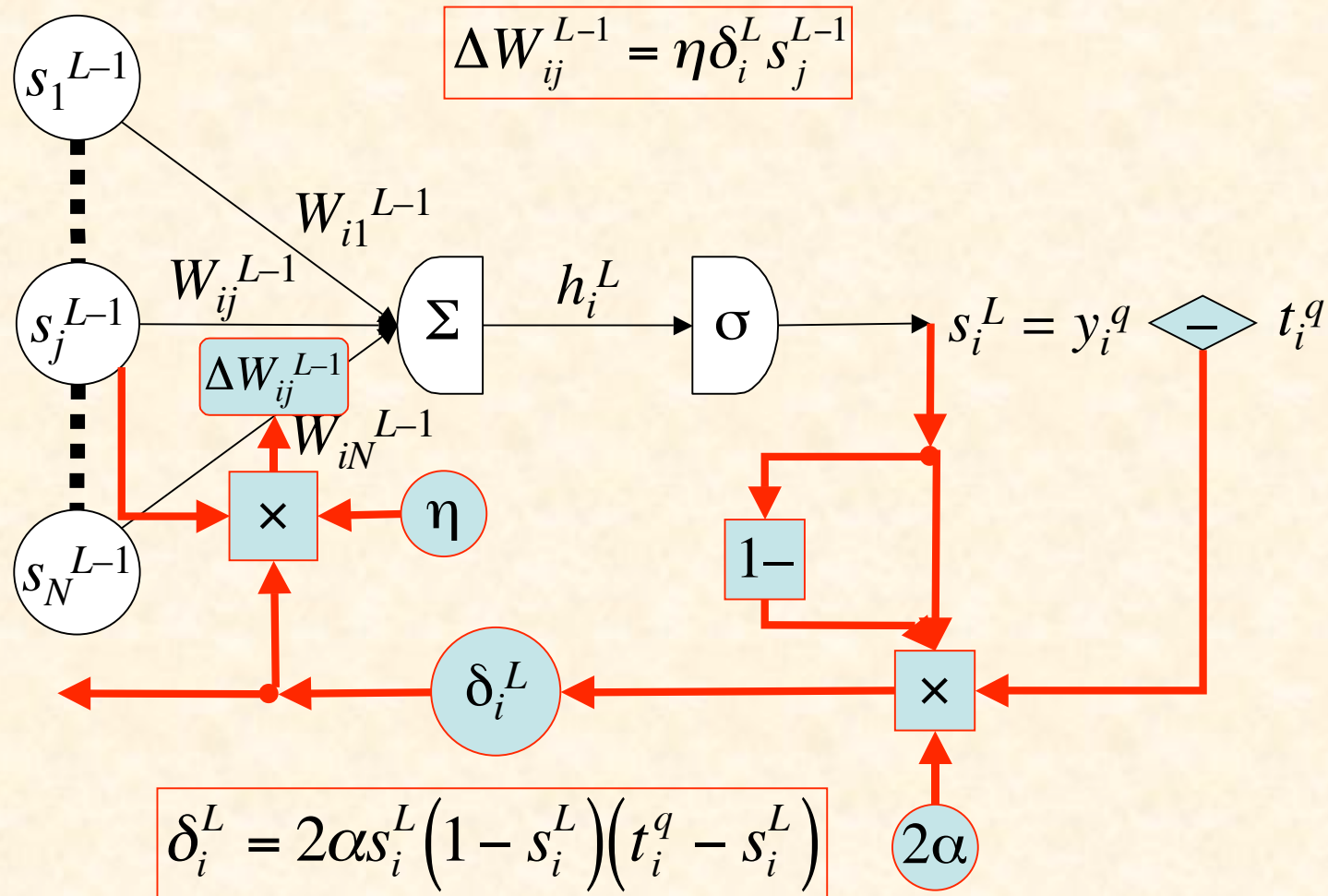
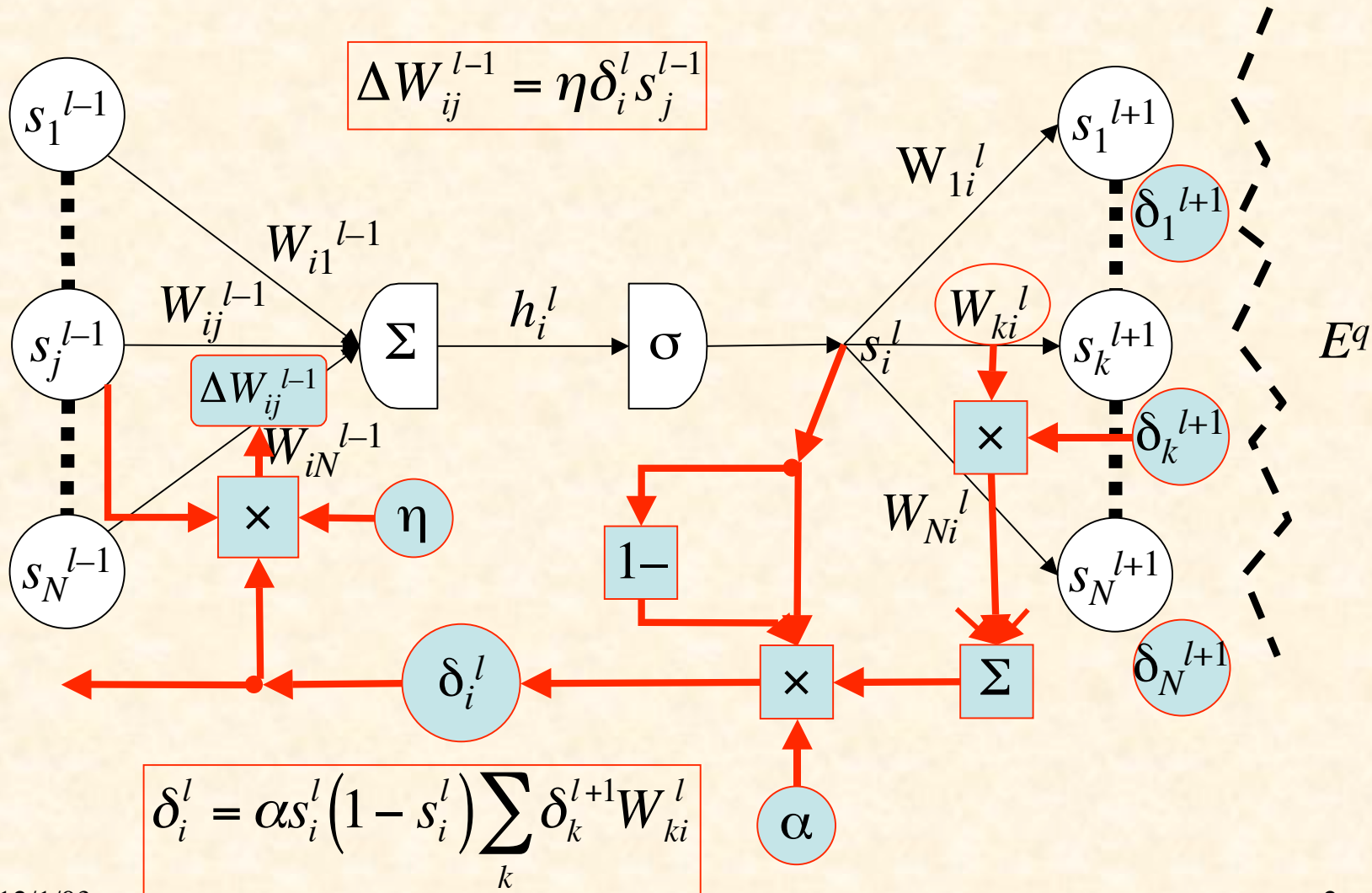


Output-Layer Computation



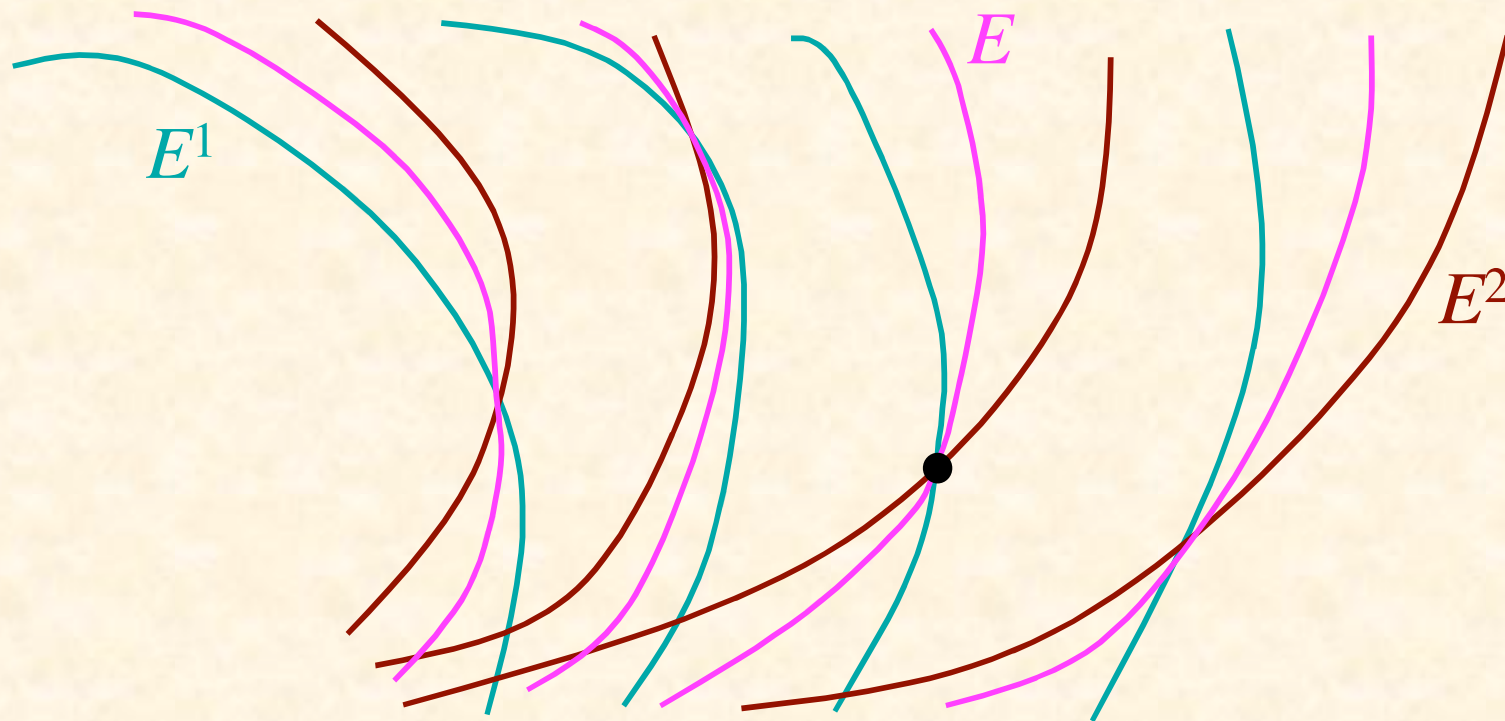
Hidden-Layer Computation



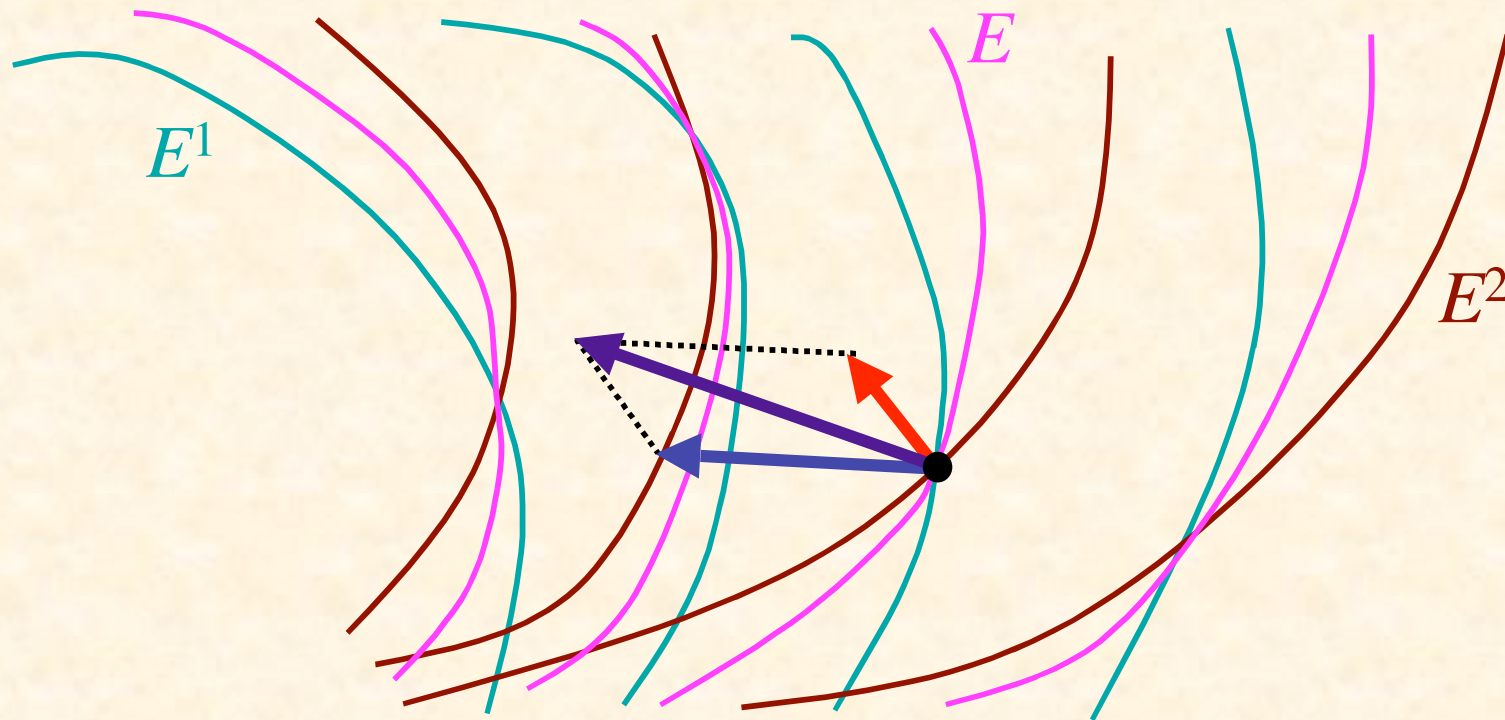
Training Procedures

- Batch Learning
 - on each *epoch* (pass through all the training pairs),
 - weight changes for all patterns accumulated
 - weight matrices updated at end of epoch
 - accurate computation of gradient
- Online Learning
 - weight are updated after back-prop of each training pair
 - usually randomize order for each epoch
 - approximation of gradient
- Doesn't make much difference

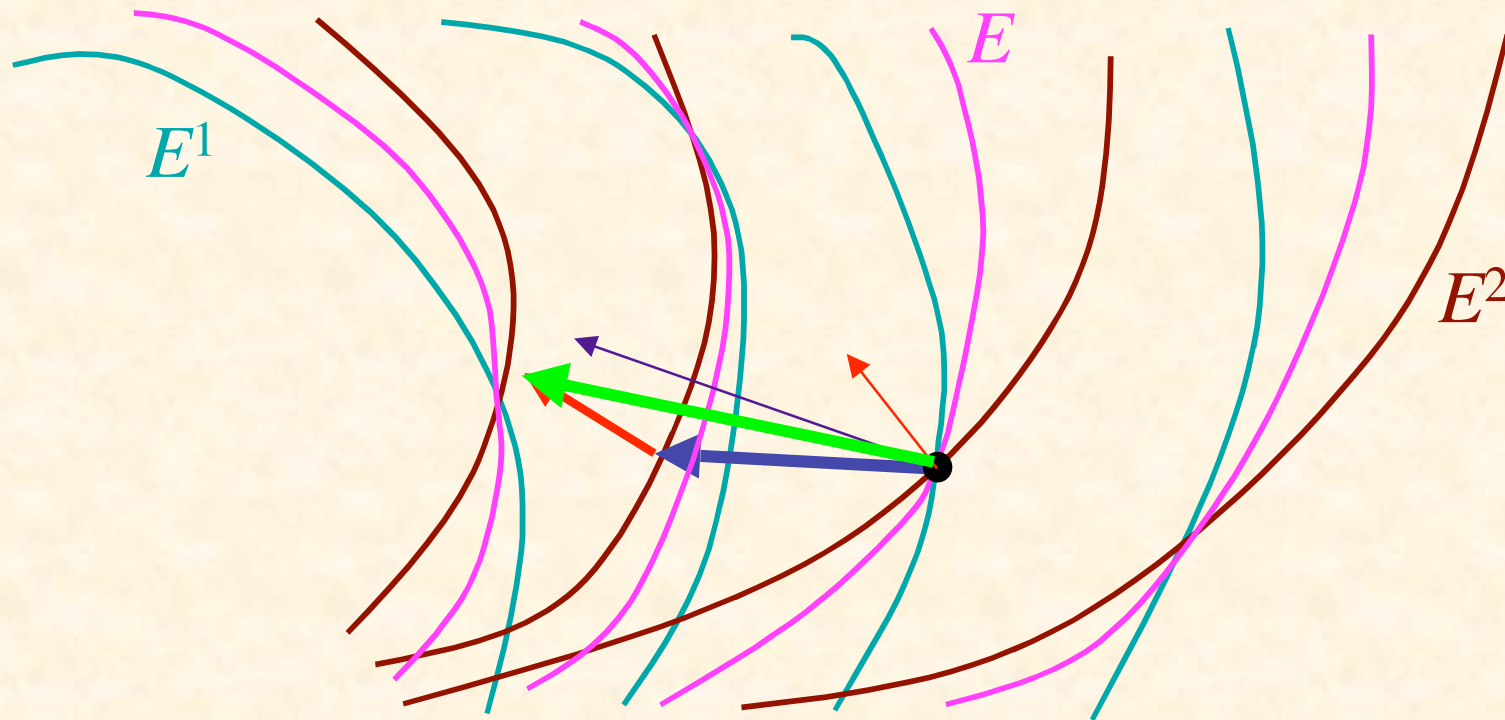
Summation of Error Surfaces



Gradient Computation in Batch Learning



Gradient Computation in Online Learning



VIII. Review of Key Concepts

Complex Systems

- Many interacting elements
- Local vs. global order: entropy
- Scale (space, time)
- Phase space
- Difficult to understand
- Open systems

Many Interacting Elements

- Massively parallel
- Distributed information storage & processing
- Diversity
 - avoids premature convergence
 - avoids inflexibility

Complementary Interactions

- Positive feedback / negative feedback
- Amplification / stabilization
- Activation / inhibition
- Cooperation / competition
- Positive / negative correlation

Emergence & Self-Organization

- Microdecisions lead to macrobehavior
- Circular causality (macro / micro feedback)
- Coevolution
 - predator/prey, Red Queen effect
 - gene/culture, niche construction, Baldwin effect

Pattern Formation

- Excitable media
- Amplification of random fluctuations
- Symmetry breaking
- Specific difference vs. generic identity
- Automatically adaptive

Stigmergy

- Continuous (quantitative)
- Discrete (qualitative)
- Coordinated algorithm
 - non-conflicting
 - sequentially linked

Emergent Control

- Stigmergy
- Entrainment (distributed synchronization)
- Coordinated movement
 - through attraction, repulsion, local alignment
 - in concrete or abstract space
- Cooperative strategies
 - nice & forgiving, but reciprocal
 - evolutionarily stable strategy

Attractors

- Classes
 - point attractor
 - cyclic attractor
 - chaotic attractor
- Basin of attraction
- Imprinted patterns as attractors
 - pattern restoration, completion, generalization, association

Wolfram's Classes

- Class I: point
- Class II: cyclic
- Class III: chaotic
- Class IV: complex (edge of chaos)
 - persistent state maintenance
 - bounded cyclic activity
 - global coordination of control & information
 - order for free

Energy / Fitness Surface

- Descent on energy surface / ascent on fitness surface
- Lyapunov theorem to prove asymptotic stability / convergence
- Soft constraint satisfaction / relaxation
- Gradient (steepest) ascent / descent
- Adaptation & credit assignment

Biased Randomness

- Exploration vs. exploitation
- Blind variation & selective retention
- Innovation vs. incremental improvement
- Pseudo-temperature
- Diffusion
- Mixed strategies

Natural Computation

- Tolerance to noise, error, faults, damage
- Generality of response
- Flexible response to novelty
- Adaptability
- Real-time response
- Optimality is secondary

Student Course Evaluation