterns: may be stored in the same weight matrix by summation:

W ? ? $\mathbf{W}_i$ (? i)

Probabilistic Formulation

We can define several spaces for the sake of developing a probabilistic formulation of associative memory:

1 Pattern space: e.g. a 2-D bipolar space:

Direct: ?1? ?1? | ? 1 ? ??1? | ?? 1? ? 1 ? | ?? 1? ??1?

Orthonormal: ?1? ?0? ?0? ?0? | ?0? ?1? ?0? ?0? | ?0? ?0? ?1? ?0? | ?0? ?0? ?0? ?1?

Dirac: $|1,1\rangle$ $|1,?1\rangle$ $|?1,1\rangle$ $|?1,?1\rangle$

2 Biased decoder possibility state space: e.g. for encoded pattern s = [1 1]ᵀ:

$|?_d\rangle$? $\frac{1}{2}|1,1\rangle$? $\frac{1}{2}|?1,?1\rangle$

Observation: One might say that the decoder is *biased* in a superposition of the encoded pattern and its complement.

In general, the probability p_j of a single pattern $|?_j\rangle$ being recalled

from decoder state $|?_d\rangle$? ? $a_i|?_i\rangle$ is p_j ? $\frac{a_j}{? a_i}$.

Furthermore, $r|?_d\rangle$? $|?_d\rangle$ since p_j ? $\frac{ra_j}{? ra_i}$? $\frac{a_j}{? a_i}$.

3 Recall probability density matrix :

D ? ? $a_i|?_i\rangle\langle ?_i|$

e.g. consider the single encoded pattern ? = [1 1]ᵀ. Then,

$\mathbf{D}_i$? $\frac{1}{2}|1,1\rangle\langle 1,1|$? $\frac{1}{2}|?1,?1\rangle\langle ?1,?1|$? $\frac{1}{2}$?1? ?0? ?0? ?0? ?1 0 0 0? ? $\frac{1}{2}$?0? ?0? ?0? ?1? ?0 0 0 1? ? ?1/2 0 0 0? ?0 0 0 0? ?0 0 0 0? ?0 0 0 1/2?

If we ask “what is the probability of recalling memory pattern ? $_c$ = [-1 -1]^T?”, we may obtain the answer by projecting the density matrix onto the *projection operator*, E$_{?c}$:

$$\mathbf{E}_{?_c} ? \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} ? \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

as follows:

$$p_{?_c} ? \text{Trace}\{DE_{?_c}\} ? \text{Trace}\left\{ \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right\} ? \frac{1}{2}$$

Observation: *Cross-pattern interference* occurs when the pattern space is of low dimensionality. This leads to “*spurious recall*”. The network still projects to one of the basis states, but it might not be one of the encoded states. This occurs for two or more memories when the bias space is a *sub-space* of the pattern space.

Encoding all possible patterns *cancels all the bias*:

$$\mathbf{D} ? \frac{1}{2}\mathbf{D}_? ? \frac{1}{2}\mathbf{D}_? ? \begin{bmatrix} .25 & 0 & 0 & 0 \\ 0 & .25 & 0 & 0 \\ 0 & 0 & .25 & 0 \\ 0 & 0 & 0 & .25 \end{bmatrix} \quad \text{where} \quad ? = [1\ 1]^T,\ ? = [1\ \text{-}1]^T$$

(**Does not** lead to spurious recall since the bias space equals pattern space.)

Observation: Interference only occurs between patterns in conventional AM since the *weighting factors are real.* Like patterns can only reinforce because phase shifts are not allowed.

4 Attractor space:

- Formed from the pattern space by the outer product
- Contains redundant copies of the pattern space
- Scales geometrically with pattern dimension rather than exponentially like the orthonormal pattern space ; i.e. n^2 rather than 2^n

Associative “entanglement” between channels?

The decoder state that results from the encoding of a single pattern cannot be factored into a product of states:

$$|?_d\rangle ? \frac{1}{2}|1,1\rangle ? \frac{1}{2}|?1,?1\rangle$$

but the *a priori* formulation of the state *can* be factored:

$$\mathbf{M} ? \mathbf{ss}^T ? \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} ? \begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix}$$

Why? Because distant cells in the memory are connected by wires.

How can we achieve the following using associative memory?

- Communicate information over a noisy channel?
- Have complex weighting factors in the probabilistic treatment?
- Alleviate quadratic and polynomial physical scaling complexity in a 2-D plane; e.g. a silicon chip?

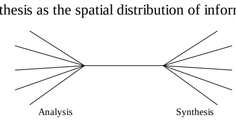
"....In both neural and silicon technologies, the active devices (synapses and transistors) occupy no more than one or two percent of the space - - "wire" fills the entire remaining space. The limitation of connectivity in nervous tissue has clearly forced the computational structure into a particular form...."

- Carver Mead, CalTech professor and founder of Neuromorphic Engineering, and student & cohort of Feynman

See also: C. Mead, *Collective Electrodynamics: Quantum Foundations of Electromagnetism*, MIT Press, Sept. 2000.

Beginnings of the Spectral Formulation

- Mass-action and long-range connectivity are present in nearly all neural network models; a focusing of spatially distributed information to single points in the network. Same for EM.
- Analysis & Synthesis:
  - Define analysis as the spatial concentration of information *to* a point.
  - Define synthesis as the spatial distribution of information *from* a point.



Analysis	Synthesis
“In-star”	“Out-star”
“Fan-in”	“Fan-out”
Focus	Diffraction
Inner-product	Outer or tensor products
$\langle a\|b\rangle$ or $\langle a\|\hat{p}\|b\rangle$	$\|a\rangle\langle b\|$ or $p_{ab}\|a\rangle\langle b\|$

We can achieve long-range, off-diagonal connectivity by *spectral convolution*:

Analysis in the frequency domain: $\mathbf{vW}^T$ $\qquad f(t)f(t)?\frac{1}{2}F(?)?F(?)$

$$v?t?? \sum_{?i} v_i e^{j?_i t} \qquad \text{Spectral decoder state}$$

$$W?t?? \sum_{l?1}^{n} \sum_{k?1}^{n} w_{kl} e^{j?_{wkl} t} \qquad \text{Attractor wave}$$

$$v?t?W?t? ? \sum_{i?1}^{n} v_i e^{j?_i t} \sum_{l?1}^{n} \sum_{k?1}^{n} w_{kl} e^{j?_{wkl} t} ? \sum_{i?1}^{n} \sum_{l?1}^{n} \sum_{k?1}^{n} v_i w_{kl} e^{jt?_i ? ?_{wkl}?}$$

we want $a_j ? \sum_{?i} v_i w_{ij}$ which leads to $?_{vi} ? ?_{wij} ? ??$

=> leads to the notion of a “*beat frequency*”.

Furthermore, long-range connectivity is made *virtually* in the frequency domain.

Synthesis in the frequency domain:

- Memory synthesis: $\mathbf{W} ? \mathbf{M} ? \mathbf{I}; \quad \mathbf{M} ? \mathbf{s}\mathbf{s}^T$

$$s_1 ?t ? ? \sum_{? i} s_i e^{j? _{s1i} t}$$

$$s_2 ?t ? ? \sum_{? i} s_i e^{j? _{s2i} t}$$

$$s_1 ?t ?s_2 ?t ? ? \sum_{i?1}^{n} s_i e^{j? _{s1i} t} \sum_{j?1}^{n} s_j e^{? j? _{s2j} t} ? \sum_{i?1}^{n} \sum_{j?1}^{n} s_i s_j e^{jt ?? _{s1i} ?? _{s2j} ?}$$

Letting band 1 be in a "*spectral column form*" and band 2 in "*spectral row form*", a "*spectral tensor product*" is formed.

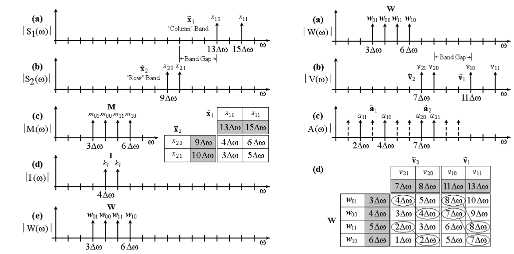
- Decoder state synthesis:

$$v ?t ? ? \sum_{? i} v_i e^{j? _{vi} t}$$

Band Structures

• Channels are separated by integer multiples of a **beat frequency**, ?? , and have **row-column form**.

Spectral Coding (AAM) Spectral Decoding (Recurrent AAD)

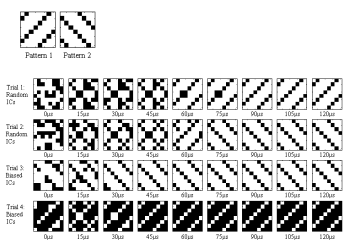


Multi-Pattern Recall

- Encoding and recall of multiple, simultaneous 8x8 binary patterns
- Interference and spurious attraction are still a problem for small networks

Encoded patterns =>

Recalled patterns =>

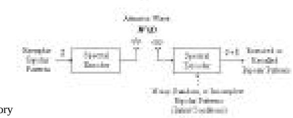


Observations:

- Notion of an *attractor wave* which replaces synaptic weight matrix
- *Bandwidth* now scales geometrically or polynomially with pattern dimension rather than physical dimension of the network
- Long-range connectivity is made *virtually* in the frequency domain
- Weighting factors may now be complex
- Aliasing may occur in memory formation and recall for non-complex formulations

Implications:

- Attractor is now separable from the decoder => telecommunication applications



- Spectral memory is now *volatile* memory
- Physical realizations now scale *linearly* with pattern dimension
- Like patterns may cancel each other if out of phase by ?
- Anti-aliasing constraints must be obeyed for non-complex (in-phase and quadrature formulations)

Overview of Spectral Formulations

Type of Pattern		Remarks
Bipolar (Digital modulators w/coding gain)		* Attractor wave caries **range space**
		* System point is "**pushed**" into basin
1. **Single-Pattern**	- *Auto*associative	* **Quantization** of recalled channels
	- *Hetero*associative	* No "**Bremsstrahlung**" to prevent
2. **Multi-Pattern**	- *Auto*associative	* **Content addressibility** is possible
	- *Hetero*associative	* Address addressibility **is not** possible
Real-Valued (Analog & digital MODEMs withcoding gain)		* Attractor wave caries **kernel space**
		* System point is "**pulled**" into basin
1. **Single-Pattern**	- *Auto*associative	* **No** quantization of recalled channels
	- *Hetero*associative	* "**Bremsstrahlung**" must be prevented
2. **Multi-Pattern**	- *Auto*associative	* **Content addressibility** is possible
		* **Address addressibility** is possible

Real-valued associative memory - Spatial formulation:

("Real-valued" refers to the pattern space, not the weighting coefficients)

Define a squared error function, *E*, and derive state update via gradient descent:

$$E ? \frac{1}{2}|\mathbf{e}|^2$$

where error vector, **e**, is defined as the residue of

$$\mathbf{e} ? \mathbf{W}\mathbf{v}; \qquad \mathbf{W} ? ?? \mathbf{I} ? \mathbf{M} ?$$

where **I** is the identity matrix, **v** is the state of the decoder, and **M** is the composite memory matrix from *n* orthonormal patterns, $\mathbf{s}_i$:

$$\mathbf{M} ? \sum_{i?1}^{n} ?_i \mathbf{M}_i ? \sum_{i?1}^{n} ?_i \mathbf{s}_i \mathbf{s}_i^T \quad (\textit{a priori} \text{ spectral composition})$$

$$\text{or} \quad \mathbf{M} ? \mathbf{U} ? \mathbf{U}^T \quad \text{where} \quad ? ? \begin{bmatrix} ?_1 & 0 & 0 & 0 \\ 0 & ?_2 & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & ?_n \end{bmatrix}$$

The state vector may be adapted along the negative gradient of the global error function:

$$\frac{d\mathbf{v}}{dt} ? ?c? E ? ?c\frac{d\mathbf{e}}{d\mathbf{v}}\mathbf{e} ? ?c\mathbf{We} ? ?c\mathbf{WWv}$$

And the eigenvalue may be adapted along the negative derivative of the global error function:

$$\frac{d?}{dt} ? ?c\frac{dE}{d?} ? ?c\mathbf{e}\frac{d\mathbf{e}}{d?} ? ?c\mathbf{e}\frac{d??\,\mathbf{I} ? \mathbf{M}?\mathbf{v}?}{d?} ? ?c\mathbf{ev} ? ?c\mathbf{Wvv}$$

Observation: Error calculation requires focusing of information from all over the network (requires long-range connectivity and quadratic scaling), but not the eigenvector and eigenvalue updates.

Observation: For single encoded memories, there is no need to adapt the eigenvalue, which acts as an address. Also **WW**=**W** in this case, thus the *error vector* ***is*** *the gradient.*

Probabilistic formulation of real-valued spectral associative memory:

- Pattern space is infinite, but the bias space is not => talk in terms of *bias space*.
- Intensity invariance : $r|?_d\rangle ? |?_d\rangle$
- Biased possibility space: e.g. $|?_d\rangle ? \sum_{?i} c_i|?_i\rangle;$ c_i may be complex
  => shows associatively "entangled" channels result from simple encodings
- Recall density matrix:

$$\mathbf{D} ? \sum_{?i} \mathbf{D}_i ? \sum_{?i} p_i|?_i\rangle\langle?_i|; \quad p_i ? \frac{|?_i|}{\sum_{?i}|?_i|}$$

a posteriori spectral decomposition

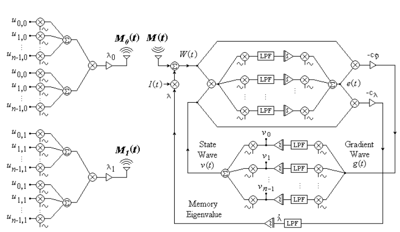
where $|?_i\rangle$ is the *a posteriori* orthonormal representation (not the pattern itself).

Observation: If observable D is "measured" by detecting one of its eigenvalues, one would say in quantum mechanics that a single measurement of D in state $|?_i\rangle$ produces a value of $?_i$ with probability $p_i ? Trace?|?_i\rangle\langle?_i|\mathbf{D}_i?$

And the expectation of ? relative to the i^{th} pattern: $Trace?|?_i\rangle\langle?_i|\mathbf{D}?$

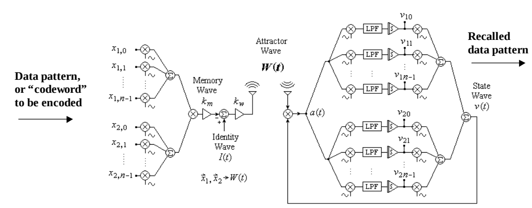
Real-valued associative memory in the frequency domain: Associative amplitude modulation

- *"Real-valued" refers to the pattern channel values, not the weighting coefficients*
- *Multi-channel, multi-carrier AM radio with content-addressibility*



Real-valued associative amplitude modulation

- *"Real-valued" refers to the pattern channel values, not the weighting coefficients*
- *Multi-channel, multi-carrier AM radio*



Spectral Coding

Associative amplitude modulation encodes data patterns into an attractor wave.

In-Phase (I):

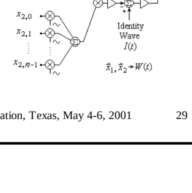
1. Synthesis: $s_{I1(2)}(t) ? \frac{1}{R}\sum_{i?0}^{n?1} x_{1?2?_i}\cos(?_{S1(2)i}t)$
2. Convolution: $M(t) ? M_{DSB}(t) ? k_m s_{I1}(t)s_{I2}(t)$
3. Kernel Generation: $W(t) ? k_w?I(t) ? M(t)?$, $I(t) ? I_{DSB}(t) ? \cos? ?\sum_{i?0}^{n?1}\cos(?_{ILi}t) ? \cos(?_{IUi}t)$

Quadrature (I/Q):

1. Synthesis: $s_{Q1(2)}(t) ? \frac{1}{R}\sum_{i?0}^{n?1} x_{1?2?}\sin(?_{S1(2)i}t)$
2. Convolution: $M(t) ? M_{SSB}(t) ? k_m ?s_{I1}?t?s_{I2}?t?? s_{Q1}?t?s_{Q2}?t??$
3. Kernel Generation: $W(t) ? k_w?I(t) ? M(t)?$, $I(t) ? I_{SSB}(t) ? \cos? ?\sum_{i?0}^{n?1}\cos(?_{ILi}t)$ $\quad \cos? ?? \frac{\mathbf{x}_1^T\mathbf{x}_2}{|\mathbf{x}_1||\mathbf{x}_2|}$

Complex:

1. Synthesis: $?_{s1(2)}(t) ? k_{s1(2)}\sum_{i?0}^{n?1} x_{1?2?_i}e^{j?_{S1(2)_i}t};\quad k_{s1(2)} ? \frac{1}{|\mathbf{x}_{1(2)}|}$
2. Convolution: $?_M(t) ? k_m ?_{s1}(t)?_{s2}(t)? ? k_m\sum_{i?0}^{n?1}\sum_{k?0}^{n?1} x_{1_i}x_{2_k}e^{jt?_{S1i}??_{S2k}?}$
3. Kernel Generation: $?_W(t) ? k_w?_I(t) ?? _M(t)?$, $?_{IL}(t) ? k_i\sum_{i?0}^{n?1} e^{j?_{IL_i}t}\quad k_i ? \cos? ?$



Coding Coefficients for Normalized Signal Power

Table 1. Coding coefficients for normalized signal power.

		Bipolar Patterns					Real-Valued Patterns				
Association	**Coding**	k_{s1}	k_{s2}	k_m	k_i	k_w	k_{s1}	k_{s2}	k_m	k_i	k_w
Auto	**Complex**	$\frac{1}{\sqrt{n}}$	$\frac{1}{\sqrt{n}}$	$\sqrt{n}$	$\frac{1}{\sqrt{n}}$	$\sqrt{\frac{P_s}{n-1}}$	$\frac{1}{\lvert\mathbf{x}\rvert}$	$\frac{1}{\lvert\mathbf{x}\rvert}$	1	1	$\sqrt{\frac{P_s}{n-1}}$
"	**In-Phase (I)**	$\frac{1}{\sqrt{n}}$	$\frac{1}{\sqrt{n}}$	$2\sqrt{n}$	$\frac{1}{\sqrt{n}}$	$\sqrt{\frac{P_s}{n-1}}$	$\frac{1}{\lvert\mathbf{x}\rvert}$	$\frac{1}{\lvert\mathbf{x}\rvert}$	2	1	$\sqrt{\frac{P_s}{n-1}}$
"	**Quadrature (I/Q)**	$\frac{1}{\sqrt{n}}$	$\frac{1}{\sqrt{n}}$	$\sqrt{n}$	$\frac{1}{\sqrt{n}}$	$\sqrt{\frac{2P_s}{n-1}}$	$\frac{1}{\lvert\mathbf{x}\rvert}$	$\frac{1}{\lvert\mathbf{x}\rvert}$	1	1	$\sqrt{\frac{2P_s}{n-1}}$
Hetero	**Complex**	$\frac{1}{\sqrt{n}}$	$\frac{1}{\sqrt{n}}$	1	0	$\sqrt{P_s}$	$\frac{1}{\lvert\mathbf{x}_1\rvert}$	$\frac{1}{\lvert\mathbf{x}_2\rvert}$	1	$\cos(\theta)$	$\sqrt{\frac{P_s}{(n-2)\cos^2(\theta)+1}}$
"	**In-Phase (I)**	$\frac{1}{\sqrt{n}}$	$\frac{1}{\sqrt{n}}$	2	0	$\sqrt{P_s}$	$\frac{1}{\lvert\mathbf{x}_1\rvert}$	$\frac{1}{\lvert\mathbf{x}_2\rvert}$	2	$\cos(\theta)$	$\sqrt{\frac{P_s}{(n-2)\cos^2(\theta)+1}}$
"	**Quadrature (I/Q)**	$\frac{1}{\sqrt{n}}$	$\frac{1}{\sqrt{n}}$	1	0	$\sqrt{2P_s}$	$\frac{1}{\lvert\mathbf{x}_1\rvert}$	$\frac{1}{\lvert\mathbf{x}_2\rvert}$	1	$\cos(\theta)$	$\sqrt{\frac{2P_s}{(n-2)\cos^2(\theta)+1}}$

Spectral Recall

Recurrent associative amplitude modulation focuses extrinsic information distributed all over the attractor band for subsequent spectral analysis, or **down conversion.**

1. Spectral Synthesis:

   $? \,_v(t) ? \sum_{l?1}^{2} \sum_{i?0}^{n?1} v_{li} e^{j? \,_{li} t}$

2. Spectral Convolution:

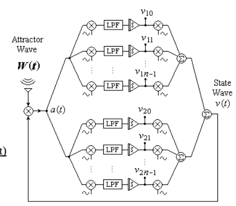
   $? \,_a(t) ?? \,_v(t)? \,_W(t)?$

3. Spectral Analysis (Down Conversion):

   $\frac{d\mathbf{u}_{1(2)}}{dt} ? ? \mathrm{LPF}?\mathrm{Re}? \,_a(t) e^{? j? \,_{A1(2)} t} ??$ (spectral gradient)

4. State Update:

   $\mathbf{v} ? f(\mathbf{u}) \quad \mathbf{u} ? ?\dot{\mathbf{u}} dt$

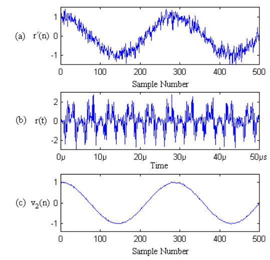


Performance: Noise Immunity (RV-SAMs)

Transient response of a 2-D autoassociative network for a sampled sine wave input:

(a) what the original signal would have been if sent directly over the noisy channel with 10dB SNR,
(b) noisy attractor wave of the first data sample, and
(c) recalled sine wave.

Comparing (c) with (a) illustrates the built-in noise immunity.



Performance: Real-Valued Recall

- Real-valued coding and recall of Marylin Monroe

Performance: Settling Time & Accuracy of Recall

Butterworth recall (B) of various order, with and without notch filter (N). $?_{min}$=40dB

Bits Accuracy	B3 T_s (?s)	B4 T_s (?s)	B5 T_s (?s)	B6 T_s (?s)
1	30	27	27	27
2	30	58	55	55
3	85	66	78	78
4	123	96	90	83
5	142	123	112	107
6	Inf	Inf	163	147
Bits Accuracy	B3 + N T_s (?s)	B4 + N T_s (?s)	B5 + N T_s (?s)	B6 + N T_s (?s)
1	19	18	17	17
2	20	18	17	17
3	50	42	38	37
4	55	45	41	42
5	80	64	55	43
6	Inf	93	Inf	81

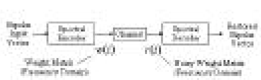
Chebyshev recall (C) of various order, with and without notch filter (N). $?_{min}$=40dB

Bits Accuracy	C3 T_s (?s)	C4 T_s (?s)	C5 T_s (?s)	C6 T_s (?s)
1	26	21	24	24
2	27	54	67	53
3	94	112	75	105
4	155	Inf	128	157
5	Inf	Inf	148	Inf
6	Inf	Inf	Inf	Inf
Bits Accuracy	C3 + N T_s (?s)	C4 + N T_s (?s)	C5 + N T_s (?s)	C6 + N T_s (?s)
1	18	14	17	15
2	25	34	17	36
3	52	41	43	45
4	55	64	47	69
5	Inf	89	75	104
6	Inf	Inf	Inf	Inf

- Inserting a **notch filter** at ?? in cascade with the LPF allows the beginning of the stopband to be doubled, which increases the width of the passband and **increases the speed of recall**.
- **Butterworth recall filters** appear to give the fastest response for higher accuracy of recall.
- **Data rates**: ~ 30kHz/ch (Nyquist) or 60kHz/ch. (baud) for 1 or 2 bits accuracy
  ~ 11kHz/ch (Nyquist) or 22kHz/ch. (baud) for 5 or 6 bits accuracy

•All results were obtained with a *beat frequency* of ?? = 2? 50k rad/sec.

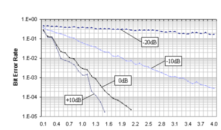
Single-Pattern Telecommunications



- But **single patterns** may be transmitted, one at a time, across a noisy channel, to a spectral neural decoder with **spurious-free recall** due to a **lack of interference** between patterns.
- Due to volatility, basins of attraction disappear when transmission of the attractor wave ceases.
- **Code rate** depends on virtual architecture (auto/hetero), which allows **dynamic reconfiguration**.
- Digital modulators with coding gain! MODEM-CODECs

Bit error rate (BER) for various degrees of SNR, decoding period, & spectrum spread for single-sideband attractor waves formed by quadrature coding

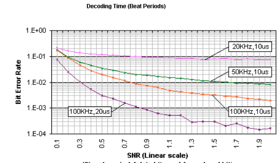
Decoding Time / Data Rate



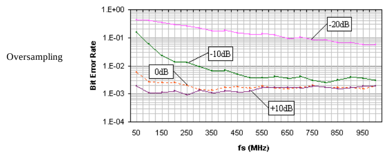
Increasing *beat frequency*

The more "*spread-out*" the spectrum is, the higher the noise immunity (*controlled by the beat frequency*)

Decoding Time / Data Rate



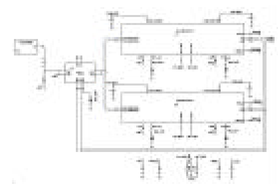
Bit error rate (BER) vs. oversampling



Conclusion:
Limited energy/information is carried per unit time by the attractor wave and Shannon's limit is obeyed

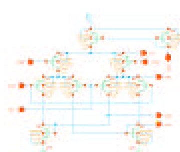
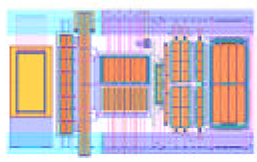
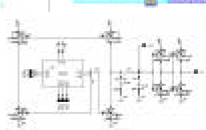
CMOS Implementation of a Bipolar Decoder

- **Addition** may be implemented in **current domain**
- **Mixers** may be implemented with **Gilbert cells**
- **Integrators** may be implemented by **OTA-C** stages
- **Oscillators** may be generated off-chip or implemented by **ring oscillators** (must be phase-locked)



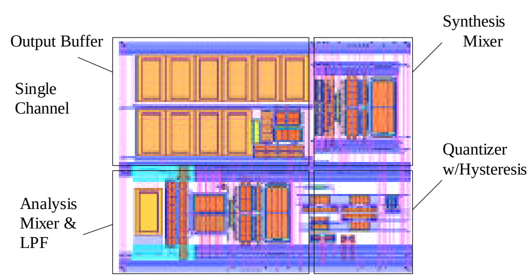
CMOS Hebbian Filter: Mixer + LPF

- An additional transconductance stage followed by a capacitor and two inverters

CMOS Decoder Channel

Technology: .35? TSMC double-poly, quadruple metal n-well, 500u x 400u

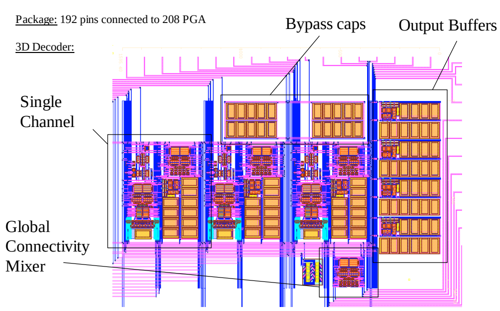


3D CMOS Decoder

Technology: .35? TSMC double-poly, quadruple metal n-well, 1.6mm x 1mm

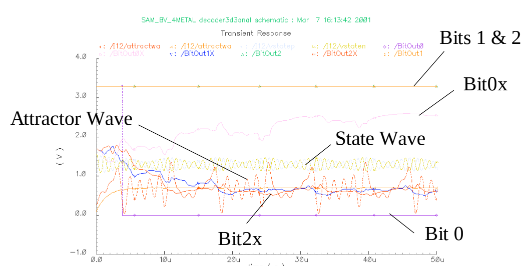
Package: 192 pins connected to 208 PGA

3D Decoder:



CMOS Implementation: Transient Response

Bit 0 encoded low, Bit 1 encoded high, Bit 2 encoded high



References

Journal Papers:

- R. Spencer, **"Bipolar Spectral Associative Memories,"** *IEEE Trans. on Neural Networks*. To appear.
- R. Spencer, **"Spectral Associative Memories: Content-Addressable Communications,"** *Int'l. J. of Smart Engineering System Design.* To appear.

Conference Papers:

- R. Spencer, **"Nonlinear Heteroassociative Spectral Memories: Virtually Partitioned Neural Encoders and Decoders for Digital Communications,"** *2000 Int'l. Symposium on Computational Intelligence, Int'l. Congress on Intelligent Systems and Applications*, Wollongong, Australia, Dec. 12-15, 2000.
- R. Spencer, **"Nonlinear Spectral Associative Memories: Neural Encoders and Decoders for Digital Communications,"** In: *Intelligent Engineering Systems Through Artificial Neural Networks*, C.H. Dagli, A.L. Buczac, J. Ghosh, M.J. Embrechts, O. Ersoy, and S. Kercel (Eds.), ASME Press, Vol. 10., pp. 971-976, 2000. ISBN 0-7918-0161-6. Best Paper Award.

Submitted/Working Papers:

- R. Spencer, H. Dinc, & T. Sumesaglam, **"Associative Amplitude Modulation with Built-In Noise Immunity,"** submitted to *ICECS 2001*, Malta.
- R. Spencer, **"Multi-Pattern Real-Valued Spectral Associative Memories,"** submitted to *IJCNN 2001.*