

Index

- 3DVAR, 125, 126, 131, 153, 155, 158
 - accuracy, 127, 128
 - algorithm, 127
 - implementation, 130
 - Kalman filter comparison, 126
 - Kalman gain, 127
 - vs. 4DVAR, 129
- 4DVAR, 125, 131
 - implementation, 130
 - strong constraint, 129
 - vs. 3DVAR, 129
 - weak constraint, 129, 130
 - small signal noise limit, 131
 - vanishing noise limit, 130
- analysis operator, 145
- analysis step, 113, 139, 144
- approximation of an integral, 63
- Armijo condition, 175
- auto-correlation, 77, 90
- Bayes' formula, xiv, 5, 20, 144
- Bayes' theorem, 4, 6, 7, 33, 60, 73
 - as optimization, 56
 - variational formulation, 56–58
- Bayesian brittleness, 20
- Bayesian inference, 20, 25, 49
 - optimization approach, 49
- Bayesian inversion, 31, 188
 - Bayesian hierarchical formulation, 34
 - function space, 30
 - optimization approach, 32
- Bayesian model evaluation, 20
- Bayesian posterior distribution, 125
- Bayesian statistics, 20
- Bernstein–von Mises theorem, 31
- BPF, 144, 146, 147, 149, 150, 153, 163
- consistency, 147
- convergence, 147, 156
- linear-Gaussian, 162
- random dynamical system, 151
 - vs. OPF, 153, 155, 162
- burn-in, 78
- coupling, 89
- credible intervals, 6
- curse of dimensionality, 163
- data assimilation, xv, 6, 9, 30, 57, 61, 109, 173
 - machine learning, 111
 - sequential, 189
 - stability, 13
 - time-ordered data, 101
 - variational methods, 132
- data model, 102, 103, 110
 - deep learning, 124
 - linear, 112
- data sequence, Y , 102
- data-misfit objective function, 173, 174, 177, 180, 185
- detailed balance, 80, 86
- Dirac mass, xvi, 6, 24, 60, 61, 73, 153, 160
- Dirac mass approximation, 6, 144
- discrete Gronwall inequality, 17, 18, 129, 150
- distance, 3
 - between random probability measures, 59, 61, 147, 156
- Hellinger, 3, 9, 11, 12, 50, 51, 66, 91, 108
 - bounds, 10
 - compared with total variation, 12
 - smallness, 12
 - upper bound, 51
- total variation, 3, 9, 11, 50, 59, 66, 108
 - bounds, 10

smallness, 12
 upper bound, 51
 variational characterization, 11, 85
 divergence
 χ^2 , 59, 65–67, 69, 162
 asymmetry, 66
 bootstrap and optimal filter, 162
 Kullback–Leibler, 19, 49–52, 56–58,
 66, 91
 asymmetry, 50
 mean-seeking, 55
 mode-seeking, 55
 dynamical system, xv, 110
 deterministic chaotic, 132
 discrete, 126
 high dimension, 137
 random, 158, 163
 stochastic, xv
 dynamics
 deterministic, 110
 linear, 113
 machine learning, 124
 noise-free, 117
 nonlinear, 126
 stochastic nonlinear, 153
 EM, *see* expectation maximization
 EnKF, 133, 137, 138, 142
 algorithm, 138
 geophysical applications, 142
 optimization, 142
 perturbed observation, 139, 142
 EnKF implementation
 N dimensions, 142
 EnKI, 191
 EnKI-SL, 177, 185, 186
 ensemble Kalman filter
 iterative
 statistical linearization, *see* IEnKF-SL
 ensemble Kalman method, 183, 189
 linear case, 142
 machine learning, 189
 ensemble Kalman methods
 nonlinear case, 143
 ergodicity, 78, 83
 coupling approach
 continuous state-space, 86
 finite state-space, 83
 geometric, 78
 pCN method, 88
 ergodicity of sample path, 76
 ExKF, 133–135, 142

algorithm, 135, 136
 high dimension, 137
 weather forecasting, 142
 ExKI, 177, 180, 181, 186
 linear, 182
 vs. IExKF, 181
 expectation maximization (EM), 167, 168
 exponential family, 58
 filtering
 Bayesian formulation, 9
 on-line, 101, 125
 filtering distribution, 102, 120, 133, 139, 144,
 153
 approximation, 109
 Gaussianity, 109, 113
 Kalman smoother, 120
 linear-Gaussian, 112
 non-Gaussian setting, 144
 optimization methods, 125
 smooth distribution marginal, 102
 smoothing marginal, 108
 true, 147
 filtering map, 146
 filtering problem, 102, 106, 109
 analysis step, 106, 107, 109
 updating, 108
 distribution, 106
 linear, 112
 prediction step, 106, 107, 109
 well-posedness, 108
 filtering update, 153
 filtering update formula
 approximate, 146
 forecasting, xv
 forward
 error, 8
 model, xiii, xiv, 3, 8, 13, 19, 21, 30
 approximate, 15
 approximation, xiii
 linearity, 21
 nonlinear, 125
 problem, xiii
 forward Euler method, 15, 16
 Gauss–Newton method, 173, 175, 177, 178,
 184, 185, 191
 implicit regularization, 177
 MAP estimator, 124
 statistical linearization, 183
 Tikhonov–Phillips objective function, 177
 Gaussian, 6, 112, 117
 isotropic, 52

- mixture, 6
- noise, 21
- Gaussian ansatz, 144
- Gaussian approximation, 6, 7, 21, 56, 57, 109
 - best, 51–54, 56
 - moment matching, 54
 - infinite-dimensional, 58
 - mixtures, 58
 - moment-matching, 57
 - small noise limit, 58
- Gaussian best fit, *see* Gaussian approximation
- Gaussian elimination, 121
- Gaussian noise, 19
- Gaussian posterior, 22
- Gaussian prior, 21
 - truncated, 86
- Gibbs sampling, 94
 - Metropolis–Hastings, 95
- GOPE, 156, 163
 - algorithm, 159
 - convergence, 156
- gradient descent, 40
 - machine learning, 46
 - stochastic, 43, 44, 46, 53
 - convergence, constant step-size, 47
 - convergence, 44
 - machine learning, 45
- gradient flow
 - preconditional, 40
- graphical model, 95
- IEnKF-SL, 177, 184, 185, 190
- IExKF, 177–179, 185, 189
 - linear, 182
 - linear case, 179
 - vs. ExKI, 181
- ill-posed model, xiv
- implicit regularization, 177
- importance resampling
 - sequential, 71
- importance sampling, xiv, 30, 59–61, 65, 69, 71, 73, 144, 147, 156
 - accuracy, 69
 - adaptive algorithms, 71
 - algorithm, 64
 - error, 67
 - estimator, 69
 - accuracy, 71
 - filtering, 71
 - sample size, 71
- independent samples, 73
- information loss, 55
- innovation, 116
- invariant distribution, 73, 80
 - Markov kernel, 74
- inverse error, 8
- inverse problem, xiii, 21
 - Bayesian formulation, xiv, 3, 6, 9, 18–20, 23, 49, 57, 59, 86, 103, 110, 174
 - algorithmic approach, 19
 - infinite-dimensional, 48
 - stability, 13, 19
 - well-posedness, 3, 8, 19, 105
 - dimension reduction, 19
 - ensemble Kalman methods, 173
 - function space, 19
 - ideal, xiv
 - ill-posedness, xiv, 177, 191
 - incomplete data, xiv
 - linear, 30
 - linear-Gaussian setting, 30, 115, 162, 178, 189
 - bound, 69
 - non-statistical approach, 18
 - nonlinear, 30
 - optimization approach, 24, 31, 47, 173
 - PDE approach, 48
 - regularization, 48
 - statistical estimation theory, 48
 - underdetermined, 31
 - well-posedness, 101
 - well-posedness under data perturbation, 91
- Ising model, 95
- Jensen inequality, 66
- Kalman filter, 109, 112, 119, 126, 136, 139
 - algorithm, 116
 - analysis step, 114
 - characterization, 113
 - constraint formulation, 139
 - EnKF, comparison, 137, 138
 - ensemble, *see* EnKF, 163, 173
 - ExKF, comparison, 134
 - extended, *see* ExKF, 173
 - iterative ensemble, *see* IEnKF-SL
 - iterative extended, *see* IExKF
 - linear-Gaussian setting, 139, 144
 - mean, 117
 - optimality, 119, 124
 - optimization approach, 126
 - prediction step, 114
 - unscented, 189
 - update formulae, 134
 - updates, 115

- Kalman gain, 116, 117, 127, 178
 - EnKF, 139
 - ExKF, 136
- Kalman inversion
 - ensemble, *see* EnKI-SL
 - statistical linearization, *see* EnKI-SL
 - extended, *see* ExKI
 - Tikhonov ensemble, *see* TEnKI-SL
 - statistical linearization, *see* TEnKI-SL
 - Tikhonov extended, *see* TExKI
- Kalman smoother, 112, 119, 124
 - characterization, 120
 - marginal, 120
 - mean, 121
- Kalman–Bucy filter, 124
- Kullback–Leibler divergence, *see* divergence
- Laplace distribution, 96
- Levenberg–Marquardt method, 173, 175–177, 185, 187, 191
 - data-misfit objective function, 177, 180
 - implicit regularization, 176
 - increment, 176
 - statistical linearization, 183
 - Tikhonov–Phillips objective function, 177, 181
- likelihood function, xiv, 4, 5, 9, 13, 23, 73, 166
 - approximate, 13
 - Gaussian, 22, 30
 - linear-Gaussian setting, 169
 - randomized, 139, 142
- likelihood map, \mathcal{A}_j , 107
- linear-Gaussian setting, 21, 23, 114, 124, 139
- Lipschitz function, 15
- log-likelihood, 33
 - pseudo, 97
- loss regularizer
 - Gaussian prior, 34
 - Laplace prior, 34
- loss term, 23, 25, 33, 52, 104, 105
 - cross-entropy, 188
 - Gaussian observational noise, 33
 - logistic, 188
- machine learning, 39
 - probabilistic, 57
- MAP estimator, 6, 7, 19, 32, 33, 36, 48, 121, 125, 174
 - attainable, 34
 - fails to capture posterior, 39
 - high dimension, 38
 - infinite-dimensional inverse problems, 48
 - linear-Gaussian, 23, 33
 - non-Gaussian prior, 48
 - optimization approach, 34
 - posterior distribution
 - single-peaked, 36
- Markov chain, 82
 - ergodicity, 73
 - continuous state-space, 86, 88
 - finite state-space, 82
 - homogeneous, 47
 - inhomogeneous, 44, 47
 - n -dimensional, 74
 - sample path, 74
 - sampling, 73, 75, 76
- Markov chain Monte Carlo, *see* MCMC
- Markov inequality, 26
- Markov kernel, 73, 74, 83, 145
 - detailed balance, 80
- MCMC, xiv, 73, 75, 78, 89
 - error, 76
 - sampling, 90
 - mean-field approximation, 57
- Meta Theorem, 8
 - failure, 19
- Metropolis–Hastings, 73, 74, 78–80, 82, 94, 96
 - acceptance probability, 78
 - detailed balance, 81
 - ergodicity, 82
 - estimator, 79
 - Gibbs sampling, 95
 - kernel, 79, 86
 - detailed balance, 80, 81
 - invariance, 82
 - pCN, 78, 82, 86
 - proposal kernel, 78
 - sampling, 73
 - target distribution, 80
- model error, 110, 111
- model parameters, 111
- Monte Carlo algorithm, 55
- Monte Carlo error, 62
- Monte Carlo estimator, 61–64
 - accuracy, 78
- Monte Carlo method, 71
 - multi-level (MCLC), 71
 - quasi-, 71
- Monte Carlo sampling, 59–63, 67, 69, 73, 144, 156
 - algorithm, 61
 - limitation, 63
- nonlinear least-squares, 174, 180
 - objective function, 173, 174

- objective function, 23, 33
 - canonical, 34
 - data-misfit, 173, 174, 177, 185
 - Gaussian, 36
 - general least-squares, 176
 - Laplace, 36
 - nonlinear least-squares, 173, 174
 - optimization
 - small ball, 36
 - posterior probability, 35
 - regularized, 176
 - rough, 37
 - sparsity, 34
 - stochastic, 43
 - Tikhonov–Phillips, 173, 174, 177, 187
 - observation distribution, 6, 7, 105
 - bound, 149
 - linear, 113, 126, 144, 153, 155
 - noise, 3, 7, 21
 - nonlinear, 153
 - small noise limit, 25, 28, 30, 162, 163
 - zero noise limit, 30
 - observation model
 - linearity, 109
 - OPF, 153, 155, 156, 158, 163
 - algorithm, 159
 - convergence, 156, 163
 - implementation
 - linear setting, 157
 - nonlinear setting, 157
 - linear-Gaussian, 162
 - optimality, 160, 161
 - vs. BPF, 153, 155, 162
 - optimal particle filter, *see* OPF
 - optimization, 6, 40, 57
 - algorithm, 40
 - gradient-based, 39, 174
 - large-scale machine learning, 48
 - optimization algorithms compared, 177
 - parameter estimation on ODEs
 - well-posedness, 15
 - parameter tuning, 90
 - particle filter, 144, 163
 - algorithms, 152
 - bootstrap, *see* BPF
 - Gaussianized optimal, *see* GOPF
 - local, 163
 - optimal, *see* OPF
 - without resampling, 160, 161
 - pCN, 74, 78, 82, 86, 88, 89
 - continuous state-space, 86
 - pendulum problem, 165
 - perturbed observation, 139
 - posterior distribution, xiv, 3–7, 15, 21, 22, 36, 52
 - approximate, 6, 15, 58
 - consistency
 - determined case, 27
 - failure, 31
 - infinite dimensions, 30
 - overdetermined, 25
 - underdetermined case, 28
 - covariance, 4, 22, 23, 25
 - Dirac mass, 25
 - expectation, 20
 - Gaussian, 22, 30, 36
 - high dimension, 5
 - Laplace, 36
 - linear-Gaussian setting, 24
 - mean estimator, 5–7, 22, 24, 25, 121
 - linear-Gaussian, 23
 - measure, 7, 8
 - mode, 5, 32
 - multiple-peaked, 36
 - on a manifold, 7
 - rough, 37
 - sensitivity, 8
 - single-peaked, 36
 - small noise limit
 - determined case, 27
 - overdetermined case, 24
 - underdetermined case, 28
 - well-posedness, 13
- precision, 115
- preconditioned Crank–Nicolson, *see* pCN
- prediction map, 106
 - ExKF, 135
- prediction step, 113
- prior distribution, 4, 5, 7, 20, 21, 52
 - conjugate, 30
 - covariance, 23
 - Gaussian, 21, 30
 - measure, 20
 - non-Gaussian, 19
- proposal distribution, 63
- proximity of target and proposal distribution, 69, 71
- quasi-random sequence, 71
- r -Polyak–Lojasiewicz (r -PL) condition, 92
- r -strongly convex, 92
- random walk Metropolis algorithm, 94
- regularization, 23, 33, 174

210

Bayesian, 23
 Bayesian approach, 19
 relative entropy, 50
 resampling, 151, 160
 sampling, 6, 7, 89
 algorithm, 58
 approximate, 153
 high-dimensional state-space, 89
 sequential importance resampling, 144, 146
 sequential Monte Carlo, 188
 signal estimation, 125
 signal sequence, *V*, 102
 small noise limit, 8, 21, 24, 30, 133
 linear-Gaussian setting, 24
 observation, 162
 smoothing distribution, 102–104, 120
 Gaussian, 120
 linear-Gaussian, 112
 linear-Gaussian setting, 121
 marginal, 102, 108
 optimization methods, 125
 smoothing problem, 102, 109, 142
 Bayesian formulation, 9
 inverse problem formulation, 103
 likelihood function, 106
 linear, 112
 linear-Gaussian setting, 119
 off-line, 101, 125
 stability
 data perturbation, 13
 well-posedness, 104, 105, 108
 state-space
 finite, 82
 ergodicity, 83

Index

state-space estimation
 high-dimensional, 139
 statistical linearization, 183–186,
 189
 stochastic dynamics model, 101, 103,
 107, 110
 linear, 109, 112
 noiseless, 162
 nonlinear, 109
 target distribution, 57, 59, 61, 79
 invariance, 80, 81
 TEnKI, 191
 TEnKI-SL, 177, 186, 187, 191
 TExKI, 177, 181, 182, 187
 linear, 182
 Tikhonov–Phillips objective function, 173, 174,
 177, 178, 180, 184, 187
 trust region, 176
 uncertainty quantification
 forward, 20
 variational characterization of total variation,
 11
 w4DVAR, *see* 4DVAR, weak constraint
 weak convergence, xvi, 50
 of Gaussian distribution, 24, 25
 of probability measure, \Rightarrow , 24
 weight collapse, 152
 well-posed model, xiv
 well-posedness
 filtering, 9
 inverse problem, 8, 9
 smoothing, 9
 Woodbury matrix identity, 178