

Contact symmetries and Hamiltonian thermodynamics [31]
 Geometrical structures of higher-order dynamical systems and field theories [157]
 Free energy and phase transition of the matrix model on a plane wave [86]
 Integrals of nonlinear equations of evolution and solitary waves [119]
 A simple observer for nonlinear systems applications to bioreactors [76]
 HIERARCHICAL CLASSIFICATION TREE MODELING OF NONSTATIONARY NOISE FOR ROBUST SPEECH
 RECOGNITION [211]
 Lasso-Kalman smoother for tracking sparse signals [9]
 Kalman filtering for compressed sensing [108]
 Optimization Viewpoint on Kalman Smoothing with Applications to Robust and Sparse Estimation [12]
 Real-time dynamic MR [158]
 A finite dimensional filter with exponential conditional density [32]
 Exact and Approximate Hidden Markov Chain Filters Based on Discrete Observations [29]
 Filtering, stability, and robustness [196]
 The wonham filter with random parameters rate of convergence and error bounds [85]
 Wonham filter with random parameters Rate of convergence and error bounds, 2003 [84]
 Asymptotic Stability of the Wonham Filter Ergodic and Nonergodic Signals [20]
 Dynamic topic models [25]
 A Recurrent Latent Variable Model for Supervised Modeling of High-Dimensional Sequential Data [50]
 Why Gaussian macro-finance term structure models are nearly unconstrained factor-VARs [106]
 Estimating parameters and hidden variables in non-linear state-space models based on ODEs for biological networks
 inference [159]
 Dynamic Factor Models [4]
 An Introduction to factor graphs [132]
 Estimation of multi-factor shadow-rate term structure models [111]
 Belief networks, hidden Markov models, and Markov random fields A [179]
 Hidden physics models Machine learning of nonlinear partial differential equations [162]
 Comparing Coefficients of Nested Nonlinear Probability Models [112]
 Hierarchically Nested Covariance Structure Models for Multitrait-Multimethod Data [205]
 The nested chinese restaurant process and bayesian nonparametric inference of topic hierarchies [24]
 Hierarchical Mixture Models for Nested Data Structures [199]
 A Rasch Hierarchical Measurement Model [134]
 A Forward-Backward Algorithm for Nested Hidden semi-Markov Model and Application to Network Traffic [208]
 Multi-task learning for sequential data via iHMMs and the nested Dirichlet process [148]
 Sharing Clusters among Related Groups Hierarchical Dirichlet Processes [188]
 State space techniques in structural equation modeling [140]
 An Investigation of Growth Mixture Models for Studying the Flynn Effect [142]
 A Nested Mixture Model for Genomic Prediction Using Whole-Genome SNP [212]
 A nested mixture model for protein identification using mass spectrometry [126]
 Bootstrap non-nested mixture model selection-Application to extreme value modeling in metal fatigue problems [197]
 Nested Gibbs sampling for mixture-of-mixture model and its application to speaker clustering [187]
 Bayesian Multilevel Latent Class Models for the Multiple Imputation of Nested Categorical Data [201]
 Multilevel Factor Analysis, Structural Equation and Mixture Models [2]
 The nested Dirichlet process [168]
 Dirichlet Process Mixture Models for Nested Categorical Data [94]
 Context-Specific Nested Effects Models [183]
 Applications of finite mixtures of regression models [80]
 Nested hidden Markov chains for modeling dynamic unobserved heterogeneity in multilevel longitudinal data [18]
 Efficient hierarchical graph-based video segmentation [81]
 A Hierarchical Adaptive Information Fusion Method Based on Multimodal Kalman Filtering [99]
 Hierarchical Bayesian Models for Regularization in Sequential Learning [60]
 Hierarchical Bayesian inference in the visual cortex [120]
 FuzzyKalman Hierarchical Horizontal Motion Control of Underactuated ROVs [161]
 A Hierarchical Identification Method for SISO Fractional-order State-space Systems [171]
 Adaptive Classification by Variational Kalman Filtering [184]
 Decentralized Kalman filter comparison for distributed-parameter systems A [90]
 Comparison of Ensemble Kalman Filters under Non-Gaussianity [121]
 A two-stage ensemble Kalman filter for smooth data assimilation [102]
 Parameterizations for ensemble Kalman inversion [43]
 Dynamic Model of Visual Recognition Predicts Neural Response Properties in the Visual Cortex [164]

A Kalman Filter Model with Sparse Matrices in Spatial-Temporal Prediction [82]
 Hierarchical estimation [49]
 Decentralized structures for parallel Kalman filtering [88]
 Tracking dynamic sparse signals using Hierarchical Bayesian Kalman filters [110]
 Hierarchical Bayesian Time Series Models [23]
 Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter [8]
 Hierarchical adaptive Kalman filtering for interplanetary orbit determination [44]
 Hierarchical control of small autonomous helicopters [173]
 Hierarchical Simultaneous Localization and Mapping Based on Discrete Event Systems [58]
 The hierarchical atlas [129]
 Space-time Kalman filter [55]
 Hierarchical shape fitting using an iterated linear filter [19]
 Hierarchical identification of lifted state-space models for general dual-rate systems [63]
 A Hierarchical Bayes Ensemble Kalman Filter [192]
 Tracking dynamic sparse signals with hierarchical Kalman filters A [69]
 Towards an Hierarchical Kalman Filter Approach to Robot Localisation and Mapping [26]
 Cortical Mechanisms of Visual Recognition and Learning A Hierarchical Kalman Filter Model [163]
 Hierarchical Ensemble Kalman Filter [143]
 AN APPROACH FOR DISTRIBUTED KALMAN FILTERING [160]
 A Kalman-Based Coordination for Hierarchical State Estimation [217]
 Understanding and Designing Surface Computing with ZOIL and Squidy [98]
 Online Model Selection Based on the Variational Bayes [174]
 Fast Multipole Method as a Matrix-Free Hierarchical Low-Rank Approximation [209]
 Hierarchical Coupled-Geometry Analysis for Neuronal Structure and Activity Pattern Discovery [137]
 A Geometric Approach to the Theory of Evidence [57]
 Hierarchical Fuzzy Feature Similarity Combination for Presentation Slide Retrieval [118]
 Alignment with Kalman filter fitted tracks and reconstruction of B [6]
 Multilevel Stochastic Dynamic Process Models and Possible Applications in Global Financial Market Analysis and Surveillance [154]
 The Spatial Semantic Hierarchy [116]
 Information Geometry of Orthogonal Initializations and Training [180]
 Inference of stochastic parametrizations for model error treatment using nested ensemble Kalman filters [175]
 A perception-driven exploration hierarchical simultaneous localization and mapping for mobile robots adapted to search and rescue environments [204]
 Mixture Kalman filters [47]
 State-space models for bio-loggers A methodological road map [104]
 An Architectural Selection Framework for Data Fusion in Sensor Platforms [136]
 Fast Kalman filter using hierarchical matrices and a low-rank perturbative approach [172]
 Geometric Observers for Dynamically Evolving Curves [150]
 Information Geometry For Nonlinear Least-Squares Data Fitting And Calculation Of The Superconducting Superheating Field [190]
 Geometric algebra and information geometry for quantum computational software [36]
 A minimum discrimination information approach for hidden Markov modeling [66]
 The geometry of low-rank Kalman filters [27]
 An Information Geometry of Statistical Manifold Learning [182]
 Works on an information geometrodynamical approach to chaos [34]
 The information geometry of chaos [35]
 Intrinsic modeling of stochastic dynamical systems using empirical geometry [185]
 Natural encoding of information through interacting impulses [125]
 Quantum processes and functional geometry new perspectives in brain dynamics [170]
 Observer motion estimation and control from optical flow [210]
 Extension of information geometry for modelling non-statistical systems [10]
 Dynamical System and Information Geometry - A Complementary Approach to Complex Systems - [72]
 Recent theoretical progress on an information geometrodynamical approach to chaos [37]
 Fisher Information and the Combination of RGB [122]
 Quantum information geometry in the space of measurements [135]
 Integrating hidden information which is observed and the observer information regularities [124]
 The Information Geometry of Hierarchical Bayesian Models [176]
 An algebraic framework for the design of nonlinear observers with unknown inputs [15]
 Stochastic observers on Lie groups a tutorial [17]

Privacy-preserving nonlinear observer design using contraction analysis [151]
 Reduced-order observer design for nonlinear systems [38]
 A Robust Nonlinear Observer for Real-Time Attitude Estimation Using Low-Cost MEMS [83]
 Introduction to geometric nonlinear control linearization, observability, decoupling [166]
 State estimation and limited communication control for nonlinear robotic systems [165]
 A nonlinear observer for 6 DOF [14]
 Nonlinear Observer for 3D Rigid Body Motion Estimation Using Doppler Measurements [30]
 Observability and local observer construction for unknown parameters in linearly and nonlinearly parameterized systems [198]
 Nonlinear observers [113]
 Online natural gradient as a Kalman filter [153]
 Computational Methods of Information Geometry with Real-Time Applications in Audio Signal Processing. Mapose [62]
 Information Geometric Nonlinear Filtering [145]
 Connecting Information Geometry and Geometric Mechanics [123]
 Bayesian smoothing and information geometry [117]
 Information-geometric approach to inferring causal directions [97]
 Information geometry in optimization, machine learning and statistical inference [95]
 Information geometric analysis of phase transitions in complex patterns the case of the Gray-Scott reaction [87]
 An infinite-dimensional statistical manifold modelled on Hilbert space [144]
 The Geometry of Data Distance on Data Manifolds [51]
 Information Geometry and Primal-Dual Interior-point Algorithms [107]
 Foliations-Webs-Hessian Geometry-Information Geometry-Entropy and Cohomology [28]
 Nonlinear dynamical factor analysis for state change detection [96]
 On the Geometry of Maximum Entropy Problems [155]
 Information geometry for shape analysis Probabilistic models for shape matching and indexing [156]
 Geometric sliding mode control the linear and linearized theory [91]
 Sliding mode control of quantum systems [64]
 Mathematical Basis of Sliding Mode Control of an Uninterruptible Power Supply [3]
 Pole placement for overdetermined 2D systems [177]
 Kquota [48]
 Information Geometry of Non-Equilibrium Processes in a Bistable System with a Cubic Damping [93]
 Dynamical system and information geometry-a complementary approach to complex systems [73]
 Deep Gaussian mixture models [202]
 GeoSeq2Seq Information Geometric Sequence-to-Sequence Networks [21]
 Deep Gaussian Processes for Regression using Approximate Expectation Propagation [33]
 Deep Linear Discriminant Analysis [65]
 Composing graphical models with neural networks for structured representations and fast inference [103]
 Deep Kalman Filters [114]
 Sequential Neural Models with Stochastic Layers [70]
 Deep Variational Bayes Filters Unsupervised Learning of State Space Models from Raw Data [109]
 Hierarchical Mixtures of Experts and the EM [105]
 A scalable parallel algorithm for training a hierarchical mixture of neural experts [67]
 Adaptively Growing Hierarchical Mixtures of Experts [71]
 A multiscale neural network based on hierarchical nested bases [68]
 Hierarchical mixtures-of-experts for generalized linear models some results on denseness and consistency [101]
 Review of hierarchical models for data clustering and visualization [200]
 Are deep neural networks the best choice for modeling source code [89]
 Information geometry on extendable hierarchical large scale neural network model [130]
 Nested CRP [186]
 Exploration of Tree-based Hierarchical Softmax for Recurrent Language Models [100]
 Dirichlet belief networks for topic structure learning [214]
 Deep Infinite Admixture Models [189]
 Hierarchical Neural Networks for Image Interpretation [22]
 Nonparametric Variational Auto-Encoders for Hierarchical Representation Learning [79]
 Latent tree models for hierarchical topic detection [46]
 A Statistical View of Deep Learning [139]
 Deep Kalman Filtering Network for Video Compression Artifact Reduction [133]
 Simple recurrent network trained by RTRL and extended Kalman filter algorithms [41]
 Recurrent neural network training with the extended kalman filter [191]

A Neural Implementation of the Kalman Filter [207]
 Information Geometric Approach to Recursive Update in Nonlinear Filtering [127]
 Some observations on the use of the extended Kalman filter as a recurrent network learning algorithm [206]
 Nonlinear Bayesian Filters for Training Recurrent Neural Networks [11]
 Deep Gaussian Processes [59]
 Finite Mixture Models in Market Segmentation A Review and Suggestions for Best Practices [193]
 A recursive filter for linear systems on Riemannian manifolds [194]
 Mathematical and Physical Models for the Estimation of Wind-Wave Power Potential in the Eastern Mediterranean Sea [74]
 A Completed Information Projection Interpretation of Expectation Propagation [203]
 Kalman Filtering With Intermittent Observations Convergence for Semi-Markov Chains and an Intrinsic Performance Measure [40]
 Invariant kalman filtering [16]
 Kalman filtering with a class of geometric state equality constraints [45]
 Initialization of the Kalman filter without assumptions on the initial state [128]
 Extended information filter on matrix Lie groups [42]
 Geometric illustration of the Kalman filter gain and covariance update algorithms [115]
 Geometric Theory of Nonlinear Filtering [138]
 Information geometry-based action decision-making for target tracking by fixed-wing unmanned aerial vehicle [215]
 Feedback Particle Filter on Riemannian Manifolds and Matrix Lie Groups [213]
 Geometric remarks on Kalman filtering with intermittent observations [39]
 Bayesian Nonlinear Filtering via Information Geometric Optimization [7]
 Penalising Model Component Complexity A Principled, Practical Approach to Constructing Priors [178]
 Information Geometry and the Dynamic Detection of Financial Bubbles and Crashes [61]
 THE [216]
 Information geometric probability models in statistical signal processing [53]
 Sensitivity Analysis for Bayesian Hierarchical Models [169]
 Mixture decompositions of exponential families - using a decomposition of their sample spaces [141]
 Bayesian hierarchically weighted finite mixture models for samples of distributions [167]
 A hierarchical finite mixture model that accommodates zero-inflated counts, non-independence, and heterogeneity [142]
 Modeling when people quit Bayesian censored geometric models with hierarchical and latent-mixture extensions [152]
 Information geometry on hierarchy of probability distributions [5]
 Hierarchical modeling for univariate spatial data [1]
 Conditional Akaike information for mixed-effects models [195]
 Scalable Deep Poisson Factor Analysis for Topic Modeling [75]
 The variational hierarchical EM [54]
 Stochastic variational inference [92]
 Kernels on Structured Objects Through Nested Histograms [56]
 Linear Response Methods for Accurate Covariance Estimates from Mean Field Variational Bayes [78]
 Guaranteed Deterministic Bounds on the total variation Distance between univariate mixtures [149]
 On the geometry of mixed states and the Fisher information tensor [52]
 Independent Factor Analysis [13]
 Mixture Model-based Statistical Pattern Recognition of Clustered or Longitudinal Data [146]
 Convergence of latent mixing measures in finite and infinite mixture models [147]
 Hierarchical Markov normal mixture models with applications to financial asset returns [77]
 Efficient, Multifidelity Perceptual Representations via Hierarchical Gaussian Mixture Models [181]
 Bayesian Mixture Models [131]

References

- [1] Hierarchical modeling for univariate spatial data. In *C&H/CRC Monographs on Statistics & Applied Probability*. Chapman and Hall/CRC, dec 2003.
- [2] Multilevel factor analysis, structural equation and mixture models. In *Wiley Series in Probability and Statistics*, pages 189–200. John Wiley & Sons, Ltd, dec 2010.
- [3] Mathematical basis of sliding mode control of an uninterruptible power supply. *Acta Polytechnica Hungarica*, 11(03), feb 2014.
- [4] Dynamic factor models, jan 2016.

- [5] S.-I. Amari. Information geometry on hierarchy of probability distributions. *IEEE Transactions on Information Theory*, 47(5):1701–1711, jul 2001.
- [6] Jan Mennis Amoraal. Alignment with kalman filter fitted tracks and reconstruction of $b\{0\}_{-s}$ to $j/\psi\phi$ decays. *Technical report*, 2011.
- [7] and Yongqiang Cheng, Xiang Li, Hongqiang Wang, Xiaoqiang Hua, and Yuliang Qin. Bayesian nonlinear filtering via information geometric optimization. *Entropy*, 19(12):655, dec 2017.
- [8] Jeffrey L. Anderson. Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter. *Physica D: Nonlinear Phenomena*, 230(1-2):99–111, jun 2007.
- [9] Daniele Angelosante, Stergios I. Roumeliotis, and Georgios B. Giannakis. Lasso-kalman smoother for tracking sparse signals. In *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*. IEEE, 2009.
- [10] Ben Anthonis. Extension of information geometry for modelling non-statistical systems. 2014.
- [11] Ienkaran Arasaratnam and Simon Haykin. Nonlinear bayesian filters for training recurrent neural networks. In *MICAI*, volume 5317 of *Lecture Notes in Computer Science*, pages 12–33. Springer, 2008.
- [12] Aleksandr Y. Aravkin, James V. Burke, and Gianluigi Pillonetto. Optimization viewpoint on kalman smoothing with applications to robust and sparse estimation. In *Compressed Sensing & Sparse Filtering*, pages 237–280. Springer Berlin Heidelberg, sep 2013.
- [13] H. Attias. Independent factor analysis. *Neural Computation*, 11(4):803–851, may 1999.
- [14] Grant Baldwin, Robert E. Mahony, and Jochen Trumpf. A nonlinear observer for 6 DOF pose estimation from inertial and bearing measurements. In *ICRA*, pages 2237–2242. IEEE, 2009.
- [15] Jean-Pierre Barbot, Michel Fliess, and Thierry Floquet. An algebraic framework for the design of nonlinear observers with unknown inputs. In *CDC*, pages 384–389. IEEE, 2007.
- [16] Axel Barrau and Silvere Bonnabel. Invariant kalman filtering. *Annual Review of Control, Robotics, and Autonomous Systems*, 1:237–257, 2018.
- [17] Axel Barrau and Silvere Bonnabel. Stochastic observers on lie groups: a tutorial. In *CDC*, pages 1264–1269. IEEE, 2018.
- [18] Francesco Bartolucci and Monia Lupparelli. Nested hidden markov chains for modeling dynamic unobserved heterogeneity in multilevel longitudinal data. 2012.
- [19] Adam Baumberg. Hierarchical shape fitting using an iterated linear filter. *Image Vision Comput.*, 16(5):329–335, 1998.
- [20] Peter Baxendale, Pavel Chigansky, and Robert Sh. Liptser. Asymptotic stability of the wonham filter: Ergodic and nonergodic signals. *SIAM J. Control and Optimization*, 43(2):643–669, 2004.
- [21] Alessandro Bay and Biswa Sengupta. Geoseq2seq: Information geometric sequence-to-sequence networks. *arXiv preprint arXiv:1710.09363*, 2017.
- [22] Sven Behnke. *Hierarchical Neural Networks for Image Interpretation*, volume 2766 of *Lecture Notes in Computer Science*. Springer, 2003.
- [23] L. Mark Berliner. Hierarchical bayesian time series models. In *Maximum Entropy and Bayesian Methods*, pages 15–22. Springer Netherlands, 1996.
- [24] David M. Blei, Thomas L. Griffiths, and Michael I. Jordan. The nested chinese restaurant process and bayesian nonparametric inference of topic hierarchies. *J. ACM*, 57(2):7:1–7:30, 2010.
- [25] David M. Blei and John D. Lafferty. Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning - ICML 06*. ACM Press, 2006.
- [26] Jeannette Bohg. Towards an hierarchical kalman filter approach to robot localisation and mapping. 12 2005.
- [27] Silvere Bonnabel and Rodolphe Sepulchre. The geometry of low-rank kalman filters. *CoRR*, abs/1203.4049, 2012.
- [28] Michel Nguiffo Boyom. Foliations-webs-hessian geometry-information geometry-entropy and cohomology. *Entropy*, 18(12):433, 2016.

- [29] Nicole B"auerle, Igor Gilitschenski, and Uwe D. Hanebeck. Exact and approximate hidden markov chain filters based on discrete observations. *CoRR*, abs/1411.0849, 2014.
- [30] Sérgio Bras, Maziar Izadi, Carlos Silvestre, Amit K. Sanyal, and Paulo Jorge Ramalho Oliveira. Nonlinear observer for 3d rigid body motion estimation using doppler measurements. *IEEE Trans. Automat. Contr.*, 61(11):3580–3585, 2016.
- [31] A. Bravetti, C.S. Lopez-Monsalvo, and F. Nettel. Contact symmetries and hamiltonian thermodynamics. *Annals of Physics*, 361:377–400, oct 2015.
- [32] D. Brigo and F. LeGland. A finite dimensional filter with exponential conditional density. In *Proceedings of the 36th IEEE Conference on Decision and Control*. IEEE.
- [33] Thang D. Bui, Daniel Hernéandez-Lobato, José Miguel Hernéandez-Lobato, Yingzhen Li, and Richard E. Turner. Deep gaussian processes for regression using approximate expectation propagation. In *ICML*, volume 48 of *JMLR Workshop and Conference Proceedings*, pages 1472–1481. JMLR.org, 2016.
- [34] Carlo Cafaro. Works on an information geometrodynamical approach to chaos. *Chaos, Solitons & Fractals*, 41(2):886–891, 2009.
- [35] Carlo Cafaro. The information geometry of chaos. *arXiv preprint arXiv:1601.07935*, 2016.
- [36] Carlo Cafaro. Geometric algebra and information geometry for quantum computational software. *Physica A: Statistical Mechanics and its Applications*, 470:154–196, mar 2017.
- [37] Carlo Cafaro, Marcelo de Souza Lauretto, Carlos Alberto de Braganca Pereira, and Julio Michael Stern. Recent theoretical progress on an information geometrodynamical approach to chaos. In *AIP Conference Proceedings*. AIP, 2008.
- [38] Daniele Carnevale, Dimitrios Karagiannis, and Alessandro Astolfi. Reduced-order observer design for nonlinear systems. In *2007 European Control Conference (ECC)*, pages 559–564. IEEE, 2007.
- [39] Andrea Censi. Geometric remarks on kalman filtering with intermittent observations. *arXiv preprint arXiv:0906.1637*, 2009.
- [40] Andrea Censi. Kalman filtering with intermittent observations: Convergence for semi-markov chains and an intrinsic performance measure. *IEEE Trans. Automat. Contr.*, 56(2):376–381, 2011.
- [41] Michal Cernansky and Lubica Benuskova. Simple recurrent network trained by rtl and extended kalman filter algorithms. *Neural Network World*, 13(3):223–234, 2003.
- [42] Josip &Cedil;esić, Ivan Marković, Mario Bukal, and Ivan Petrović. Extended information filter on matrix lie groups. *Automatica*, 82:226–234, aug 2017.
- [43] Neil K Chada, Marco A Iglesias, Lassi Roininen, and Andrew M Stuart. Parameterizations for ensemble kalman inversion. *Inverse Problems*, 34(5):055009, apr 2018.
- [44] W.S. Chaer, R.H. Bishop, and J. Ghosh. Hierarchical adaptive kalman filtering for interplanetary orbit determination. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3):883–896, jul 1998.
- [45] Paul Chauchat, Axel Barrau, and Silvere Bonnabel. Kalman filtering with a class of geometric state equality constraints. In *CDC*, pages 2581–2586. IEEE, 2017.
- [46] Peixian Chen, Nevin L. Zhang, Tengfei Liu, Leonard K. M. Poon, Zhourong Chen, and Farhan Khawar. Latent tree models for hierarchical topic detection. *Artif. Intell.*, 250:105–124, 2017.
- [47] Rong Chen and Jun song Liu. Mixture kalman filters. 2000.
- [48] Jaehyung Choi and Andrew P. Mullhaupt. K"ahlerian information geometry for signal processing. *Entropy*, 17(4):1581–1605, 2015.
- [49] Chee-Yee Chong. Hierarchical estimation. In *Proc. MIT/ONR Workshop on C3*, pages 205–220, 1979.
- [50] Panayiotis Christodoulou, Sotirios P. Chatzis, and Andreas S. Andreou. A recurrent latent variable model for supervised modeling of high-dimensional sequential data. In *2018 Innovations in Intelligent Systems and Applications (INISTA)*. IEEE, jul 2018.
- [51] Casey Chu. The geometry of data: Distance on data manifolds. 2016.

- [52] I. Contreras, E. Ercolelli, and M. Schiavina. On the geometry of mixed states and the fisher information tensor. *Journal of Mathematical Physics*, 57(6):062209, jun 2016.
- [53] Russell Costa. Information geometric probability models in statistical signal processing. 2016.
- [54] Emanuele Coviello, Antoni B. Chan, and Gert R. G. Lanckriet. The variational hierarchical EM algorithm for clustering hidden markov models. In *NIPS*, pages 413–421, 2012.
- [55] Noel Cressie and Christopher K Wikle. Space-time kalman filter. *Wiley StatsRef: Statistics Reference Online*, 2014.
- [56] Marco Cuturi and Kenji Fukumizu. Kernels on structured objects through nested histograms. In *NIPS*, pages 329–336. MIT Press, 2006.
- [57] Fabio Cuzzolin. A geometric approach to the theory of evidence. *IEEE Trans. Systems, Man, and Cybernetics, Part C*, 38(4):522–534, 2008.
- [58] Xufeng Dai and Juan Du. Hierarchical simultaneous localization and mapping based on discrete event systems. In *2009 International Conference on Measuring Technology and Mechatronics Automation*. IEEE, 2009.
- [59] Andreas C. Damianou and Neil D. Lawrence. Deep gaussian processes. In *AISTATS*, 2013.
- [60] João F. G. de Freitas, Mahesan Niranjan, and Andrew H. Gee. Hierarchical bayesian models for regularization in sequential learning. *Neural Computation*, 12(4):933–953, 2000.
- [61] Guilherme DL Demos. *Information Geometry and the Dynamic Detection of Financial Bubbles and Crashes*. PhD thesis, ETH Zurich, 2017.
- [62] Arnaud Dessein. *Computational Methods of Information Geometry with Real-Time Applications in Audio Signal Processing. (Méthodes Computationnelles en Géométrie de l’Information et Applications Temps Réel au Traitement du Signal Audio)*. PhD thesis, Pierre and Marie Curie University, Paris, France, 2012.
- [63] Feng Ding and Tongwen Chen. Hierarchical identification of lifted state-space models for general dual-rate systems. *IEEE Trans. on Circuits and Systems*, 52-I(6):1179–1187, 2005.
- [64] Daoyi Dong and Ian R Petersen. Sliding mode control of quantum systems. *New Journal of Physics*, 11(10):105033, 2009.
- [65] Matthias Dorfer, Rainer Kelz, and Gerhard Widmer. Deep linear discriminant analysis. *CoRR*, abs/1511.04707, 2016.
- [66] Yariv Ephraim, Amir Dembo, and Lawrence R. Rabiner. A minimum discrimination information approach for hidden markov modeling. *IEEE Trans. Information Theory*, 35(5):1001–1013, 1989.
- [67] Pablo A. Estévez, Héléne Paugam-Moisy, Didier Puzeat, and Manuel Ugarte. A scalable parallel algorithm for training a hierarchical mixture of neural experts. *Parallel Computing*, 28(6):861–891, 2002.
- [68] Yuwei Fan, Jordi Feliu-Fabà, Lin Lin, Lexing Ying, and Leonardo Zepeda-Núñez. A multiscale neural network based on hierarchical nested bases. *Research in the Mathematical Sciences*, 6(2), mar 2019.
- [69] Jason Filos, Evripidis Karseras, Wei Dai, and Shulin Yan. Tracking dynamic sparse signals with hierarchical kalman filters: A case study. In *DSP*, pages 1–6. IEEE, 2013.
- [70] Marco Fraccaro, Søren Kaae Sønderby, Ulrich Paquet, and Ole Winther. Sequential neural models with stochastic layers. In *NIPS*, pages 2199–2207, 2016.
- [71] Jürgen Fritsch, Michael Finke, and Alex Waibel. Adaptively growing hierarchical mixtures of experts. In *NIPS*, pages 459–465. MIT Press, 1996.
- [72] Masatoshi Funabashi. Dynamical system and information geometry - a complementary approach to complex systems -. 2010.
- [73] Masatoshi Funabashi. *Dynamical system and information geometry-a complementary approach to complex systems*. PhD thesis, Ecole Polytechnique X, 2010.
- [74] George Galanis, George Emmanouil, Christina Kalogeri, and George Kallos. Mathematical and physical models for the estimation of wind-wave power potential in the eastern mediterranean sea. In *Lecture Notes in Earth System Sciences*, pages 561–564. Springer Berlin Heidelberg, oct 2013.
- [75] Zhe Gan, Changyou Chen, Ricardo Henao, David E. Carlson, and Lawrence Carin. Scalable deep poisson factor analysis for topic modeling. In *ICML*, volume 37 of *JMLR Workshop and Conference Proceedings*, pages 1823–1832. JMLR.org, 2015.

- [76] J.P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems applications to bioreactors. *IEEE Transactions on Automatic Control*, 37(6):875–880, jun 1992.
- [77] John Geweke and Gianni Amisano. Hierarchical markov normal mixture models with applications to financial asset returns. *Journal of Applied Econometrics*, 26(1):1–29.
- [78] Ryan Giordano, Tamara Broderick, and Michael I. Jordan. Linear response methods for accurate covariance estimates from mean field variational bayes. In *NIPS*, pages 1441–1449, 2015.
- [79] Prasoon Goyal, Zhiting Hu, Xiaodan Liang, Chenyu Wang, Eric P. Xing, and Carnegie Mellon. Nonparametric variational auto-encoders for hierarchical representation learning. In *ICCV*, pages 5104–5112. IEEE Computer Society, 2017.
- [80] Bettina Grün, Friedrich Leisch, et al. Applications of finite mixtures of regression models. *URL: <http://cran.r-project.org/web/packages/flexmix/vignettes/regression-examples.pdf>*, 2007.
- [81] Matthias Grundmann, Vivek Kwatra, Mei Han, and Irfan Essa. Efficient hierarchical graph-based video segmentation. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, jun 2010.
- [82] Bingjing Gu. A kalman filter model with sparse matrices in spatial-temporal prediction.
- [83] Jos&e-Fermi Guerrero-Castellanos, Heberto Madrigal-Sastre, Sylvain Durand, Flor Lizeth Torres Ortiz, and Germ&an Ardul Mu&oz Hern&andez. A robust nonlinear observer for real-time attitude estimation using low-cost MEMS inertial sensors. *Sensors*, 13(11):15138–15158, 2013.
- [84] X Guo and G Yin. Wonham filter with random parameters: Rate of convergence and error bounds, 2003. *Preprint*.
- [85] Xin Guo and George Yin. The wonham filter with random parameters: rate of convergence and error bounds. *IEEE Trans. Automat. Contr.*, 51(3):460–464, 2006.
- [86] Shirin Hadizadeh, Bojan Ramadanovic, Gordon W. Semenoff, and Donovan Young. Free energy and phase transition of the matrix model on a plane wave. *Physical Review D*, 71(6), mar 2005.
- [87] Omri Har-Shemesh, Rick Quax, Alfons G Hoekstra, and Peter M A Slood. Information geometric analysis of phase transitions in complex patterns: the case of the gray-scott reaction–diffusion model. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(4):043301, apr 2016.
- [88] H.R. Hashemipour, S. Roy, and A.J. Laub. Decentralized structures for parallel kalman filtering. *IEEE Transactions on Automatic Control*, 33(1):88–94, jan 1988.
- [89] Vincent J. Hellendoorn and Premkumar Devanbu. Are deep neural networks the best choice for modeling source code? In *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering - ESEC/FSE 2017*. ACM Press, 2017.
- [90] Zulkifli Hidayat, Robert Babuska, Bart De Schutter, and Alfredo N&u&nez. Decentralized kalman filter comparison for distributed-parameter systems: A case study for a 1d heat conduction process. In *ETFA*, pages 1–8. IEEE, 2011.
- [91] R.M. Hirschorn and A.D. Lewis. Geometric sliding mode control: the linear and linearized theory. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*. IEEE.
- [92] Matthew D Hoffman, David M Blei, Chong Wang, and John Paisley. Stochastic variational inference. *The Journal of Machine Learning Research*, 14(1):1303–1347, 2013.
- [93] Rainer Hollerbach and Eun-Jin Kim. Information geometry of non-equilibrium processes in a bistable system with a cubic damping. *Entropy*, 19(6):268, 2017.
- [94] Jingchen Hu. Dirichlet process mixture models for nested categorical data. 2015.
- [95] Shun ichi Amari. Information geometry in optimization, machine learning and statistical inference. *Frontiers of Electrical and Electronic Engineering in China*, 5(3):241–260, jul 2010.
- [96] Alexander Ilin, Harri Valpola, and Erkki Oja. Nonlinear dynamical factor analysis for state change detection. *IEEE Trans. Neural Networks*, 15(3):559–575, 2004.
- [97] Dominik Janzing, Joris M. Mooij, Kun Zhang, Jan Lemeire, Jakob Zscheischler, Povilas Daniusis, Bastian Steudel, and Bernhard Sch"olkopf. Information-geometric approach to inferring causal directions. *Artif. Intell.*, 182-183:1–31, 2012.
- [98] Hans-Christian Jetter, Werner A. König, and Harald Reiterer. Understanding and designing surface computing with zoil and squid. 2009.

- [99] Zhu Ji, Jin Hua Bian, and Luhua Zhang. A hierarchical adaptive information fusion method based on multimodal kalman filtering. 2016.
- [100] Nan Jiang, Wenge Rong, Min Gao, Yikang Shen, and Zhang Xiong. Exploration of tree-based hierarchical softmax for recurrent language models. In *IJCAI*, pages 1951–1957. ijcai.org, 2017.
- [101] Wenxin Jiang and Martin A. Tanner. Hierarchical mixtures-of-experts for generalized linear models: some results on denseness and consistency. In *AISTATS*. Society for Artificial Intelligence and Statistics, 1999.
- [102] Craig J. Johns and Jan Mandel. A two-stage ensemble kalman filter for smooth data assimilation. *Environmental and Ecological Statistics*, 15(1):101–110, sep 2007.
- [103] Matthew J. Johnson, David K. Duvenaud, Alexander B. Wiltschko, Ryan P. Adams, and Sandeep R. Datta. Composing graphical models with neural networks for structured representations and fast inference. In *NIPS*, pages 2946–2954, 2016.
- [104] I.D. Jonsen, M. Basson, S. Bestley, M.V. Bravington, T.A. Patterson, M.W. Pedersen, R. Thomson, U.H. Thygesen, and S.J. Wotherspoon. State-space models for bio-loggers: A methodological road map. *Deep Sea Research Part II: Topical Studies in Oceanography*, 88-89:34–46, apr 2013.
- [105] Michael I. Jordan and Robert A. Jacobs. Hierarchical mixtures of experts and the EM algorithm. *Neural Computation*, 6(2):181–214, 1994.
- [106] Scott Joslin, Anh Le, and Kenneth J Singleton. Why gaussian macro-finance term structure models are (nearly) unconstrained factor-vars. *Journal of Financial Economics*, 109(3):604–622, 2013.
- [107] Satoshi Kakiyama, Atsumi Ohara, and Takashi Tsuchiya. Information geometry and primal-dual interior-point algorithms. 2009.
- [108] Dimitri Kanevsky, Avishy Carmi, Lior Horesh, Pini Gurfil, Bhuvana Ramabhadran, and Tara N. Sainath. Kalman filtering for compressed sensing. In *FUSION*, pages 1–8. IEEE, 2010.
- [109] Maximilian Karl, Maximilian S"olch, Justin Bayer, and Patrick van der Smagt. Deep variational bayes filters: Unsupervised learning of state space models from raw data. In *ICLR*. OpenReview.net, 2017.
- [110] Evripidis Karseras, Kin K. Leung, and Wei Dai. Tracking dynamic sparse signals using hierarchical bayesian kalman filters. In *ICASSP*, pages 6546–6550. IEEE, 2013.
- [111] Don H Kim and Marcel Pribsch. Estimation of multi-factor shadow-rate term structure models. *Federal reserve board discussion paper series, Federal Reserve Board*, 2013.
- [112] Ulrich Kohler, Kristian Bernt Karlson, and Anders Holm. Comparing coefficients of nested nonlinear probability models. *The Stata Journal: Promoting communications on statistics and Stata*, 11(3):420–438, oct 2011.
- [113] AJ Krener. Nonlinear observers. 2003.
- [114] Rahul G. Krishnan, Uri Shalit, and David Sontag. Deep kalman filters. *CoRR*, abs/1511.05121, 2015.
- [115] T. Kronhamn. Geometric illustration of the kalman filter gain and covariance update algorithms. *IEEE Control Systems Magazine*, 5(2):41–43, may 1985.
- [116] Benjamin Kuipers. The spatial semantic hierarchy. *Artificial Intelligence*, 119(1-2):191–233, may 2000.
- [117] Rudolf Kulhavy. Bayesian smoothing and information geometry. *NATO SCIENCE SERIES SUB SERIES III COMPUTER AND SYSTEMS SCIENCES*, 190:319–340, 2003.
- [118] Azadeh Kushki, Mohammed Ajmal, and Konstantinos N. Plataniotis. Hierarchical fuzzy feature similarity combination for presentation slide retrieval. *EURASIP J. Adv. Sig. Proc.*, 2008, 2008.
- [119] Peter D. Lax. Integrals of nonlinear equations of evolution and solitary waves. *Communications on Pure and Applied Mathematics*, 21(5):467–490, sep 1968.
- [120] Tai Sing Lee and David Mumford. Hierarchical bayesian inference in the visual cortex. *Journal of the Optical Society of America A*, 20(7):1434, jul 2003.
- [121] Jing Lei, Peter Bickel, and Chris Snyder. Comparison of ensemble kalman filters under non-gaussianity. *Monthly Weather Review*, 138(4):1293–1306, apr 2010.
- [122] Reiner Lenz and Vasileios Zografos. Fisher information and the combination of RGB channels. In *CCIW*, volume 7786 of *Lecture Notes in Computer Science*, pages 250–264. Springer, 2013.

- [123] Melvin Leok and Jun Zhang. Connecting information geometry and geometric mechanics. *Entropy*, 19(10):518, 2017.
- [124] Vladimir S. Lerner. Integrating hidden information which is observed and the observer information regularities. 2013.
- [125] Vladimir S. Lerner. Natural encoding of information through interacting impulses. In *2016 Future Technologies Conference (FTC)*. IEEE, dec 2016.
- [126] Qunhua Li, Michael J. MacCoss, and Matthew Stephens. A nested mixture model for protein identification using mass spectrometry. *The Annals of Applied Statistics*, 4(2):962–987, jun 2010.
- [127] Yubo Li, Yongqiang Cheng, Xiang Li, Xiaoqiang Hua, and Yuliang Qin. Information geometric approach to recursive update in nonlinear filtering. *Entropy*, 19(2):54, 2017.
- [128] Magnus Linderöth, Kristian Soltesz, Anders Robertsson, and Rolf Johansson. Initialization of the kalman filter without assumptions on the initial state. In *ICRA*, pages 4992–4997. IEEE, 2011.
- [129] B. Lisien, D. Morales, D. Silver, G. Kantor, I. Rekleitis, and H. Choset. The hierarchical atlas. *IEEE Transactions on Robotics*, 21(3):473–481, jun 2005.
- [130] Yun-Hui Liu, Si-Wei Luo, Ai-Jun Li, Hua Huang, and Jin-Wei Wen. Information geometry on extendable hierarchical large scale neural network model. In *Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693)*. IEEE.
- [131] Zhihui Liu et al. *Bayesian Mixture Models*. PhD thesis, 2010.
- [132] H. Loeliger. An introduction to factor graphs. *IEEE Signal Processing Magazine*, 21(1):28–41, jan 2004.
- [133] Guo Lu, Wanli Ouyang, Dong Xu, Xiaoyun Zhang, Zhiyong Gao, and Ming-Ting Sun. Deep kalman filtering network for video compression artifact reduction. In *ECCV (14)*, volume 11218 of *Lecture Notes in Computer Science*, pages 591–608. Springer, 2018.
- [134] Kimberly S. Maier. A rasch hierarchical measurement model. *Journal of Educational and Behavioral Statistics*, 26(3):307–330, sep 2001.
- [135] Warner A. Miller. Quantum information geometry in the space of measurements. In Michael Hayduk and Eric Donkor, editors, *Quantum Information Science, Sensing, and Computation X*. SPIE, may 2018.
- [136] Atif R. Mirza. An architectural selection framework for data fusion in sensor platforms. 2006.
- [137] Gal Mishne, Ronen Talmon, Ron Meir, Jackie Schiller, Maria Lavzin, Uri Dubin, and Ronald R. Coifman. Hierarchical coupled-geometry analysis for neuronal structure and activity pattern discovery. *J. Sel. Topics Signal Processing*, 10(7):1238–1253, 2016.
- [138] Sanjoy K. Mitter. Geometric theory of nonlinear filtering. Technical report, mar 1983.
- [139] Shakir Mohamed. A statistical view of deep learning. 2015.
- [140] Peter CM Molenaar. State space techniques in structural equation modeling. Retrieved from the Pennsylvania State Department of Human Development and Family Studies website: <http://www.hhdev.psu.edu/hdfs/faculty/docs/StateSpaceTechniques.pdf>, 2003.
- [141] Guido Mont’ufar. Mixture decompositions of exponential families - using a decomposition of their sample spaces. *Kybernetika*, 49(1):23–39, 2013.
- [142] Grant Morgan and A. Beaujean. An investigation of growth mixture models for studying the flynn effect. *Journal of Intelligence*, 2(4):156–179, oct 2014.
- [143] Inge Myrseth and Henning Omre. Hierarchical ensemble kalman filter. *SPE Journal*, 15(02):569–580, jun 2010.
- [144] Nigel J. Newton. An infinite-dimensional statistical manifold modelled on hilbert space. *Journal of Functional Analysis*, 263(6):1661–1681, sep 2012.
- [145] Nigel J. Newton. Information geometric nonlinear filtering. *CoRR*, abs/1502.04638, 2015.
- [146] Shu Kay Angus Ng and Geoffrey J. McLachlan. Mixture model-based statistical pattern recognition of clustered or longitudinal data. 2005.
- [147] XuanLong Nguyen. Convergence of latent mixing measures in finite and infinite mixture models. *The Annals of Statistics*, 41(1):370–400, feb 2013.

- [148] Kai Ni, Lawrence Carin, and David B. Dunson. Multi-task learning for sequential data via ihmms and the nested dirichlet process. In *ICML*, volume 227 of *ACM International Conference Proceeding Series*, pages 689–696. ACM, 2007.
- [149] Frank Nielsen and Ke Sun. Guaranteed deterministic bounds on the total variation distance between univariate mixtures. In *MLSP*, pages 1–6. IEEE, 2018.
- [150] Marc Niethammer, Patricio A. Vela, and Allen R. Tannenbaum. Geometric observers for dynamically evolving curves. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(6):1093–1108, 2008.
- [151] Jerome Le Ny. Privacy-preserving nonlinear observer design using contraction analysis. In *CDC*, pages 4499–4504. IEEE, 2015.
- [152] Kensuke Okada, Joachim Vandekerckhove, and Michael D. Lee. Modeling when people quit: Bayesian censored geometric models with hierarchical and latent-mixture extensions. *Behavior Research Methods*, 50(1):406–415, mar 2017.
- [153] Yann Ollivier. Online natural gradient as a kalman filter. *Electronic Journal of Statistics*, 12(2):2930–2961, 2018.
- [154] Heping Pan. Multilevel stochastic dynamic process models and possible applications in global financial market analysis and surveillance. In *JCIS*. Atlantis Press, 2006.
- [155] Michele Pavon and Augusto Ferrante. On the geometry of maximum entropy problems. *SIAM Review*, 55(3):415–439, 2013.
- [156] Adrian M Peter. *Information geometry for shape analysis: Probabilistic models for shape matching and indexing*. University of Florida, 2008.
- [157] Pedro Daniel Prieto-Martínez. Geometrical structures of higher-order dynamical systems and field theories. 2014.
- [158] Chenlu Qiu, Wei Lu, and Namrata Vaswani. Real-time dynamic MR image reconstruction using kalman filtered compressed sensing. In *ICASSP*, pages 393–396. IEEE, 2009.
- [159] Minh Quach, Nicolas Brunel, and Florence d’Alché-Buc. Estimating parameters and hidden variables in non-linear state-space models based on odes for biological networks inference. *Bioinformatics*, 23(23):3209–3216, 2007.
- [160] Rogério Bastos Quirino and Celso Pascoli Bottura. An approach for distributed kalman filtering. 2001.
- [161] Francesco M. Raimondi and Maurizio Melluso. Fuzzy/kalman hierarchical horizontal motion control of underactuated ROVs. *International Journal of Advanced Robotic Systems*, 7(2):12, jun 2010.
- [162] Maziar Raissi and George E. Karniadakis. Hidden physics models: Machine learning of nonlinear partial differential equations. *J. Comput. Physics*, 357:125–141, 2018.
- [163] Rajesh P. N. Rao. Cortical mechanisms of visual recognition and learning: A hierarchical kalman filter model. 2007.
- [164] Rajesh P. N. Rao and Dana H. Ballard. Dynamic model of visual recognition predicts neural response properties in the visual cortex. *Neural Computation*, 9(4):721–763, 1997.
- [165] Henrik Reh binder. *State estimation and limited communication control for nonlinear robotic systems*. PhD thesis, Matematik, 2001.
- [166] Witold Respondek. Introduction to geometric nonlinear control; linearization, observability, decoupling. Technical report, 2002.
- [167] A. Rodriguez, D. B. Dunson, and J. Taylor. Bayesian hierarchically weighted finite mixture models for samples of distributions. *Biostatistics*, 10(1):155–171, may 2008.
- [168] Abel Rodriguez, David B Dunson, and Alan E Gelfand. The nested dirichlet process. *Journal of the American Statistical Association*, 103(483):1131–1154, 2008.
- [169] Małgorzata Roos, Thiago G. Martins, Leonhard Held, and Håvard Rue. Sensitivity analysis for bayesian hierarchical models. *Bayesian Analysis*, 10(2):321–349, jun 2015.
- [170] Sisir Roy and Menas Kafatos. Quantum processes and functional geometry: new perspectives in brain dynamics. *FORMA-TOKYO-*, 19(1):69–84, 2004.
- [171] Behrooz Safarinejadian, Mojtaba Asad, and Amin Torabi Jahromi. A hierarchical identification method for siso fractional-order state-space systems. *The Modares Journal of Electrical Engineering*, 13(4):17–28, 2014.
- [172] Arvind K Saibaba, Eric L Miller, and Peter K Kitanidis. Fast kalman filter using hierarchical matrices and a low-rank perturbative approach. *Inverse Problems*, 31(1):015009, jan 2015.

- [173] C.P. Sanders, P.A. DeBitetto, E. Feron, H.F. Vuong, and N. Leveson. Hierarchical control of small autonomous helicopters. In *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No.98CH36171)*. IEEE.
- [174] Masa-aki Sato. Online model selection based on the variational bayes. *Neural Computation*, 13(7):1649–1681, 2001.
- [175] Guillermo Scheffler, Juan Ruiz, and Manuel Pulido. Inference of stochastic parametrizations for model error treatment using nested ensemble kalman filters. *Quarterly Journal of the Royal Meteorological Society*, apr 2019.
- [176] Cheng Dong Seon. The information geometry of hierarchical bayesian models.
- [177] Liran Shaul and Victor Vinnikov. Pole placement for overdetermined 2d systems. 2017.
- [178] Daniel Simpson, Håvard Rue, Andrea Riebler, Thiago G. Martins, and Sigrunn H. Sørbye. Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, 32(1):1–28, feb 2017.
- [179] Padhraic Smyth. Belief networks, hidden markov models, and markov random fields: A unifying view. *Pattern Recognition Letters*, 18(11-13):1261–1268, 1997.
- [180] Piotr A. Sokol and Il Memming Park. Information geometry of orthogonal initializations and training. *CoRR*, abs/1810.03785, 2018.
- [181] Shobhit Srivastava and Nathan Michael. Efficient, multifidelity perceptual representations via hierarchical gaussian mixture models. *IEEE Transactions on Robotics*, 35(1):248–260, feb 2019.
- [182] Ke Sun and Stéphane Marchand-Maillet. An information geometry of statistical manifold learning. In *ICML*, volume 32 of *JMLR Workshop and Conference Proceedings*, pages 1–9. JMLR.org, 2014.
- [183] Yuriy Sverchkov, Yi-Hsuan Ho, Audrey P. Gasch, and Mark Craven. Context-specific nested effects models. In *RECOMB*, volume 10812 of *Lecture Notes in Computer Science*, pages 194–210. Springer, 2018.
- [184] Peter Sykacek and Stephen J. Roberts. Adaptive classification by variational kalman filtering. In *NIPS*, pages 737–744. MIT Press, 2002.
- [185] Ronen Talmon and Ronald R. Coifman. Intrinsic modeling of stochastic dynamical systems using empirical geometry. *Applied and Computational Harmonic Analysis*, 39(1):138–160, jul 2015.
- [186] Xi Tan, Vinayak Rao, and Jennifer Neville. Nested CRP with hawkes-gaussian processes. In *AISTATS*, volume 84 of *Proceedings of Machine Learning Research*, pages 1289–1298. PMLR, 2018.
- [187] Naohiro Tawara, Tetsuji Ogawa, Shinji Watanabe, and Tetsunori Kobayashi. Nested gibbs sampling for mixture-of-mixture model and its application to speaker clustering. *APSIPA Transactions on Signal and Information Processing*, 5, 2016.
- [188] Yee Whye Teh, Michael I. Jordan, Matthew J. Beal, and David M. Blei. Sharing clusters among related groups: Hierarchical dirichlet processes. In *NIPS*, pages 1385–1392, 2004.
- [189] Lavanya Sita Tekumalla. Deep infinite admixture models. 2018.
- [190] Mark Transtrum. Information geometry for nonlinear least-squares data fitting and calculation of the superconducting superheating field. 2011.
- [191] Peter Trebatıcký. Recurrent neural network training with the extended kalman filter. In *IIT. SRC 2005: Student Research Conference*, page 57, 2005.
- [192] Michael Tsyrlunikov and Alexander Rakitko. A hierarchical bayes ensemble kalman filter. *Physica D: Nonlinear Phenomena*, 338:1–16, jan 2017.
- [193] Michael Nche Tuma and Reinhold Decker. Finite mixture models in market segmentation: A review and suggestions for best practices. 2013.
- [194] Ambrish Tyagi and James W. Davis. A recursive filter for linear systems on riemannian manifolds. In *CVPR*. IEEE Computer Society, 2008.
- [195] Florin Vaida and Suzette Blanchard. Conditional akaike information for mixed-effects models. *Biometrika*, 92(2):351–370, jun 2005.
- [196] Ramon van Handel. Filtering, stability, and robustness. 2007.
- [197] Pierre Vandekerckhove, Jagan Padbiri, and David McDowell. Bootstrap non-nested mixture model selection-application to extreme value modeling in metal fatigue problems. 2013.

- [198] D. Del Vecchio and R.M. Murray. Observability and local observer construction for unknown parameters in linearly and nonlinearly parameterized systems. In *Proceedings of the 2003 American Control Conference, 2003*. IEEE.
- [199] Jeroen K. Vermunt and Jay Magidson. Hierarchical mixture models for nested data structures. In *GfKl, Studies in Classification, Data Analysis, and Knowledge Organization*, pages 240–247. Springer, 2004.
- [200] D Vicente and Alfredo Vellido. Review of hierarchical models for data clustering and visualization. *Tendencias de la Minera de Datos en Espaa, Espaa de Minera de Datos*, 2004.
- [201] Davide Vidotto, Jeroen K. Vermunt, and Katrijn van Deun. Bayesian multilevel latent class models for the multiple imputation of nested categorical data. *Journal of Educational and Behavioral Statistics*, page 107699861876987, apr 2018.
- [202] Cinzia Viroli and Geoffrey J. McLachlan. Deep gaussian mixture models. *Statistics and Computing*, 29(1):43–51, 2019.
- [203] John MacLaren Walsh. A completed information projection interpretation of expectation propagation. 2007.
- [204] Hongling Wang, Chengjin Zhang, Yong Song, and Bao Pang. A perception-driven exploration hierarchical simultaneous localization and mapping for mobile robots adapted to search and rescue environments. *Advances in Mechanical Engineering*, 10(3):168781401876587, mar 2018.
- [205] Keith F. Widaman. Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, 9(1):1–26, mar 1985.
- [206] Ronald J Williams. *Some observations on the use of the extended Kalman filter as a recurrent network learning algorithm*. Citeseer, 1992.
- [207] Robert C. Wilson and Leif H. Finkel. A neural implementation of the kalman filter. In *NIPS*, pages 2062–2070. Curran Associates, Inc., 2009.
- [208] Y. Xie, J. Hu, S. Tang, and X. Huang. A forward-backward algorithm for nested hidden semi-markov model and application to network traffic. *The Computer Journal*, 56(2):229–238, sep 2012.
- [209] Rio Yokota, Huda Ibeid, and David E. Keyes. Fast multipole method as a matrix-free hierarchical low-rank approximation. *CoRR*, abs/1602.02244, 2016.
- [210] Liangyin Yu and Charles R. Dyer. Observer motion estimation and control from optical flow. In *ICIP (2)*, pages 941–944. IEEE, 2001.
- [211] Petr Zelinka and Milan Sigmund. Hierarchical classification tree modeling of nonstationary noise for robust speech recognition. 2015.
- [212] Jian Zeng, Dorian J. Garrick, Jack C. Dekkers, and Rohan L Fernando. A nested mixture model for genomic prediction using whole-genome SNP genotypes. Technical report, 2016.
- [213] Chi Zhang, Amirhossein Taghvaei, and Prashant G. Mehta. Feedback particle filter on riemannian manifolds and matrix lie groups. *IEEE Trans. Automat. Contr.*, 63(8):2465–2480, 2018.
- [214] He Zhao, Lan Du, Wray L. Buntine, and Mingyuan Zhou. Dirichlet belief networks for topic structure learning. In *NeurIPS*, pages 7966–7977, 2018.
- [215] Yunyun Zhao, Xiangke Wang, Yirui Cong, and Lincheng Shen. Information geometry-based action decision-making for target tracking by fixed-wing unmanned aerial vehicle. *International Journal of Advanced Robotic Systems*, 15(4):172988141878706, jul 2018.
- [216] Fengwei Zhong, Huafei Sun, and Zhenning Zhang. THE GEOMETRY OF THE DIRICHLET MANIFOLD. *Journal of the Korean Mathematical Society*, 45(3):859–870, may 2008.
- [217] Saman A. Zonouz and William H. Sanders. A kalman-based coordination for hierarchical state estimation: algorithm and analysis. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*. IEEE, jan 2008.