Bivariate Markov Processes and Their Estimation

Bivariate Markov Processes and Their Estimation

Yariv Ephraim

George Mason University
Fairfax, VA 22030
USA
yephraim@gmu.edu

Brian L. Mark

George Mason University
Fairfax, VA 22030
USA
bmark@gmu.edu

the essence of knowledge

Boston – Delft

Foundations and Trends® in Signal Processing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is Y. Ephraim and B. L. Mark, Bivariate Markov Processes and Their Estimation, Foundations and Trends® in Signal Processing, vol 6, no 1, pp 1–95, 2012

ISBN: 978-1-60198-674-0

Foundations and Trends® in Signal Processing
Volume 6 Issue 1, 2012

Editorial Scope

Foundations and Trends® in Signal Processing will publish survey and tutorial articles on the foundations, algorithms, methods, and applications of signal processing including the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital and multirate signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
  - classification and detection
  - estimation and regression
  - tree-structured methods

Information for Librarians
Foundations and Trends® in Signal Processing, 2012, Volume 6, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.

Foundations and Trends® in
Signal Processing
Vol. 6, No. 1 (2012) 1–95

DOI: 10.1561/2000000043

Bivariate Markov Processes and Their Estimation

Yariv Ephraim[1] and Brian L. Mark[2]

[1] *George Mason University, 4400 University Drive, Fairfax, VA 22030, USA, yephraim@gmu.edu*

[2] *George Mason University, 4400 University Drive, Fairfax, VA 22030, USA, bmark@gmu.edu*

Abstract

A bivariate Markov process comprises a pair of random processes which are *jointly* Markov. One of the two processes in that pair is observable while the other plays the role of an underlying process. We are interested in three classes of bivariate Markov processes. In the first and major class of interest, the underlying and observable processes are continuous-time with finite alphabet; in the second class, they are discrete-time with finite alphabet; and in the third class, the underlying process is continuous-time with uncountably infinite alphabet, and the observable process is continuous-time with countably or uncountably infinite alphabet. We refer to processes in the first two classes as bivariate Markov chains. Important examples of continuous-time bivariate Markov chains include the Markov modulated Poisson process, and the batch Markovian arrival process. A hidden Markov model with finite alphabet is an example of a discrete-time bivariate Markov chain. In the third class we have diffusion processes observed in Brownian motion, and diffusion processes modulating the rate of a Poisson process. Bivariate Markov processes play central roles in the theory and applications of estimation, control, queuing, biomedical engineering, and reliability. We review properties of bivariate Markov processes, recursive estimation of their statistics, and recursive and iterative parameter estimation.

Contents

1

Introduction

A bivariate Markov process comprises a pair of random processes which are *jointly Markov*. One of the two processes is observable, while the other plays the role of an underlying process. The underlying process affects the statistical properties of the observable process. Usually, the observable process is not Markov, but the underlying process is often conveniently chosen to be Markov. The theory of bivariate Markov processes does not require either process to be Markov.

The family of bivariate Markov processes is very rich, and has produced powerful models in many applications. Perhaps the most familiar bivariate Markov process stems from the hidden Markov model, see, e.g., [13, 40]. The underlying process of a hidden Markov model is a discrete-time finite-state Markov chain, and the observable process comprises a collection of conditionally independent random variables, e.g., normal, given the underlying Markov chain. Together, the two processes form a bivariate Markov process. Another example follows from the Markov modulated Poisson process, see, e.g., [43, 78, 93]. Here, the underlying process is a continuous-time finite-state Markov chain, and the observable process is conditionally Poisson given the underlying Markov chain. A generalization of this process is given by a Poisson process whose rate is modulated by an underlying diffusion process, see, e.g., [16, 103, 120]. As a final example, we mention the bivariate Markov process formed by an underlying diffusion process, and the same process observed in Brownian motion. Here the underlying process is a continuous-time continuous-alphabet Markov process. Bivariate Markov processes play central roles in the theory and applications of estimation, control, queuing, economics, biomedical engineering, and reliability.

In general, each of the two process components of a bivariate Markov process may be discrete-time or continuous-time, with finite, countably infinite, or uncountably infinite alphabet. We shall focus on three classes of bivariate Markov processes. In the first class, the pair of processes comprising the bivariate Markov process are continuous-time with finite alphabet; in the second class, they are discrete-time with finite alphabet; and in the third class, both processes are continuous-time with a diffusion underlying process and an observable process with a countably or uncountably infinite alphabet. Our primary focus in this paper will be on the first class of processes, which we refer to as *continuous-time bivariate Markov chains* or simply as *bivariate Markov chains*. We shall refer to processes from the second class as *discrete-time bivariate Markov chains*. The processes in the third class are assumed to be diffusion processes observed in Brownian motion, in a counting process, or in a mixture of Brownian motion and a counting process. Some of the results reported here for finite alphabet processes, apply to bivariate Markov processes with countably infinite alphabet, by resorting to modulo arithmetic.

The theory of *univariate* Markov processes applies to bivariate Markov processes. Excellent sources for that theory may be found in Doob [31], Breiman [15] and Todorovic [109]. Application of the theory of univariate Markov processes to bivariate Markov processes, with the observable and underlying processes playing different roles, is not straightforward. Research on various forms of bivariate Markov processes has been ongoing for more than four decades. The research has focused on two main interrelated estimation problems, namely, parameter and signal estimation. In parameter estimation, the maximum likelihood approach has dominated the field. Here, identifiability of the parameter of the bivariate Markov process was studied; iterative estimation approaches in the form of the expectation-maximization (EM) algorithm were developed; and consistency and asymptotic normality were proven for parameter estimation of some bivariate Markov chains. Application of the EM approach requires minimum mean square error recursive estimation of several statistics of the bivariate Markov chain. In particular, estimation of the number of jumps from one state to another, and the total sojourn time of the process in each state, in a given interval, are required. In other applications, estimation of the state of the underlying process is of primary interest.

In this paper we present some of the fundamentals of the theory of bivariate Markov processes, and review the various parameter and signal estimation approaches. Our goals are to provide a comprehensive introduction to bivariate Markov chains, along with the details of the various estimation algorithms. While proofs are generally omitted, an interested reader should be able to implement the estimation algorithms for bivariate Markov chains straight out of this paper. Most of the material in this paper should be accessible to the signal processing community. It requires some familiarity with Markov chains and the intricacies of the theory of hidden Markov models. The discussion on diffusion processes requires some further knowledge in nonlinear estimation theory.

Our presentation in Sections 1 to 7 focuses on continuous-time bivariate Markov chains with a finite or countably infinite number of states. In Section 8 we discuss finite alphabet discrete-time bivariate Markov chains. In Section 9 we consider a bivariate Markov chain observed through Brownian motion. In Section 10, we provide a glimpse into the fascinating topic of bivariate Markov processes with underlying diffusion processes. Several applications are discussed in Section 11, and some concluding remarks are given in Section 12.

References

[1] A. Albert, "Estimating the infinitesimal generator of a continuous time, finite state Markov process," *Ann. Math. Stat.*, vol. 33, pp. 727–753, 1962.

[2] S. Asmussen, *Applied Probability and Queues.* New York: Springer-Verlag, 2nd Edition, 2003.

[3] S. Asmussen, O. Nerman, and M. Olsson, "Fitting phase-type distributions via the EM algorithm," *Scand. J. Stat.*, vol. 23, no. 4, pp. 419–441, 1996.

[4] A. Bain and D. Crisan, *Fundamentals of Stochastic Filtering.* New York: Springer, 2009.

[5] F. Ball, R. K. Milne, and G. F. Yeo, "Continuous-time Markov chains in a random environment, with applications to ion channel modelling," *Adv. Appl. Probab.*, vol. 26, pp. 919–946, 1994.

[6] F. Ball and G. F. Yeo, "Lumpability and marginalisability for continuous time Markov chains," *J. Appl. Prob.*, vol. 30, no. 3, pp. 518–528, 1993.

[7] F. G. Ball and J. A. Rice, "Stochastic models for ion channels: Introduction and bibliography," *Math. Biosci.*, vol. 112, pp. 189–206, 1992.

[8] F. G. Ball and M. S. P. Sansom, "Ion-channel gating mechanisms: model identification and parameter estimation from single channel recordings," *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 236, no. 1285, pp. 385–416, 1989.

[9] V. S. Barbu and N. Limnios, *Semi-Markov Chains and Hidden Semi-Markov Models Towards Applications.* New York: Springer, 2008.

[10] L. E. Baum, "An inequality and associated maximization technique in statistical estimation of probabilistic functions of Markov processes," *Inequalities*, vol. 3, no. 1, pp. 1–8, 1972.

[11] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *Ann. Math. Statist.*, vol. 37, pp. 1554–1563, 1966.

[12] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *Ann. Math. Statist.*, vol. 41, pp. 164–171, 1970.

[13] P. J. Bickel, Y. Ritov, and T. Rydén, "Asymptotic normality of the maximum-likelihood estimator for general hidden Markov models," *Ann. of Statist*, vol. 26, no. 4, pp. 1614–1635, 1998.

[14] P. Billingsley, *Convergence of Probability Measures*. New York: John-Wiley & Sons, 1968.

[15] L. Breiman, *Probability*. SIAM, Philadelphia, 1992.

[16] P. Brémaud, *Point Processes and Queues: Martingale Dynamics*. New York: Springer-Verlag, 1981.

[17] L. Breuer, "An EM algorithm for batch Markovian arrival processes and its comparison to a simpler estimation procedure," *Ann. Oper. Res.*, vol. 112, no. 1–4, pp. 123–138, 2002.

[18] D. Bryant, N. Galtier, and M. A. Poursat, "Likelihood calculation in molecular phylogenetics," in *Mathematics of Evolution and Phylogeny*, (O. Gascuel, ed.), New York: Oxford, 2007.

[19] P. Bryant and J. A. Williamson, "Asymptotic behaviour of classification maximum likelihood estimates," *Biometrika*, vol. 65, no. 2, pp. 273–281, 1978.

[20] A. Budhiraja, L. Chen, and C. Lee, "A survey of numerical methods for nonlinear filtering problems," *Physica D*, vol. 230, pp. 27–36, 2007.

[21] O. Cappé, E. Moulines, and T. Rydén, *Inference in Hidden Markov Models*. New York: Springer, 2005.

[22] G. Casale, E. Zhang, and E. Smirni, "Trace data characterization and fitting for Markov modeling," *Performance Evaluation*, vol. 67, pp. 61–79, 2010.

[23] B. Chen and P. A. Zadrozny, "Analytic derivatives of the matrix exponential for estimation of linear continuous-time models," *Journal of Economic Dynamics & Control*, vol. 25, pp. 1867–1879, 2001.

[24] S. H. Chung, V. Krishnamurthy, and J. B. Moore, "Adaptive processing techniques based on hidden Markov models for characterizing very small channel currents buried in noise and deterministic interference," *Phil. Trans. R. Soc. Lond. B*, vol. 334, pp. 357–384, 1991.

[25] E. Çinlar, "Markov renewal theory: A survey," *Management Science*, vol. 21, no. 7, pp. 727–752, 1975.

[26] J. M. C. Clark, "The design of robust approximations to the stochastic differential equations of nonlinear filtering," in *Communication Systems and Random Processes Theory*, (J. K. Skwirzynski, ed.), Darlington, 1977. Alphen aan den Rijn, The Netherlands: Sijthoff and Noordhoff, 1978; 721–734.

[27] D. Colquhoun and A. G. Hawkes, "On the stochastic properties of bursts of single ion channel openings and of clusters of bursts," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 300, no. 1098, pp. 1–59, 1982.

[28] D. R. Cox, "Some statistical methods connected with series of events," *Journal of the Royal Statistical Society*, Series B, vol. 17, no. 2, pp. 129–164, 1955.

[29] A. Dembo and O. Zeitouni, "Parameter estimation of partially observed continuous time stochastic processes via the EM algorithm," *Stochastic Processes and their Applications*, vol. 23, pp. 91–113, 1986.

[30] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc. B*, vol. 39, pp. 1–38, 1977.

[31] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1990.

[32] R. Douc, E. Moulines, J. Olsson, and R. Van Handel, "Consistency of the maximum likelihood estimator for general hidden Markov models," *Ann. Statist.*, vol. 39, pp. 474–513, 2011.

[33] R. Douc, E. Moulines, and T. Rydén, "Asymptotic properties of the maximum likelihood estimator in autoregressive models with Markov regime," *Ann. Statist.*, vol. 32, pp. 2254–2304, 2004.

[34] T. E. Duncan, "Evaluation of likelihood functions," *Information and Control*, vol. 13, pp. 62–74, 1968.

[35] G. Elfving, "Zur theorie der markoffschen ketten," *Acta Social Science Fennicaen.*, vol. series A.2, no. 8, pp. 1–17, 1937.

[36] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*. New York: Springer-Verlag, 1997.

[37] Y. Ephraim and B. L. Mark, "Explicit forward recursive estimators for Markov modulated Markov processes," *Stoch. Models*, vol. 28, no. 3, pp. 359–387, 2012.

[38] Y. Ephraim and B. L. Mark, "On forward recursive estimation for bivariate Markov chains," in *Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, March 2012.

[39] Y. Ephraim and B. L. Mark, "Consistency of maximum likelihood parameter estimation for bivariate Markov chains," *Stoch. Models*, vol. 29, no. 1, pp. 89–111, 2013.

[40] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1518–1569, 2002.

[41] Y. Ephraim and W. J. J. Roberts, "An EM algorithm for Markov modulated Markov processes," *IEEE Trans. Signal Process.*, vol. 57, pp. 463–470, 2009.

[42] J. D. Ferguson, ed., *Proc. of the Symposium on the Applications of Hidden Markov Models to Text and Speech*. IDA-CRD, Princeton, NJ, 1980.

[43] W. Fischer and K. Meier-Hellstern, "The Markov-modulated Poisson process (MMPP) cookbook," *Perform. Eval.*, vol. 18, pp. 149–171, 1992.

[44] D. R. Fredkin and J. A. Rice, "On aggregated Markov processes," *J. Appl. Prob.*, vol. 23, no. 1, pp. 208–214, 1986.

[45] D. R. Fredkin and J. A. Rice, "Bayesian restoration of single-channel patch clamp recordings," *Biometrics*, vol. 48, no. 2, pp. 427–448, 1992.

[46] D. R. Fredkin and J. A. Rice, "Maximum likelihood estimation and identification from single-channel recordings," *Proc. Royal. Soc. London B*, vol. 249, no. 1325, pp. 125–132, 1992.

[47] R. Frey, T. Schmidt, and L. Xu, "On Galerkin approximations for the Zakai equation with diffusive and point process observations," Preprint available at http://www.math.uni-leipzig.de/∼frey/publications-frey.html. Also in arXiv:1303.0975, 2011.

[48] N. Galtier and A. Jean-Marie, "Markov-modulated Markov processes and the covarion process of molecular evolution," *J. Computational Biology*, vol. 11, no. 4, pp. 727–733, 2004.
[49] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in the time domain: Modeling and exploiting white space," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 66–72, May 2007.
[50] A. Germani and M. Piccioni, "A Galerkin approximation for the Zakai equation," *Lecture Notes in Control and Information Sciences*, vol. 59, pp. 415–423, 1984.
[51] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes.* Oxford: Oxford Science Publications, 2001.
[52] R. Gusella, "Characterizing the variability of arrival processes with indexes of dispersion," *IEEE J. Selected Areas Comm.*, vol. 9, pp. 203–211, 1991.
[53] U. Holst and G. Lindgren, "Recursive estimation in mixture models with Markov regime," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1683–1690, Nov. 1991.
[54] U. Holst, G. Lindgren, J. Holst, and M. Thuvesholmen, "Recursive estimation in switching autoregressive with a Markov regime," *J. Time Ser. Anal*, vol. 15, no. 5, pp. 489–506, 1994.
[55] R. A. Horn and C. R. Johnson, *Matrix Analysis.* New York: Cambridge University Press, 2013.
[56] H. Ito, S.-I. Amari, and K. Kobayashi, "Identifiability of Hidden Markov information sources and their minimum degrees of freedom," *IEEE Trans. on Inform. Theory*, vol. 38, pp. 324–333, Mar. 1992.
[57] M. R. James, V. Krishnamurthy, and F. Le Gland, "Time discretization of continuous-time filters and smoothers for HMM parameter estimation," *IEEE Trans. Inform. Theory*, vol. 42, pp. 593–605, Mar. 1996.
[58] J. F. C. Kingman, "The imbedding problem for finite Markov Chains," *Zeitschrift fur Wahrscheinlichkeitstheorie*, vol. 1, pp. 14–24, 1962.
[59] F. C. Klebaner, *Introduction to Stochastic Calculus with Applications.* London: Imperial College Press, 2006.
[60] A. Klemm, C. Lindemann, and M. Lohmann, "Modeling IP traffic using the batch Markovian arrival process," *Performance Evaluation*, vol. 54, pp. 149–173, 2003.
[61] T. G. Kurtz and D. L. Ocone, "Unique characterization of conditional distributions in nonlinear filtering," *The Annals of Probability*, vol. 16, no. 1, pp. 80–107, 1988.
[62] H. J. Kushner, "Dynamical equations for optimal nonlinear filtering," *Journal of Differential Equations*, vol. 3, no. 2, pp. 179–190, 1967.
[63] P. Lancaster and M. Tismenetsky, *The Theory of Matrices.* New York: Academic Press, 1985.
[64] B. Larget, "A canonical representation for aggregated Markov processes," *J. Appl. Prob.*, vol. 35, pp. 313–324, 1998.
[65] G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling.* American Statistical Association (ASA) and Society for Industrial and Applied Mathematics (SIAM), 1999.

[66] B. G. Leroux, "Maximum likelihood estimation for hidden Markov models," *Stochastic Processes and Their Applications*, vol. 40, pp. 127–143, 1992.
[67] S. Levinson, "Continuously variable duration hidden Markov models for automatic speech recognition," *Computer, Speech and Language*, vol. 1, pp. 29–45, 1986.
[68] Z. Li, Y. Ephraim, and B. L. Mark, "Recursive parameter estimation for bivariate Markov chains," in preparation.
[69] G. Lindgren and U. Holst, "Recursive estimation of parameters in Markov-modulated Poisson processes," *IEEE Trans. Commun.*, vol. 43, no. 11, pp. 2812–2820, Nov. 1995.
[70] R. S. Liptser and A. N. Shiryaev, *Statistics of Random Processes, Part I.* New York: Springer-Verlag, 2001.
[71] D. M. Lucantoni, "New results on the single server queue with a batch Markovian arrival process," *Stochastic Models*, vol. 7, no. 1, pp. 1–46, 1991.
[72] D. M. Lucantoni, "The BMAP/G/1 queue: A tutorial," in *Models and techniques for Performance Evaluation of Computer and Communications Systems*, (L. Donatiello and R. Nelson, eds.), pp. 330–358, Springer Verlag, 1993.
[73] D. G. Luenberger, *Linear and Nonlinear Programming.* Reading, Massachusetts: Addison-Wesley, 1984.
[74] W. P. Malcolm, R. J. Elliott, and J. Van Der Hoek, "A deterministic discretisation-step upper bound for state estimation via Clark transformations," *Journal of Applied Mathematics and Stochastic Analysis*, vol. 4, pp. 371–384, 2004.
[75] B. L. Mark and Y. Ephraim, "Explicit forward recursive estimators for continuous-time bivariate Markov chains," in preparation.
[76] B. L. Mark and Y. Ephraim, "An EM algorithm for continuous-time bivariate Markov chains," *Computational Statistics and Data Analysis*, vol. 57, pp. 504–517, 2013.
[77] B. L. Mark and A. O. Nasif, "Estimation of maximum interference-free transmit power level for opportunistic spectrum access," *IEEE Trans. on Wireless Communications*, vol. 8, no. 5, pp. 2505–2513, May 2009.
[78] K. S. Meier, "A statistical procedure for fitting Markov-modulated Poisson processes," Ph.D. Dissertation, University of Delaware, 1984.
[79] C. Moler and C. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty five years later," *SIAM Review*, vol. 45, pp. 3–49, 2003.
[80] R. E. Mortensen, "Optimal control of continuous-time stochastic systems," Ph.D. dissertation, University of California, Berkeley, 1966.
[81] M. F. Neuts, "A versatile Markovian point process," *J. Appl. Prob.*, vol. 16, pp. 764–779, 1979.
[82] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models.* Baltimore: Johns Hopkins University Press, 1981.
[83] T. Nguyen, B. L. Mark, and Y. Ephraim, "Spectrum sensing using a hidden bivariate Markov model," submitted for publication.
[84] T. Nguyen, B. L. Mark, and Y. Ephraim, "Hidden Markov process based dynamic spectrum access for cognitive radio," in *Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, March 2011.

[85] L. Pachter and B. Sturmfels, "The mathematics of phylogenomics," *SIAM Review*, vol. 49, no. 1, pp. 3–31, Mar. 2007.
[86] T. Petrie, "Probabilistic functions of finite state Markov chains," *Ann. Math. Statist.*, vol. 40, pp. 97–115, 1969.
[87] F. Qin, A. Auerbach, and F. Sachs, "Maximum likelihood estimation of aggregated Markov processes," *Proc. Royal Society Lond. B*, vol. 264, pp. 375–383, 1997.
[88] N. I. Ramesh, "Statistical analysis on Markov-modulated Poisson processes," *Environmetrics*, vol. 6, no. 2, pp. 165–179, March/April 1995.
[89] W. J. J. Roberts and Y. Ephraim, "An EM algorithm for ion-channel current estimation," *IEEE Trans. Sig. Proc.*, vol. 56, no. 1, pp. 26–33, 2008.
[90] W. J. J. Roberts, Y. Ephraim, and E. Dieguez, "On Rydén's EM algorithm for estimating MMPPs," *IEEE Sig. Proc. Let.*, vol. 13, pp. 373–377, 2006.
[91] M. H. Rossiter, "Characterizing a random point process by a switched Poisson process," Ph.D. Dissertation, Monash University, 1989.
[92] M. Rudemo, "State estimation for partially observed Markov chains," *J. Math. Analysis and Applications*, vol. 44, pp. 581–611, 1973.
[93] T. Rydén, "Parameter estimation for Markov modulated Poisson processes," *Commun. Statist. Stoch. Models*, vol. 10, no. 4, pp. 795–829, 1994.
[94] T. Rydén, "Consistent and asymptotically normal parameter estimates for Markov modulated Poisson processes," *Scand. J. Statist.*, vol. 22, pp. 295–303, 1995.
[95] T. Rydén, "An EM algorithm for estimation in Markov-modulated Poisson processes," *Computational Statistics and Data Analysis*, vol. 21, pp. 431–447, 1996.
[96] T. Rydén, "On identifiability and order of continuous-time aggregated Markov chains, Markov-modulated Poisson processes, and phase-type distributions," *J. Appl. Prob.*, vol. 33, pp. 640–653, Sept. 1996.
[97] T. Rydén, "On recursive estimation for hidden Markov models," *Stochastic Processes and their Applications*, vol. 66, pp. 79–96, 1997.
[98] T. Rydén, "Asymptotically efficient recursive estimation for incomplete data models using the observed information," *Metrika*, vol. 47, pp. 119–145, 1998.
[99] K. Salamatian and S. Vaton, "Hidden Markov modelling for network communication channels," in *Proc. ACM Sigmetrics*, pp. 92–101, June 2001.
[100] A. Segall and T. Kailath, "The modeling of randomly modulated jump processes," *IEEE Trans. Inform. Theory*, vol. 21, pp. 135–143, 1975.
[101] A. Segall and T. Kailath, "Nonlinear filtering with counting observations," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 2, pp. 143–149, 1975.
[102] B. Singer and S. Spilerman, "The representation of social processes by Markov models," *The American Journal of Sociology*, vol. 82, no. 1, pp. 1–54, 1976.
[103] D. L. Snyder, "Filtering and detection for doubly stochastic Poisson processes," *IEEE Trans. Information Theory*, vol. IT-18, no. 1, pp. 91–102, Jan. 1972.
[104] D. L. Snyder and M. I. Miller, *Random Point Processes in Time and Space*. New York: Springer-Verlag, 1991.
[105] C. Stiller and G. Radons, "Online estimation of hidden Markov models," *IEEE Sig. Proc. Let.*, vol. 6, no. 8, pp. 213–215, 1999.

[106] Z. Sun, G. J. Bradford, and J. N. Laneman, "Sequence detection algorithms for PHY-layer sensing in dynamic spectrum access networks," *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 1, pp. 97–109, Feb. 2011.

[107] D. M. Titterington, "Comments on 'Application of the conditional population-mixture model to image segmentation'," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 5, pp. 656–658, Sept. 1984.

[108] D. M. Titterington, "Recursive parameter estimation using incomplete data," *J. R. Statist. Soc. B*, vol. 46, no. 2, pp. 257–267, 1984.

[109] P. Todorovic, *An Introduction to Stochastic Processes and Their Applications*. New York: Springer-Verlag, 1994.

[110] C. F. Van Loan, "Computing integrals involving the matrix exponential," *IEEE Trans. Automat. Control*, vol. 23, pp. 395–404, 1978.

[111] J. H. van Schuppen, "Filtering, prediction and smoothing for counting process observations, a martingale approach," *SIAM J. Appl. Math*, vol. 32, no. 3, pp. 552–570, 1977.

[112] W. J. Vetter, "Derivative operations on matrices," *IEEE Trans. Automatic Control*, vol. AC-15, pp. 241–244, April 1970.

[113] S. Wang and Y. Zhao, "Almost sure convergence of Titterington's recursive estimator for mixture models," *Statistics & Probability Letters*, vol. 76, pp. 2001–2006, 2006.

[114] W. Wei, B. Wang, and D. Towsley, "Continuous-time hidden Markov models for network performance evaluation," *Performance Evaluation*, vol. 49, pp. 129–146, Sep. 2002.

[115] E. Weinstein, M. Feder, and A. V. Oppenheim, "Sequential algorithms for parameter estimation based on the Kullback-Leibler information measure," *IEEE Trans. Acoustic., Speech and Signal Processing*, vol. 38, no. 9, pp. 1652–1654, Sept. 1990.

[116] L. R. Welch, "Hidden Markov models and the Baum-Welch algorithm," *IEEE Information Theory Society Newsletter*, vol. 53, no. 4, Dec. 2003.

[117] C. J. Willy, W. J. J. Roberts, T. A. Mazzuchi, and S. Sarkani, "On recursive MMPP parameter estimation," *IEEE Sig. Proc. Let.*, vol. 15, pp. 883–886, 2008.

[118] C. J. Willy, W. J. J. Roberts, T. A. Mazzuchi, and S. Sarkani, "Recursions for the MMPP score vector and observed information matrix," *Stoch. Models*, vol. 26, no. 4, pp. 649–665, 2010.

[119] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. New Jersey: Prentice-Hall, 1989.

[120] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*. New York: Springer-Verlag, 1984.

[121] W. M. Wonham, "Some applications of stochastic differential equations to optimal non-linear filtering," *SIAM J. Contr., Series A*, vol. 2, pp. 347–369, 1965.

[122] C. F. J. Wu, "On the convergence properties of the *EM* algorithm," *Ann. Statist.*, vol. 11, no. 1, pp. 95–103, 1983.

[123] A. I. Yashin, "Filtering of jump processes," *Automat. Remote Control*, vol. 31, pp. 725–730, 1970.

[124] S. Z. Yu, "Hidden semi-Markov models," *Artificial Intelligence*, vol. 174, pp. 215–243, 2010.

[125] M. Zakai, "On the optimal filtering of diffusion processes," *Z. Wahrsch Verw. Gebiete.*, vol. 11, pp. 230–243, 1969.

[126] M. Zakai, "Stochastic processes in communication and control theory," unpublished notes, Technion-Israel Institute of Technology, Haifa, 1983.

[127] M. Zakai, "The optimal filtering of Markov jump processes in additive white noise," Research Note No. 563, Sylvania Electronic System, 1965. Also in arXiv:0904.2888, 2009.

[128] O. Zeitouni and A. Dembo, "Exact filters for the estimation of the number of transitions of finite-state continuous-time Markov processes," *IEEE Trans. Inform. Theory*, vol. 34, pp. 890–893, 1988.