Index

**bold** fonts mark sections or subsections in the text

**italic** fonts mark margin labels

A-current, *see* conductances

action potential, 4, *see also*

  - Connor-Stevens model
  - Hodgkin-Huxley model

propagation along a myelinated axon, 222

propagation along an unmyelinated axon, 220

propagation velocity, 222, 223

refractory period, *see* refractory period

  - saltatory propagation, 222

  - two moving in opposite directions, 222

activation, *see* conductances

activation function $F(I)$, 234, *see also* firing-rate models

  - Connor-Stevens model, 197, *see also* Connor-Stevens model

  - integrate-and-fire models, 164, *see also* integrate-and-fire models

activity-dependent development, 293–309, *see also*

  - causal models
  - feature-based models

  - arbor function, 300

  - of ocular dominance, 298

  - ocular dominance stripes, 302

  - orientation domains, 309

  - orientation selectivity, 299

actor, 347, *see also*

  - critic

  - delayed rewards, problem of policy

  - action matrix, 351

  - action values, 341

actor, direct, 344, 349

as Monte Carlo method, 357

average reward ($r$), 344

  - covariance rule, three-term, 351

  - delayed reward, temporal difference learning rule for, 349, 357

  - exploration-exploitation control by $\beta$, 345

  - immediate reward, learning rule, 345

  - model of basal ganglia, 351

  - multiple actions, 345

  - reinforcement comparison, 345, 349

  - stochastic gradient ascent, 344

  - vs. indirect actor, 345

actor, indirect, 342

action values, 342

  - exploration-exploitation control by $\beta$, 341

  - vs. direct actor, 345

actor-critic algorithm, 350, *see also*

  - actor

  - critic

  - dynamic programming

  - generalizations, 350

  - policy iteration, 356, *see also*

    - policy iteration

    - adaptation, *see* spike-rate
adaptation
after-hyperpolarization, see conductances
aliasing, 60
α function, 14, 183, 188, see also synaptic models
AMPA receptor, see synaptic conductances
amplification, selective, see also recurrent networks for orientation, 252, see also simple cell
linear, 246
nonlinear, 251
oscillatory, 272
analysis by synthesis, 396, see also causal models
anti-Hebb rule, 310, see also contrastive Hebb rule
Goodall rule
Hebb rule
learning rules
antidromic propagation, 221, see also action potential
arbor function, 300
associability, 333, see also learning rate
associative memory, 261
capacity, 262, 264
covariance learning rule, 264
recall, 262
spurious fixed points, 263
synaptic weight matrix, 263–264
asymptotic consistency, see estimation theory
attenuation, see dendrites, electrical properties
membrane, electrical properties
autoassociative memory, see associative memory
autocorrelation, see also correlation
cross-correlation
spike-train $Q_{su}$, 28
stimulus $Q_{su}$, 22, 47
autocovariance; 28
average firing rate $\langle r \rangle$, 11
axon, 4
azimuth, see retinal coordinate
system
backpropagation, 329
balanced excitation and inhibition, see integrate-and-fire models, irregular firing mode
bandwidth, 64
basal ganglia, 351
basin of attraction, see point attractor
stability, network
basis functions, 244, 317, see also function approximation
complete, 317
overcomplete, 317
basis, vector space, 402
Bayes theorem, 88, 87–88, 364
Bayesian decision, see decision theory
Bayesian inference, see estimation theory
BCM learning rule, see sliding threshold learning rule
Bernoulli distribution, 417
bias (of estimator), see estimation theory
e bifurcation, see stability, network
bit, 124, see also entropy
information
Boltzmann machine, 273, 322
Boltzmann distribution, 274
contrastive Hebb rule, 324, 325, 326
ergy function, 274, 323, 325
Gibbs sampling, 274, 322, 325, 326
Glauber dynamics, 274
Lyapunov function, 276
mean-field approximation, 274, 325, 372
mean-field distribution, 275
partition function, 274, 323, 325
sleep phase, 324, 326
supervised learning, 322
unsupervised learning, 325
wake phase, 324, 326
branching, see cable theory
compartmental models
buffer, Ca$^{2+}$, 203
bursts, 34, 198, 199, 200, see also
calcium spike
Bussgang’s theorem, 51, 83
cable equation, 203, 204, 206
boundary conditions, 206
constant radius segment, 207
impulse ‘velocity’, 211
solution for an infinite cable, 208
solution for an isolated
branching node, see also
compartmental models
solution for an isolated
branching node, 211
voltage attenuation and
frequency, 217
cable theory, 203, 203–223, see also
cable theory, 3/2 law
branching node, 206, see also
linear cable, 206
morphoelectronic transform,
215
voltage attenuation, 212, 215
calcium buffer, 203
calcium conductances, see
conductances, Ca$^{2+}$
calcium spike, 199
calculus of variations, 81
Cauchy distribution, 380, 417
Cauchy-Schwarz inequality, 120
causal kernel, see filter, causal
causal models, 361, see also
causal models
factor analysis
generative models
Helmholtz machine
ICA
K-means algorithm
mixture of Gaussians
recognition models
sparse coding model
unsupervised learning
analysis by synthesis, 360
cause $v$, 361
causes as representations, 360
Gaussian and non-Gaussian, 377
generation, 359, 361
heuristics, 360, 363, see also
re-representation
hidden variable, 360
hierarchical, 382
invertible, 367
latent variable, 360
noninvertible, 367
projective fields, 382
recognition, 360
structure learning, 362
summary of probability
distributions, 368
unsupervised learning, 363
center-surround structure, 54,
77–78, see also receptive field
for continuous attractor, 259
in activity-dependent
development, 304
information theory and, 140
central limit theorem, 417
central pattern generator, 201
cerebral system, 97, see also neural
coding
channels, see ion channels
chaos, 410
classical conditioning, 331, 332, see
also
critic
delayed rewards, problem of
instrumental conditioning
Rescorla-Wagner rule
acquisition, 333
blocking, 334
conditioned stimulus and
response, 332
extinction, 333
inhibitory conditioning, 335
overshadowing, 335
partial reinforcement, 334
secondary conditioning, 336
unconditioned stimulus and
response, 332
classification, see perceptron
clustering, see mixture of
Gaussians
coding, see
neural coding
re-representation
Index

coefficient of variation, see
  interspike intervals, coefficient of variation
coincidence detection, 183
compartamental models, see also
cable theory
conductance-based, 195
equivalent circuit, 162, 213–215
morphoelectrotonic transform, 215
multi-compartment, 217, 225
Rall model, 212
reduction, 218
single compartment, 161
competitive Hebb rule, 304, 306, see also
  Hebb rule
  learning rules
  recurrent networks, competitive dynamics
  competition and cooperation, 306
feature-based rule, 308
ocular dominance stripes, 306
complementary error function, see error function
complex cell, 54, 74, see also
  Hubel-Wiesel model
energy model, 76
feedforward model, 80
recurrent model, 254
complex exponential, 405
complex logarithmic map, 57, see also cortical magnification factor
conductances, see also
  Connor-Stevens model
  Hodgkin-Huxley model
  ion channels
  Ca$^{2+}$, L, T, N and P type, 199
  Ca$^{2+}$, transient, 198, 225
  Ca$^{2+}$-dependent, 167
  K$^{+}$, A-type current, 197, 197–198, 200
  K$^{+}$, delayed-rectifier, 168, 171, 175, 197
  K$^{+}$, Ca$^{2+}$-dependent, 201, 225
  Na$^{+}$, fast, 172, 197
  activation, 168, 172
variable $m$, 172
variable $n$, 168
active, 161
after-hyperpolarization (AHP), 201
deactivation, 168, 172
deinactivation, 172
delayed-rectifier, see conductances, K$^{+}$
hyperpolarization-activated, 172
inactivation, 172
variable $h$, 172
integration of gating variables, 192
noninactivating, see conductance, persistent
  passive, 161
  persistent, 168
reversal potential, see reversal potential
shunting, 160, 189
shunting conductances and division, 189
synaptic, see synaptic conductances
transient, 171
voltage-dependent, 166, 167
Connor-Stevens model, 196, 224, see also Hodgkin-Huxley model
  A-current, role in, 197
continuous attractor, see also
  stability, network
linear, 247
nonlinear, 251–259
continuous labeling, 136, 240, see also
  firing-rate models
  recurrent networks
complex cell, 255
density of coverage $\rho_0$, 240
linear recurrent network, 248
nonlinear recurrent network, 251
oscillatory network, 272
simple cell, 252
contrast, 58
contrast saturation, 73, 393
contrastive Hebb rule, 322, see also
  anti-Hebb rule
Boltzmann machine
delta rule
Hebb rule
learning rules
Boltzmann machine, 324, 326
delta rule, 321
wake and sleep phases, 324
convolution, 406
coordinate transformation, 241, see
also gain modulation
correlation, see also
autocorrelation
cross-correlation
firing-rate stimulus \( Q_{ss} \), 20, 47
matrix \( Q \), 286
reverse, see reverse correlation
correlation code, see neural coding
correlation-based learning rule, 286, see also
Hebb rule
learning rules
principal component
correlation matrix \( Q \), 286
covariance learning rule,
compared with, 288, 298
cortex, see neocortex
cortical magnification factor, 56, see also complex logarithmic map
cortical map, 293, 309, see also
ocular dominance stripes
orientation domains
retinal coordinate system
pattern formation, 293
counterphase grating, 58, 66
covariance learning rule, 287, 288, see also
Hebb rule
learning rules
principal component
sliding threshold learning rule, see plasticity
correlation-based learning rule, compared with, 288, 298
covariance matrix \( C \), 288
instability, 288
three-term for direct actor, 351
covariance, of random variables, 416
Cox process, 34
Cramér-Rao bound, 108, 120, see also estimation theory
Fisher information
Crank-Nicholson method, 226, see also numerical methods
critic, 347, see also
actor
delayed rewards, problem of policy
as Monte Carlo method, 356
model of dopaminergic activity, 339
prediction \( v(t) \), 336, see also
policy iteration, policy evaluation
temporal difference learning
rule, 348, 356
cross-correlation, 28, 314, see also
autocorrelation
correlation
current, see intracellular electrical properties
leakage current
membrane current
Dale’s law, 239
deactivation, see conductances
decision theory, 89, see also
discriminability
estimation theory
Neyman-Pearson lemma
signal detection theory
Bayesian decision, 96
false alarm rate, 91
hit rate, 91
likelihood ratio test, 95, 112
loss function, 96
perceptron, 314, see also
perceptron
test power, 91
test size, 91
two-alternative forced choice, 93
decoding, see
estimation theory
neural decoding
decorrelation, 135
anti-Hebb rule, 311
Goodall rule, 311
deinactivation, see conductances
del operator \( \nabla \), 402
delayed rewards, problem of, 332, 340
actor-critic solution, 349
discounting, 351
dynamic programming solution, see dynamic programming policy iteration
Markov decision problem, 350, 355
maze task, 347
recursive principle, 337, 336
secondary conditioning, 336
sequential action choice, 346
summed future reward, 336, 347, 355
temporal difference solution, 337
water maze task, 352
delayed-rectifier, see conductances, K+
δ function, 9, 404
delta rule, 319, 320, 381, see also contrastive Hebb rule
gradient descent learning rules
perceptron learning rule
Rescorla-Wagner rule
temporal difference learning rule
as Rescorla-Wagner rule, 333
contrastive Hebb rule, 321
function approximation, 320
dendrites, 4
dendrites, electrical properties, see also compartmental models
membrane, electrical properties
apical versus basal dendrites, 217
electrotone compactness, 156, 218
electrotone length λ, 207, 208, 213
input resistance, 209
morphoelectrotone transform, 215
density estimation, 322, 368
and optimal coding, 369
maximum likelihood, 323, 369
supervised learning, 323, see also supervised learning
unsupervised learning, 325, see also unsupervised learning
depolarization, 4, 160
depression, synaptic, see LTD
plasticity
diagonalization, matrix, 403
difference equation, 413
mode, 413
difference of Gaussians, see center-surround structure
differential entropy, see entropy, continuous variable
differential equation, 410, see also recurrent networks
stability, network mode, 411
diffusion equation, 208, see also cable equation
dimension reduction, see re-representation
Dirac δ function, see δ function
direct actor, see actor, direct
direction selectivity, 72, see also receptive field, nonseparable
directional derivative, 402, see also del operator ∇
discriminability d’, 91, 96, see also decision theory
defined by maximum likelihood, 112
non-Gaussian distributions, 94
disparity, 16
dopaminergic activity, 339
and reward, 339
temporal difference prediction error model of, 339, see also critic
dot product, 99, 399
driving force, 160, see also conductances
reversal potential
dynamic programming, 347, see also policy iteration
Bellman equation, 355
eccentricity, see retinal coordinate system
efficiency (of estimator), see estimation theory
eigensystem, 402
degeneracy, 402
eigenfunction \( \epsilon(t) \), 405
eigenvalue \( \lambda \), 245, 402
eigenvector \( \mathbf{e} \), 245, 402
eigenvector expansion, 245
eigenvector orthonormality, 403
eigenvector, principal, 294, 376, see also PCA
principal component
translation invariance, 304, 405
elastic net, see feature-based models
electrical circuit theory, 413
electrodes
extracellular, 7
patch, 6
sharp intracellular, 6
electrotomic length, see
cable theory
dendrites, electrical properties
morphoelectricotropic transform
EM (expectation maximization), 364, see also density estimation
as coordinate ascent, 370
E phase, 365
likelihood lower bound \( J \), 369
M phase, 365
theory of, 369
energy model, 76, see also complex cell
entropy, 124, 125, see also
entropy rate
information
noise entropy
continuous variable, 130
entropy maximization, 130, see also
histogram equalization
information maximization
effect of noise, 138–141
populations of neurons, 133
single neuron, 131
entropy rate, 145
estimation by direct method, 146–147
Poisson process, 146
EPSC (excitatory postsynaptic current), 181, see also synaptic conductances
EPSP (excitatory postsynaptic potential), 204, see also synaptic conductances
equilibrium point, see
continuous attractor
fixed-point
point attractor
stability, network
equilibrium potential, 158, see also
reversal potential
equivalent cable, see
compartmental models, Rall model
equivalent circuit, see
compartmental models
error function, 94
error-correcting learning rules, 318, see also
contrastive Hebb rule
delta rule
learning rules
perceptron learning rule
temporal difference learning rule
estimation theory, see also
Cramér-Rao bound
decision theory
Fisher information
neural decoding
asymptotic consistency, 109
Bayesian inference, 102
bias, 107
bias-variance tradeoff, 109
efficiency, 109
estimation error, 108
maximum \( a \ posteriori \) (MAP)
estimate, 107
inference, 103
maximum likelihood (ML)
estimate, 106
inference, 103
recurrent network ML inference, 258
unbiased estimator, 101, 108, 109
variance (of estimator), 108
Euler method, 226, see also
numerical methods
Index

reverse Euler method
excitatory postsynaptic current, 181, see also synaptic conductances
excitatory-inhibitory networks, see recurrent networks
exercises, xiv, see mitpress.mit.edu/dayan-abbott
expectation maximization, see EM
exponential distribution, 417
double exponential distribution, 380
eye position integration, 248, see also integrator, neural
facilitation, synaptic, see LTP
plasticity
factor analysis, 366, 374, see also causal models
EM algorithm, 367, see also EM
factorial re-representation, 374
generation, 374
learning rule, 395
nonlinear, see ICA
sparse coding model
PCA limit, 375
recognition model, 374
vs. PCA, 376–377
factorial code, see neural coding
factorial representation, see re-representation
false alarm rate, see decision theory
Fano factor, 27, 27, 32, see also spike count, distribution
feature selectivity, see neural coding
tuning curve
receptive field
feature-based models, 307, see also activity-dependent development
competitive Hebb rule, 308
elastic net, 308
features in, 307
ocular dominance stripe
development, 309
orientation domain
development, 309
self-organizing map, 308
synaptic weights in, 307
Fechner’s law, 18
feedforward networks, 238, 241, 301, see also function approximation
coordinate transformation, see coordinate transformation
filter, see also noise filter
spike decoding
Wiener kernel
causal, 14
linear, 13, 337, 404
filter kernel, 13, 46
firing rate, 8, 24, see also spike count rate
neural response function
r(t), 10
activation function F(I_s), 234
average, 11
estimate r_{est}(t), 46
estimate with static nonlinearity, 49
estimation, 11, 45, see also neural decoding
instantaneous, 164
interspike-interval, 164
time-dependent, 10
firing-rate models, 231
activation function F(I_s), 234
comparison with spiking models, 230, 274
continuous model, 240
current dynamics, 235, 260
excitatory-inhibitory networks, 239
firing rate, relationship to synaptic current of, 234
mean-field Boltzmann machine, 274
rate dynamics, 236, 238
sparse coding model, 380
Fisher information, 108, 109, 130, see also estimation theory for neural population, 110
fixed point, see also continuous attractor
point attractor
synaptic normalization
synaptic competition
averaged, 286
classification, 315
competition deficit in, 284
dynamic solution, 294
function approximation, 318
instability, 284, 287, 288
ocular dominance development, 298
ocular dominance stripe
development, 302
orientation selectivity
development, 299
perceptron, 315
subtractively normalized, 290, 296
supervised learning, 313
timing-based, 291, 292
Helmholtz machine, 387, see also
Boltzmann machine
causal models
approximate free energy \( - \tilde{F} \), 388
approximate probabilistic
recognition, 387, see also Monte
Carlo method
free energy, \( - F \), 388
generation, 387
learning rule, 395
probabilistic recognition, 389
wake-sleep algorithm, 389
hidden variable, see causal models
histogram
interspike interval, 32
spike-time, 12, see also firing rate
histogram equalization, 132, see also
entropy maximization
anti-Hebb rule, 311
Goodall rule, 311
hit rate, see decision theory
Hodgkin-Huxley model, see also
Connor-Stevens model
of action potential, 173, 198, 220
of delayed-rectifier K\(^+\)
conductance, 171, 175
of fast Na\(^+\) conductance, 172, 177
Hopf bifurcation, see stability,
network
Hubel-Wiesel model
complex cell, 80, see also complex
cell
simple cell, 79, see also simple
cell
hypercolumn, orientation, 252
hyperpolarization, 4, 160, 172, 198, 199
ICA (independent components
analysis), 384, see also
causal models
factor analysis
as limit of sparse coding model, 384
direct likelihood maximization, 384
EM inadequacy, 384
exact deterministic recognition, 384
factorial re-representation, 385
free energy, \( - F \), 384
information maximization, 385
learning rule, 395
natural gradient learning rule, 385
ideal observer, see
decision theory
estimation theory
signal detection theory
independence, linear, 402
independence, statistical, 416
independent-neuron code, 36, see also
neural coding
independent-spike code, 35, see also
neural coding
indirect actor, see actor, indirect
information, 125, 127, see also
entropy
noise entropy
continuous variable, 130
information maximization, 130, 385
limitations, 135
retinal ganglion cell, 135
information rate, 145
estimation by direct method,
inhomogeneous Poisson process, see Poisson process
input resistance, see dendrites, electrical properties
instrumental conditioning, 331, see also actor
classical conditioning
delayed rewards, problem of dynamic programming policy
exploration-exploitation dilemma, 341
foraging, 340, 342
maze task, 347
two-armed bandit, 341
water maze task, 352
integrate-and-fire models, 162
adaptation, 165
integration, 191
irregular firing mode, 189, 236
passive, 163
refractory period, 165
regular firing mode, 189
threshold potential, see threshold potential with constant current, 163
with synaptic input, 188
integrator, neural, 247, see also recurrent networks
eye position, 248
interneuron, 4
interspike intervals
coefficient of variation \( C_V \), 27, 33, 190
distribution, 27, 32
histogram, 32
intracellular electrical properties
current, longitudinal \( I_L \), 155, 204
resistance, longitudinal \( R_L \), 155
resistivity \( \rho_L \), 155, 203
invertible recognition models, see recognition models
ion channels, 4, 154, see also conductances
synaptic conductances
activation gate, 168
gating equation, 170, 196
kinetics, 169
models, 175
open probability \( P \), 168
probabilistic models, 176–177
selectivity, 154
single-channel conductance, 156
state diagrams, 175
state-dependent inactivation, 177
stochastic opening, 167
subunits, 168, 175
ion pumps, 4, 154, 158, 161, see also resting potential
IPSC (inhibitory postsynaptic current), see synaptic conductances
IPSP (inhibitory postsynaptic potential), see synaptic conductances
ISI, see interspike intervals
Jacobian matrix, 411
Jensen’s inequality, 150
K-means algorithm, 373, see also causal models
as limit of mixture of Gaussians, 374
kernel, see filter kernel
kernel, optimal, see spike decoding
Wiener kernel
Kirchhoff’s laws, 414
KL divergence, see Kullback-Leibler divergence
Kronecker delta, 400, see also \( \delta \) function
Kullback-Leibler divergence, 128, 150, see also entropy
information
‘flipped’ in Helmholtz machine, 388
and free energy, – \( f \), 370
density estimation, 323, 368
mean-field approximation and, 275, 277
kurtosis, 379, see also sparse distributions
Lagrange multipliers, 408, 408
latent variable, see causal models
lateral geniculate nucleus, see LGN
leakage current, 161
learning models, see
  reinforcement learning
  supervised learning
  unsupervised learning
learning rate, 286, 287, see also
delta rule
  Rescorla-Wagner rule
associability, 333
overshadowing, 335
learning rules, see
  anti-Hebb rule
  associative memory
  competitive Hebb rule
  contrastive Hebb rule
  correlation-based learning rule
covariance learning rule
delta rule
error-correcting learning rules
  Goodall rule
Hebb rule
  Oja rule
perceptron learning rule
  plasticity
Rescorla-Wagner rule
sliding threshold learning rule
temporal difference learning
rule
timing-based learning rule
trace learning
LGN (lateral geniculate nucleus), 45, 51
information theoretic
  characterization, 141
response properties, 77
thalamic relay neuron, 200
likelihood maximization, 323, 369, see also density estimation
likelihood ratio, 95, see also
decision theory
  and ROC curve, 95
likelihood ratio test, see
decision theory
  Neyman-Pearson lemma
limit cycle, see
  oscillations
recurrent networks
line attractor, see continuous attractor
linear algebra, 399
linear response estimate, 61, see also firing rate, estimation
linear separability, see perceptron
local cortical circuits, see recurrent networks
log likelihood, 323, 369
logistic function, 16, see also tuning curve, sigmoidal
LTD (long-term depression), 184, 282, see also
  learning rules
  plasticity
heterosynaptic depression, 288
homosynaptic depression, 288
in cerebellum, 310
timing-based, 291
LTP (long-term potentiation), 184, 282, see also
  learning rules
  plasticity
timing-based, 291
Lyapunov function, 260, 261, see also stability, network
  associative memory, 263
  Boltzmann machine, 275, 276
MAP (maximum a posteriori)
  estimation, see estimation theory
  marginal distribution, 363
  marginalization, 363
Markov chain, 274, see also
  Gibbs sampling
  stochastic networks
Markov decision problem, 350, 355, see also
  delayed rewards, problem of dynamic programming
Markov chain
  absorbing state, 355
Markov models
  of ion channels, see ion channels
  of synapses, see synaptic models
maximum likelihood
density estimation, see density
estimation
estimation, see estimation theory
maximum a posteriori
estimation, see estimation theory
maze task, 347, see also
delayed rewards, problem of
reinforcement learning
water maze task
reinforcement learning solution
of, 350
mean, of random variable, 416
mean-field distribution, 274, 372,
387, see also
Boltzmann machine
stochastic networks
firing rate model, 275
membrane current, 160, 160, see
e also conductances
injected, 161, 208
per unit area \( i_m \), 160, 205, 207
sign convention, 162
membrane potential, 4, 154
membrane structure, 153, see also
ion channels
membrane, electrical properties,
see also dendrites, electrical
properties
capacitance \( C_m \), 156, 204
capacitance, specific \( c_m \), 156, 204
conductance, 158
conductance, specific \( g_t \), 160
filtering, 233, 237
resistance \( R_m \), 157
resistance, specific \( r_m \), 158
time constant \( \tau_m \), 158, 163, 207, 208, 235
memory, associative, see
associative memory
Mexican hat connections, see
center-surround structure
mixture of Gaussians, 362, 373, see
e also
causal models
E phase, 365
generation, 373
generative distribution, 362
K-means algorithm limit, 374
learning rule, 395
M phase, 365
marginal distribution, 362
mixing proportions, 362
prior distribution, 362
recognition model, 373
responsibilities, 365
ML (maximum likelihood), see
maximum likelihood
moment-generating function, 41
Monte Carlo method, 324, 356, 357
morphoelectrotonic transform,
215, 216, see also
cable theory
compartmental models
motion coherence, 89
multiplication, see gain modulation
multiplicative normalization, 290,
see also
Oja rule
synaptic normalization
dynamical effect of, 296
multiresolution decomposition,
389
coding, 390
computational translation
invariance, 392
overcomplete representations,
392
representational
interdependence, 392
mutual information, see
information
myelin, 222
natural gradient, 385
natural scene statistics, 138, 142,
381
neocortex, 229, see also
cortical map
feedforward networks
recurrent networks
cortical column, 229
IT (inferotemporal), 378
M1, primary motor cortex, 15
MT (medial temporal area), 32,
89
posterior parietal cortex, 242
premotor cortex, see coordinate
transformation
V1, primary visual cortex, 14, 45,
51
Nernst equation, 159, see also equilibrium potential
Nernst potential, see equilibrium potential
neural coding, see also
  re-representation
correlation code, 35
factorial code, 134
independent-neuron code, 36
independent-spike code, 35
population code, 97
rate code, 38
redundancy, 134
temporal code, 37
neural decoding, see also
  estimation theory
causality constraint, 115, 117
optimal kernel, 116, 121
population vector, 99, 99–101
prediction delay $t_0$, 114
recurrent model, 258
stimulus estimate, 114
using spike times, see spike decoding
vector method, 99
neural recordings, 6, see also electrodes
extracellular, 7
intracellular, 6
voltage clamp, 171
neural response function $\rho(t)$, 9,
  see also firing rate
neuromodulator, 179, 201, see also dopaminergic activity
neuronal models, see
  compartmental models
  firing-rate models
  integrate-and-fire models
neurotransmitter, 5, 179, see also synapse
  synaptic conductances
GABA (γ-aminobutyric acid), 179
  glutamate, 179
Neyman-Pearson lemma, 95, 119,
  see also decision theory
NMDA receptor, see synaptic conductances
  nodes of Ranvier, 222, see also action potential, saltatory propagation
  noise, see also variability
    additive, 17
    multiplicative, 17
  neuronal, 17
  noise entropy, 126, see also entropy
    mutual information
    continuous variable, 130
  noise filter, 140
  norm, vector, 399
nullcline, see phase plane analysis
numerical methods, 191, 192, 225,
  see also differential equations
  Hines method, 227
  tridiagonal solution method, 227
Nyquist frequency, 59
ocular dominance, 294
  Hebbian development, 298
  subtractive normalization in, 299
ocular dominance stripes, 294, 302, 309
  competitive Hebb rule, 306
  feature-based developmental model, 309
  Hebbian development, 302
  relationship to orientation domains, 309
ocularity, see
  disparity
  ocular dominance
  ocular dominance stripes
odor analysis, see olfactory bulb
OFF responses, 53
Ohm’s law, 413
Oja rule, 290, 296, see also
  Hebb rule
  learning rules
  multiplicative normalization principal component stability, 291
olfactory bulb, 270
  excitatory-inhibitory network model of, 270
ON responses, 53
Index

operant conditioning, see instrumental conditioning
optical modulation transfer function, 138
optimal control, see dynamic programming
optimal kernel, see spike decoding orientation
preferred \( \theta \), 15, 65
orientation domains, 309
feature-based developmental model, 309
linear zones, 309
pinwheels, 309
relationship to ocular dominance stripes, 309
orientation selectivity, see complex cell
orientation domains
simple cell
orthodromic propagation, 221, see also action potential
oscillations, 36, see also
phase plane analysis
recurrent networks
stability, network
amplification, selective, 272
frequency, 268
limit cycle, 269, 410
olfactory bulb, 270, 272
phase-locked, 270
reciprocal inhibition, 188
overcomplete, see basis functions
multiresolution decomposition

Parseval’s theorem, 407
passive cable models, see cable equation
cable theory
morphoelectrotonic transform
Pavlovian conditioning, see classical conditioning
PCA (principal components analysis), 297, 375, see also
causal models, see also principal component
as limit of factor analysis, 375
computational properties, 297, 377
EM, 376
free energy, \(-F\), 375
generation, 375
learning rule, 395
recognition model, 375
vs. factor analysis, 376–377
perceptron, 314, see also supervised learning
capacity, 316
Hebb rule, 315
linear separability, 315
perceptron convergence theorem, 327
perceptron learning rule, 319, 324, 327, see also
contrastive Hebb rule
delta rule
learning rules
periodogram, 41, see also power spectrum
phase plane analysis, 266, 266, see also
oscillations
stability, network
nullcline, 266
phase-locking, see oscillations
place cells, hippocampal, 36
plasticity via a timing based learning rule, 313
water maze, 352
plasticity, 284, see also
LTD, LTP
learning rules
covariance rule, 264
discrete update, 287
multiple timecourses of, 281
non-Hebbian plasticity, 283
short-term, 184, 184
depression, 184
facilitation, 184
Poisson input, 185–187
point attractor, 261, 267, 410, 412, see also stability, network
basin of attraction, 261
point process, 25
Poisson distribution, 26, 417
mean and variance, 41
Poisson process, 25
computer simulation, 30
entropy rate, 146
homogeneous, 25, 41
inhomogeneous, 29, 42
Poisson spike train
autocorrelation, 28
interspike interval distribution, 27
policy, 340, see also actor
stochastic, 341, 347
policy iteration, 347, 356, see also
actor-critic algorithm
dynamic programming
policy evaluation, 348, 348, 356, see also critic
policy improvement, 349, 357, see also actor
population code, 97, see also neural
coding
population vector, 99, 99–101, see also neural decoding
post-inhibitory rebound, 198, 200
potassium conductances, see
conductances, K+
potentiation, long-term, see LTP
potentiation, synaptic, see LTD, LTP
plasticity
power spectral density, see power
spectrum
power spectrum, 22, 40
prediction delay to, see neural
decoding
prediction error
for delayed reward, see temporal
difference prediction error
for immediate reward, 333
principal component, 294, 296, see also PCA
ocular dominance stripes, 304
prior probability, see probability
probability
conditional, 87
joint, 88
prior, 88
probability density, 24, 417, see also
density estimation
probability equalization, 134
probability measure, 416
probability theory, 416
sample space, 416
projective field, 382, see also causal
models
pyramidal cell, 4, see also neocortex
compartmental reduction, 218,
see also compartmental models
long term plasticity in, 291
morphoelectrotonic transform,
217, see also morphoelectrotonic
transform
short term plasticity in, 184
quadrature pair, 76, see also
complex cell
energy model
Rall model, see compartmental
models
random variable, 416
continuous, 417
covariance, 416
independence, 416
mean, 416
variance, 416
rate code, see
firing-rate models
neural coding
re-representation, 359, 359, see also
causal models
information theory
as recognition, 360
factorial, 363
hierarchical, 382, 393
interdependence, 392
lossy vs. lossless, 391
low-dimensional, 363
overcomplete, 392
sparse, 363
rebound, see postinhibitory
rebound
receiver operating characteristic,
92, 92, see also decision theory
signal detection theory
receptive field, 14, 53, see also
center-surround structure
tuning curve
center $x_0$, $y_0$, 65
difference-of-Gaussians model, 77
Index

nonseparable, 61, 71
separable, 61
size $\sigma_x, \sigma_y$, 58, 62
space-time, 61, 68
recognition models, 360, see also
causal models
EM
generative models
approximate distribution $Q$, 367, 369
approximate, using $J$, 372
as expectation phase of EM, 365
as inverse to generative models, 363
deterministic, 360
invertible, 367, 370
noninvertible, 367
noninvertible deterministic
models, 371
noninvertible probabilistic
models, 372
probabilistic, 360
recognition distribution
$P[v|u; G]$, 363
variational method, 372
recording, see neural recordings
rectification, see
half-wave rectification
threshold function
recurrent networks, 238, 244, 301
competitive dynamics, 255, 305
complex cell model, 254, see also
complex cell
continuous labeling, 248
excitatory-inhibitory networks, 265
fixed point, see
continuous attractor
point attractor
gain modulation, 256
limit cycle, 269, see also
oscillations
linear, 245
ML inference, 258, see also
estimation theory
nonlinear, 250
olfactory bulb model, 270
oscillatory, 268
simple cell model, 252, see also
simple cell
stability, see stability, network
steady state $v_\infty$, 246
sustained activity, 257, see also
integrator, neural
symmetric coupling, 239
refractory period, 4, 33, 221, see also
action potential
integrate-and-fire models
regression, see function
approximation
reinforcement learning, 283, 331,
331, see also
actor-critic algorithm
classical conditioning
dynamic programming
instrumental conditioning
temporal difference learning
asynchronous and model free,
356, 357
exploration-exploitation
dilemma, 341
subjective utility, 343
relay neuron, see thalamic relay
neuron
renewal process, 25
Rescorla-Wagner rule, 332, 333, see
also
delta rule
learning rules
temporal difference learning rule
and gradient descent, 333
as delta rule, 333
blocking, 334
indirect actor, use for, 342
inhibitory conditioning, 335
multiple stimuli, 334
secondary conditioning
difficulties, 336
response function, see neural
response function
response variability, see variability,
in spike count
resting potential, 4, 161, 207
retina, 51
retinal circuitry, 52
retinal coordinate system, see also
complex logarithmic map
azimuth $a$, 55
eccentricity $e$, 55
retinal disparity, see disparity
retinal ganglion cells, 52
information theoretic
characterization, 135
receptive fields, 77
reversal potential, 159, see also
conductances
equilibrium potential
reverse correlation, 20, 47, see also
spike decoding
spike-triggered average
Wiener kernel
complex cells, 74
simple cells, 60
reverse Euler method, 226, see also
numerical methods
ROC, see receiver operating
characteristic
saddle-node bifurcation, see
stability, network
saltatory propagation, see action
potential
sample space, 416
sampling theorem, 408, see also
Nyquist frequency
score $Z(r)$, 97, see also decision
theory
estimation theory
second messenger, 179
secondary conditioning, 336, see also
classical conditioning
delayed rewards, problem of
selective amplification, see
amplification, selective
self-organizing map, see
feature-based models
self-supervised learning, see
unsupervised learning
shunting, see conductances,
shunting
sigmoidal function, 50, see also
tuning curves, sigmoidal
signal detection theory, 90–97, see also
decision theory
signal-to-noise ratio, 141, 144
simple cell, 54, 73, see also
Hubel-Wiesel model
orientation selectivity
feedforward model, 79
Hebbian development, 299
ICA development of, 386
recurrent model, 252
sparse coding model
development, 381
single channel, see ion channels
sleep phase, see
Boltzmann machine
Helmholtz machine
sliding threshold learning rule,
288, see also
covariance learning rule
Hebb rule
learning rules
projection pursuit, 328
stability, 289
synaptic competition, 289
sliding window, 13
sodium conductances, see
conductances, Na$^+$
soma, 4
sparse coding model, 378, see also
causal models
factor analysis
approximate deterministic
recognition, 380–381
dynamical recognition, 380
EM, 381
factorial re-representation, 379
free energy, $-F$, 380
generation, 379
ICA limit, 384
learning rule, 395
multiresolution decomposition,
389, see also multiresolution
decomposition
sparseness prior, 380
synaptic normalization, 381
sparse distributions, 378
kurtosis, 379
sparse representation, see
re-representation
sparseness, 262, see also kurtosis
spatial frequency $K$, 58
preferred spatial frequency $k$, 62
spatial phase $\Phi$, 58
invariance, 74
preferred spatial phase $\phi$, 62
spike, see action potential
spike count, 9
distribution, 31–32, see also Fano factor
spike decoding, 113, see also estimation theory
firing rate
Wiener kernel
optimal kernel, 47, 81
spike train, 8, see also interspike intervals
spike count
computer simulation, 51, see also Poisson process, computer simulation
spike-count rate $r$, 9, see also firing rate
spike-rate adaptation, 165, 201, see also integrate-and-fire models
spike-triggered average $C(r)$, 19, 19, 47, 60, see also firing rate, estimation
spike decoding
multiple-spike triggers, 23
other moments, 23
spiking probability, 10
spine, dendritic, 6, see also synapse
stability, network, 260, see also differential equation
bifurcation
Hopf, 269
saddle-node, 269
continuous attractor, 247, 251, see also continuous attractor
fixed point, 260, 261, 267, 412, see also point attractor
linear instability, 246
Lyapunov function, 260, see also Lyapunov function
marginal stability, 412
nonlinear instability, 252
oscillations, 268, see also oscillations
phase plane analysis, 266, see also phase plane analysis
point attractor, 260, 261, 267, see also point attractor
stability, network
also point attractor
stability matrix, 268, 271
strange attractor, 410
stationary state, see continuous attractor
point attractor
recurrent networks
stability, network
steady state, see continuous attractor
point attractor
recurrent networks
stability, network
stimulus
flashed or moving vs. static, 69–70
Gaussian white noise, 23
maximally effective, 48, 83
periodic, 19
white-noise, 22, 40, 47
white-noise image, 58
stimulus reconstruction, see spike decoding
stochastic gradient ascent, see gradient ascent
stochastic gradient descent, see gradient descent
stochastic networks
Boltzmann machine, 273, 322, see also Boltzmann machine
Helmholtz machine, 388, see also Helmholtz machine
Markov chain Monte Carlo sampling, 274
probabilistic input-output relationship, 322
wake-sleep algorithm, 389
stomachogastrial ganglion, 201, 201
STP (short-term potentiation), see plasticity
strange attractor, 410
striatum, 351
subtractive normalization, 290, see also synaptiastic normalization
dynamical effect of, 296, 299
in Hebb rule, 296
ocular dominance development, 299
supervised learning, 283, 313, see
also density estimation
surface attractor, see continuous attractor
synapse, 5, see also gap junction
in integrate-and-fire models, 188, see also integrate-and-fire models
inhibitory and excitatory, 160
synaptic competition, 284, see also
plasticity
synaptic normalization
sliding threshold learning rule
obviatio of, 289
timing-based learning rule, 293
synaptic conductances, 167, 178, see also
conductances
ion channels
AMPA, 179, 181
GABA_A, 180, 182
GABA_B, 180
ionotropic, 179
metabotropic, 179
NMDA, 179, 182, 183
synaptic current I_s, 232, 232, see also firing-rate models
synaptic depression, see LTD
plasticity
synaptic facilitation, see LTD
plasticity
synaptic kernel, 233, see also filter
firing-rate models
synaptic models
α function, 182
difference of exponentials, 182
exponential, 181
probabilistic, 180
transmitter release, 180
synaptic normalization, 289, 381, see also
multiplicative normalization
plasticity
subtractive normalization
dynamic imposition, 289, see also Öja rule
in supervised learning, 314
of square norm, 289
of sum, 289
rigid imposition, 289
synaptic open probability P_o, 179
synaptic plasticity, see plasticity
synaptic potentiation, see LTP
plasticity
synaptic receptors, see synaptic conductances
synaptic saturation, 284, see also plasticity
dynamical effect of, 295
synaptic time constant τ_o, 234
synaptic transmission, 178
synaptic weights, 233, 285, see also firing-rate models
plasticity
associative memory, 263–264, see also associative memory
synchrony, 36, 188, see also oscillations
systems identification, see reverse correlation
Taylor series, 411
TD (temporal difference), see entries under temporal difference
temporal code, see neural coding
temporal difference error, see temporal difference prediction error
temporal difference learning rule, 336, 337, 356, see also
actor-critic algorithm
delayed rewards, problem of learning rules
and delta rule, 337
critic, 348
recursive formula for summed reward, 337
secondary conditioning, 338
stimulus traces, 352
TD(λ) rule, 352
temporal difference prediction error δ(t), 337
direct actor, use for, 349
discounted, 352
model of dopaminergic activity, see decision theory

unbiased estimator, see estimation theory

unsupervised learning, 283, 293, 359, see also
causal models
density estimation
EM
causal models, 360, 363
input whitening, 381

variability, see also noise in
ISI, 27, 189–191, see also
interspike intervals, coefficient of variation
spike count, 16, see also Fano factor
spike times, 34

variance (of estimator), see estimation theory

variance (of random variable), 416

variance equalization, 135

variational method, 372

vector derivative $\nabla$, 402

vector method, see neural decoding

velocity (of grating)
preferred velocity, 72

ventral tegmental area, see dopaminergic activity

voltage attenuation, see cable equation

voltage clamp, see neural recordings

Volterra expansion, 46, 51

wake phase, see

Boltzmann machine

Helmholtz machine

wake-sleep algorithm, 389, see also

Helmholtz machine

water maze task, 352, see also
delayed rewards, problem of,
see also maze task

reinforcement learning solution, 354

Weber’s law, 18

white noise, see stimulus,

white-noise

<table>
<thead>
<tr>
<th>Index</th>
<th>459</th>
</tr>
</thead>
<tbody>
<tr>
<td>model of dopaminergic activity,</td>
<td>see decision theory</td>
</tr>
<tr>
<td>test power, see decision theory</td>
<td></td>
</tr>
<tr>
<td>test size, see decision theory</td>
<td></td>
</tr>
<tr>
<td>thalamic relay neuron, 200</td>
<td></td>
</tr>
<tr>
<td>thermal energy, 154</td>
<td></td>
</tr>
<tr>
<td>threshold function, 50, 234, see also</td>
<td></td>
</tr>
<tr>
<td>half-wave rectification</td>
<td></td>
</tr>
<tr>
<td>threshold potential, 162, see also</td>
<td></td>
</tr>
<tr>
<td>integrate-and-fire models</td>
<td></td>
</tr>
<tr>
<td>tight frame, 318, see also</td>
<td></td>
</tr>
<tr>
<td>basis functions</td>
<td></td>
</tr>
<tr>
<td>function approximation</td>
<td></td>
</tr>
<tr>
<td>timing-based learning rule, 291, see also</td>
<td></td>
</tr>
<tr>
<td>Hebb rule</td>
<td></td>
</tr>
<tr>
<td>learning rules</td>
<td></td>
</tr>
<tr>
<td>plasticity</td>
<td></td>
</tr>
<tr>
<td>causality, 291</td>
<td></td>
</tr>
<tr>
<td>Hebbian, 292</td>
<td></td>
</tr>
<tr>
<td>prediction, 311</td>
<td></td>
</tr>
<tr>
<td>synaptic competition in, 293</td>
<td></td>
</tr>
<tr>
<td>temporal asymmetry, 291</td>
<td></td>
</tr>
<tr>
<td>trace learning, see trace learning</td>
<td></td>
</tr>
<tr>
<td>Töplitz matrix, see translation invariance</td>
<td></td>
</tr>
<tr>
<td>trace learning, 300, 301</td>
<td></td>
</tr>
<tr>
<td>translation invariance, 136, 240, 304, 404, see also eigensystem</td>
<td></td>
</tr>
<tr>
<td>computational, 392</td>
<td></td>
</tr>
<tr>
<td>Töplitz matrix, 401</td>
<td></td>
</tr>
<tr>
<td>transmitter release probability $P_{rel}$, 179, 184, 185</td>
<td></td>
</tr>
<tr>
<td>transmitter-gated channels, see synaptic conductances</td>
<td></td>
</tr>
<tr>
<td>trial average ($\langle \rangle$), 10</td>
<td></td>
</tr>
<tr>
<td>trial average rate ($r$), see firing rate tuning curve, 14, see also</td>
<td></td>
</tr>
<tr>
<td>feature-based models</td>
<td></td>
</tr>
<tr>
<td>neural coding</td>
<td></td>
</tr>
<tr>
<td>receptive field</td>
<td></td>
</tr>
<tr>
<td>body-based, 242, 244</td>
<td></td>
</tr>
<tr>
<td>cosine, 15, 98</td>
<td></td>
</tr>
<tr>
<td>dynamic extension, 51</td>
<td></td>
</tr>
<tr>
<td>Gaussian, 14</td>
<td></td>
</tr>
<tr>
<td>invariance, 254</td>
<td></td>
</tr>
<tr>
<td>optimal width, 110–112</td>
<td></td>
</tr>
<tr>
<td>sigmoidal, 16</td>
<td></td>
</tr>
<tr>
<td>two-alternative forced choice test,</td>
<td></td>
</tr>
<tr>
<td>see decision theory</td>
<td></td>
</tr>
</tbody>
</table>
white-noise kernel, 47
whitening filter, 137, 138, see also
decorrelation
whitening, of input for
unsupervised learning, 381
Wiener expansion, 46, 51
Wiener kernel, 46, see also spike
decoding, optimal kernel
Wiener-Hopf filter, 117
window function, 13
winner-takes-all, see recurrent
networks, competitive
dynamics