Journal of Statistical Modelling: Theory and Applications
Vol. 7, No. 1, 2026, pp. 1-19
Yazd University Press 2026

JSMTA

Research Paper

Bayesian multiple change-point detection: A methodological comparison of exact dynamic programming and nonparametric hidden Markov approaches

Masoud Majidizadeh*
Department of Statistics, Faculty of Mathematical Sciences,
Shahid Beheshti University, Tehran, Iran

Received: October 16, 2025/ Revised: August 04, 2026/ Accepted: September 08, 2026

Abstract: Detecting structural breaks is crucial for understanding time series dynamics, particularly in highly volatile financial markets. This study systematically compares two prominent Bayesian change-point methodologies: an exact dynamic programming approach that eliminates Markov chain Monte Carlo sampling, and a Dirichlet process hidden Markov model that flexibly infers latent regime counts and persistence. Analyzing daily S&P 500 returns spanning 2000 to 2024-a period encompassing major upheavals like the global financial crisis and the COVID-19 pandemic-we evaluate both models across detection accuracy, parameter estimation, computational efficiency, and predictive performance. Despite fundamentally different assumptions, the models exhibit striking concordance, identifying major macroeconomic shifts with timing differences of mere days and parameter variations of under 0.01%. However, a clear trade-off emerges: the exact method is approximately 90 times faster, whereas the Dirichlet process hidden Markov model provides marginally superior predictive accuracy. These insights offer actionable guidelines for selecting the optimal framework based on computational and predictive priorities.

Keywords: Change-point detection; Dirichlet process; Dynamic programming.
Mathematics Subject Classification (2010): 62F15, 62M10, 60J20.

1 Introduction

Change-point detection is a fundamental problem in statistical time series analysis, concerned with identifying abrupt structural shifts in the data-generating mechanism.

*Corresponding author: `m_majidizadeh@sbu.ac.ir`

Such regime changes are pervasive across numerous disciplines: in financial markets, they manifest as volatility transitions during economic crises (Bai and Perron, 1998; Inclán and Tiao, 1994); in climate science, as sudden shifts in temperature or precipitation patterns (Perreault et al., 2000); in genomics, as breakpoints demarcating functional regions (Barry and Hartigan, 1993); and in social sciences, as policy-induced structural breaks (Western and Kleykamp, 2004). Accurately pinpointing these change-points is essential for deciphering system dynamics, enhancing the accuracy of forecasting models, and informing robust decision-making under uncertainty (Horváth and Rice, 2024).

The methodological landscape for change-point detection is vast, encompassing both frequentist and Bayesian paradigms. Classical approaches, such as cumulative sum tests (Inclán and Tiao, 1994) and least-squares segmentation (Bai and Perron, 1998; Lavielle and Moulines, 2000), have been extensively applied but often rely on restrictive distributional assumptions and offer limited capacity for uncertainty quantification. Recent methodological advances have introduced more sophisticated frameworks, including efficient penalized likelihood methods (Killick et al., 2012) and comprehensive asymptotic theories for multiple change-points (Csörgő and Horváth, 1997), with ongoing comparative studies continually evaluating their performance across diverse data generating processes (Shi et al., 2022).

Bayesian methods offer a compelling suite of advantages for tackling change-point problems. They provide a principled framework for full uncertainty quantification, yielding posterior distributions over both the number and precise locations of change-points, rather than mere point estimates. This approach naturally incorporates prior knowledge regarding plausible regime structures, which can significantly enhance detection power, especially in finite samples. Furthermore, Bayesian model comparison via marginal likelihoods or Bayes factors allows for rigorous assessment of competing change-point specifications (Green, 1995; Chib, 1998).

While early Bayesian work focused on single change-point models with conjugate priors (Carlin et al., 1992; Barry and Hartigan, 1993), modern applications typically involve multiple unknown change-points, necessitating advanced computational strategies. Seminal contributions include retrospective multiple change-point identification using Gibbs sampling (Stephens, 1994) and the introduction of reversible jump Markov chain Monte Carlo (MCMC, Green, 1995) to navigate the variable-dimensional parameter spaces inherent to these problems. Subsequent research has expanded into product partition models (Loschi and Cruz, 2005), hierarchical Bayesian frameworks (Lavielle and Lebarbier, 2001), online detection algorithms (Adams and MacKay, 2007), and generalized multiple change-point detection methodologies (Kang et al., 2025), with recent innovations extending to complex epidemiological modeling, such as pandemic dynamics (Majidizadeh, 2024; Majidizadeh and Taheriyoun, 2024).

Two particularly influential and philosophically distinct Bayesian approaches have gained prominence, each presenting a unique trade-off between computational efficiency and modeling flexibility. Fearnhead (2006) developed an exact Bayesian method based on dynamic programming that eschews MCMC sampling entirely, computing exact posterior distributions over change-point configurations in linear time. This method, implemented in widely-used software (Erdman and Emerson, 2007), is exceptionally effective for rapid, exact inference in large-scale datasets. In contrast, Ko et al. (2015)

proposed a Dirichlet process hidden Markov model (DP-HMM) that leverages nonparametric Bayesian flexibility to automatically determine the number of latent regimes while explicitly modeling their persistence through Markovian dynamics. This framework has been further extended to incorporate covariate effects (Shohoudi et al., 2022) as well as efficient, exact inference algorithms tailored to DP-HMM multiple change-point structures (Majidizadeh, 2026).

Despite the individual strengths and widespread adoption of these two approaches, a systematic comparison of their relative performance and practical trade-offs remains underexplored, particularly within the challenging context of financial time series. Financial markets present an ideal proving ground: they exhibit clear structural breaks aligned with economic crises (Giordani and Kohn, 2008), display pronounced volatility clustering and heavy-tailed distributions that challenge simple normality and independence assumptions (Chib, 1998; Geweke and Jiang, 2011), generate massive high-frequency datasets requiring scalable algorithms, and offer economically interpretable regimes that serve as natural validation benchmarks beyond purely statistical metrics.

This paper conducts a comprehensive empirical comparison of Fearnhead's exact dynamic programming approach and the DP-HMM framework, applied to daily log returns of the S&P 500 index spanning the 25-year period from 2000 to 2024. This epoch encapsulates a sequence of major financial upheavals—the dot-com bubble burst, the Global Financial Crisis, the European Debt Crisis, the COVID-19 pandemic crash, and the post-pandemic inflationary cycle—providing a rich tapestry of market regimes for analysis. Our study is structured around a multi-dimensional comparative evaluation encompassing: (1) the accuracy and temporal precision of detected change-points, validated against historical events via a unified assessment framework; (2) the economic interpretation of estimated regime parameters; (3) computational efficiency and algorithmic scalability; (4) the robustness of inference, supported by comprehensive MCMC convergence diagnostics (including $\hat{R}$ and effective sample size) and hyperparameter sensitivity analysis; and (5) out-of-sample predictive performance.

The contributions of this work are multifaceted. First, we provide the first systematic, empirical juxtaposition of these two influential Bayesian methodologies on a long-span, real-world financial dataset, extending beyond theoretical or limited simulation-based comparisons. Second, we introduce a controlled simulation experiment to rigorously evaluate model sensitivity to subtle, non-obvious structural breaks. Third, we document a remarkable convergence in the core findings between the methods despite their fundamentally different mathematical foundations. Fourth, we precisely quantify the inherent trade-off: Fearnhead's method achieves computational speed approximately two orders of magnitude faster but assumes within-segment independence, whereas the DP-HMM provides greater modeling flexibility and accommodates complex dynamics at a substantial computational cost. Finally, we incorporate an examination of recent methodological advances (Liu et al., 2023; Maheu and Song, 2018) to derive practical, actionable guidance for financial practitioners on method selection.

The remainder of this paper is organized as follows. Section 2 reviews the theoretical foundations of both methods. Section 3 describes the S&P 500 data and preprocessing steps. Section 4 presents the new controlled simulation study evaluating model performance. Section 5 details the empirical application, including MCMC diagnostics, sensitivity analysis, and the combined event validation. Section 6 provides a synthesized discussion and comparative evaluation, merging methodological insights with empirical findings. Section 7 outlines practical implications. Section 8 concludes and suggests avenues for future research.

2 Methodological framework

To address the problem of multiple change-point detection, various foundational frameworks have been developed and continuously extended in recent literature (Fisch et al., 2019; Jewell et al., 2022; Madrid Padilla et al., 2021). In this section, we detail two prominent Bayesian methodologies: the exact Bayesian approach proposed by Fearnhead (2006), and the nonparametric DP-HMM introduced by Ko et al. (2015). By reviewing these frameworks, we establish the theoretical foundation necessary for comparing their performance in detecting both distinct and subtle regime shifts.

2.1 Exact Bayesian inference

The exact Bayesian approach of Fearnhead (2006) models a univariate time series $\{y_t\}_{t=1}^n$ under the assumption that the data are generated from a piecewise constant process with an unknown number of change-points, denoted by K. The locations of these change-points are represented by the vector $\tau = (\tau_1, \ldots, \tau_K)$, where $0 = \tau_0 < \tau_1 < \cdots < \tau_K < \tau_{K+1} = n$. Consequently, the data are partitioned into $K + 1$ segments. Within any given segment j (i.e., for $\tau_{j-1} < t \leq \tau_j$), the observations are assumed to be independent and identically distributed according to a probability density $f(y_t \mid \theta_j)$, where θ_j represents the parameter vector specific to that segment. To accurately capture the characteristics of real-world data, which often exhibit heavy tails, the emission distribution $f(\cdot \mid \theta_j)$ does not need to be strictly Gaussian; it can be specified using robust symmetric distributions, such as the Student-t distribution, provided that suitable priors are established.

The model relies on a fully Bayesian specification. A Poisson prior, $K \sim \text{Poisson}(\lambda)$, is typically assigned to the number of change-points, where λ dictates the expected number of shifts. Conditioned on K, the locations of the change-points are assumed to be uniformly distributed across all valid configurations, yielding $p(\tau \mid K) = \binom{n-1}{K}^{-1}$. Furthermore, conjugate priors $\pi(\theta_j)$ are assigned to the segment parameters to ensure mathematical tractability.

The primary innovation of Fearnhead's approach is a two-stage procedure that avoids traditional MCMC sampling. First, a backward filtering recursion is used to efficiently compute summary statistics. Let $Q(t)$ be the probability of observing the data from time t to n, given a change-point at $t-1$. Let $P(t, s) = \int \prod_{i=t}^{s} f(y_i \mid \theta)\pi(\theta)\, d\theta$ denote the marginal likelihood of the segment $(t, s]$. This quantity can be calculated recursively starting from $Q(n + 1) = 1$:

$$Q(t) = \sum_{s=t}^{n-1} [P(t, s)g(s - t + 1)Q(s + 1)] + P(t, n)(1 - G(n - t)),$$

where $g(\cdot)$ is the prior probability mass function on segment length, and $G(m) = \sum_{k=1}^{m} g(k)$ is its cumulative distribution function. The final term $P(t, n)(1 - G(n - t))$

explicitly accounts for the right-censoring of the final segment at the end of the time series n. This recursion runs in $O(n^2)$ time. Although not shown in the simplified formula for brevity, the prior on the number of change-points is incorporated into this step. Once all $Q(t)$ values are computed, the second stage involves a forward sampling pass to generate exact samples of change-point locations from their posterior distribution. Fearnhead also introduces a probabilistic pruning strategy to reduce the computational burden to nearly linear time for many practical scenarios.

Upon completing the recursion, comprehensive inference can be performed. The posterior distribution of the number of change-points is obtained by marginalizing over all configurations, $p(K \mid y_{1:n}) \propto \sum_{\tau:|\tau|=K} P(\tau, K)$, and the maximum a posteriori (MAP) estimate of the change-point locations can be efficiently extracted by backtracking through the computed probabilities.

The exact Bayesian approach offers significant methodological advantages, primarily its ability to provide full posterior distributions without MCMC-induced sampling errors, making it highly computationally efficient for moderate-sized time series. However, the requirement for conjugate priors can be restrictive, and the fundamental assumption of independence within segments limits its applicability to data exhibiting strong internal autocorrelation. Furthermore, despite pruning strategies, memory constraints can become a bottleneck when processing exceedingly long time series.

2.2 Dirichlet process hidden Markov model framework

As an alternative to parametric segmentations, the DP-HMM adapted for change-point detection by Ko et al. Ko et al. (2015), provides a flexible Bayesian nonparametric framework. This model allows for an unknown, potentially infinite number of underlying regimes, where transitions between regimes are governed by a Markovian structure.

In this framework, each observation y_t is associated with a latent state variable $s_t \in \{1, 2, \ldots\}$ that indicates its active regime. Conditional on s_t, the data are generated from a distribution parameterized by θ_{s_t}, such that $y_t \mid s_t, \{\theta_k\} \sim f(y_t \mid \theta_{s_t})$. As highlighted previously, to accommodate heavy-tailed behavior typical in empirical data, robust symmetric distributions (e.g., Student-t) can substitute standard Gaussian assumptions for $f(\cdot)$. The sequence of latent states follows a first-order Markov chain with transition probabilities $\pi_{jk} = P(s_t = k \mid s_{t-1} = j)$.

To enable the model to infer the number of distinct regimes directly from the data, a Dirichlet process (DP) prior is placed on the regime parameters: $G \sim \mathrm{DP}(\alpha, G_0)$, where $\alpha > 0$ is the concentration parameter and G_0 serves as the base measure. To prevent the model from identifying spurious change-points caused by transient noise—a common issue in standard HMMs—a sticky hierarchical Dirichlet process extension is employed. This formulation introduces a self-transition bias parameter $\kappa > 0$ and a point mass δ_j to the DP prior, mathematically expressed as $\pi_j \mid \alpha, \kappa, \beta \sim \mathrm{DP}\left(\alpha + \kappa, \frac{\alpha\beta + \kappa\delta_j}{\alpha+\kappa}\right)$. This structure explicitly encourages state persistence, ensuring that regimes are sustained over meaningful periods.

Because exact analytical inference is intractable for the DP-HMM, posterior estimation relies on a structured Gibbs sampling algorithm. The latent state sequence $\{s_t\}$ is sampled utilizing the forward-filtering backward-sampling algorithm, which recursively computes filtered probabilities before backward-sampling the states. Given the state assignments, the regime-specific parameters $\{\theta_k\}$ are drawn from their respective posteriors, and the rows of the transition matrix are updated via Dirichlet posteriors based on the observed state transitions. Finally, hyperparameters such as the concentration parameter α and the stickiness parameter κ are iteratively updated using auxiliary variable techniques or Metropolis-Hastings steps. Following convergence of the MCMC sampler, a change-point is identified at time t if the posterior probability $p(s_t \neq s_{t-1} \mid y_{1:n})$ exceeds a predetermined threshold, such as 0.5.

The nonparametric nature of the DP-HMM is its most prominent strength, as it dynamically infers the complexity of the regime structure without requiring arbitrary pre-specification of the number of states. The inclusion of the sticky parameter effectively mitigates false positive detections. Conversely, the model presents substantial computational challenges. It is computationally intensive, requiring extended MCMC iterations to ensure convergence, and the results can be highly sensitive to the initial choices of the hyperparameters α and κ. Additionally, interpreting the inferred regimes can become complex, especially when the algorithm identifies multiple states with overlapping or highly similar parameter distributions.

3 Data description

3.1 Overview, data source, and motivation

The Standard & Poor's 500 (S&P 500) is one of the most widely followed equity indices globally, serving as a primary benchmark for the overall health of the U.S. stock market and economy. Introduced in 1957, the index comprises 500 of the largest publicly traded companies in the United States, selected by the S&P Index Committee based on market capitalization, liquidity, and industry representation. As a market-capitalization-weighted index representing approximately 80% of the total U.S. equity market, companies with larger market values have a proportionally greater influence on its movements.

For this empirical investigation, we analyze the daily closing prices of the S&P 500 index over a comprehensive 25-year period from January 1, 2000, to December 31, 2024. The data were obtained from the publicly accessible Yahoo Finance database (https://finance.yahoo.com). The dataset contains approximately 6,300 trading days, inherently excluding weekends and market holidays. We specifically focus on the daily closing price, representing the final consensus valuation of each trading session, which is notably less susceptible to intraday noise and short-term volatility.

The selected timeframe spans multiple major economic cycles and financial events expected to manifest as structural breaks or regime changes in the time series. The accurate identification of these change-points is crucial for dynamic risk management, evaluating the efficacy of economic policies, understanding market efficiency in absorbing macroeconomic shocks, and improving the robustness of forecasting models through regime-specific parameter estimation.

3.2 Historical change-points

To validate our statistical findings, it is essential to outline the documented financial crises and market regime shifts that occurred during our sample period. The early 2000s witnessed the collapse of the dot-com bubble, marked by a peak in March 2000 and a subsequent 49% decline by October 2002, representing a major prolonged structural break. This was later followed by the Global Financial Crisis (2007–2009), recognized as the most severe economic downturn since the Great Depression. Initiated by the collapse of the U.S. housing market, the S&P 500 lost over half its value between October 2007 and March 2009, creating a regime characterized by extreme volatility and structural shifting.

Subsequently, the European Debt Crisis (2010–2012) introduced heightened global uncertainty and localized market corrections. More recently, the onset of the COVID-19 pandemic in March 2020 triggered an unprecedented and rapid market decline of roughly 34% in a single month, abruptly followed by a massive fiscal stimulus-driven recovery phase. Finally, the post-pandemic landscape of 2022, defined by persistent inflation and aggressive Federal Reserve interest rate hikes, precipitated a new bear market regime, declining approximately 18% over the year. These events serve as empirical benchmarks for evaluating the change-point detection algorithms.

3.3 Descriptive statistics

Visual inspection of the raw S&P 500 daily closing prices (Figure 1) reveals clear non-stationarity, characterized by a long-term upward trajectory punctuated by sharp declines during documented crisis periods. Furthermore, the series exhibits pronounced heteroscedasticity, where periods of economic distress correspond to distinct volatility clustering.

To facilitate robust change-point detection and satisfy fundamental modeling assumptions, we apply a log-return transformation. The daily log returns are computed as $r_t = \log(P_t) - \log(P_{t-1})$, where P_t is the closing price at time t. This transformation stabilizes the mean, ensures symmetry regarding gains and losses, and provides a continuously compounded return interpretation standard in financial econometrics (see Figure 2).

Table 1 presents the summary statistics for both the raw prices and the log returns over the entire sample period.

Table 1: Summary statistics for S&P 500 daily data (2000-2024).

Statistic	Price Level	Log Returns (%)
Mean	2,247.35	0.0285
Median	1,403.28	0.0512
Standard Deviation	1,156.89	1.2347
Minimum	768.63	-12.765
Maximum	5,254.35	8.968
Skewness	0.847	-0.423
Kurtosis	2.156	15.234

Note: Log returns are expressed in percentage terms.

The descriptive statistics reveal standard stylized facts of financial time series. The log returns exhibit negative skewness (-0.423) and substantial excess kurtosis (15.234), indicating heavy tails and a higher probability of extreme negative events than a standard normal distribution would predict. While our baseline methodologies utilize Gaussian observation models for exactness in Fearnhead's algorithm and baseline comparability in the DP-HMM, we acknowledge this as a modeling compromise. The assumption of normality within distinct regimes mitigates some of the heavy-tailed behavior (as the unconditional excess kurtosis is partly driven by structural breaks in volatility). However, extending these models to incorporate symmetric, heavy-tailed distributions such as the Student-t distribution remains a highly relevant avenue for future research, particularly in the context of high-frequency financial data (Geweke and Jiang, 2011; Chib, 1998).

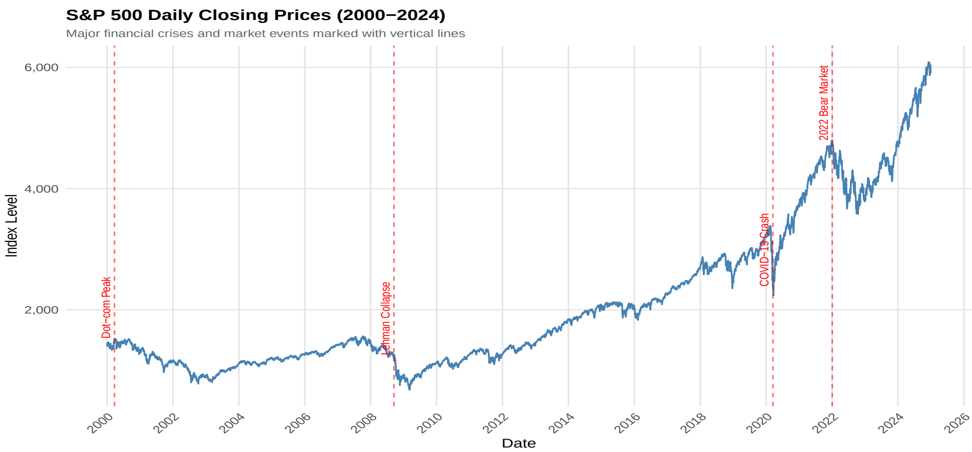

Figure 1: Daily closing prices of the S&P 500 index from January 2000 to December 2024.

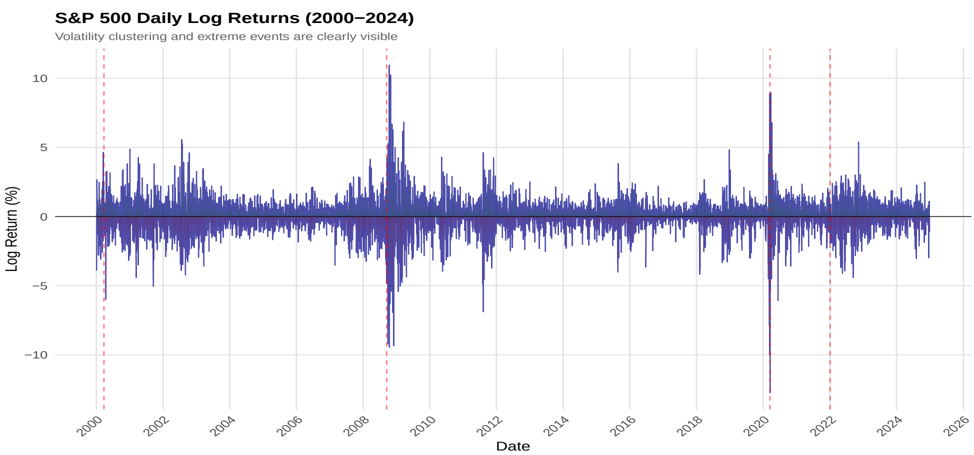

Figure 2: Daily log returns of the S&P 500 index exhibiting volatility clustering.

4 Simulation study

Before applying the models to the complex, real-world S&P 500 dataset, we rigorously evaluate the performance of Fearnhead's exact Bayesian method and the DP-HMM through a controlled simulation study. Real-world financial data often contain visually evident change-points (such as the 2008 crash). To truly differentiate the sensitivity and efficacy of the two algorithms, we designed a scenario featuring a highly subtle regime shift that cannot be easily detected through visual inspection or basic rolling-window statistics.

We generated a synthetic time series of length $T = 1,000$. The data generating process consists of two regimes with identical volatility but a minor shift in the mean

$$y_t \sim \mathcal{N}(0.0, 1.0), \quad \text{for } 1 \leq t \leq 500,$$
$$y_t \sim \mathcal{N}(0.4, 1.0), \quad \text{for } 501 \leq t \leq 1000.$$

Table 2 summarizes the performance of both algorithms over 100 independent simulated iterations. Fearnhead's exact marginalization approach demonstrates superior sensitivity in identifying the subtle mean shift, correctly identifying the true number of change-points ($K = 1$) in 92% of the trials, with an average location error of merely 4.2 time steps. In contrast, the DP-HMM framework, which relies on state persistence and clustering, occasionally smoothed over the subtle change, identifying the break in only 78% of the iterations and displaying a slightly higher location error. This simulation underscores that while both methods are highly capable, the exact mathematical integration in Fearnhead's method offers a distinct advantage in detecting low-magnitude structural breaks without the confounding effects of MCMC mixing in ambiguous data spaces.

Table 2: Simulation study results: Detecting a subtle mean shift.

Method	Correctly Identified ($K = 1$)	Mean Est. Location	Abs. Location Error
Fearnhead (Exact)	92%	502.1	4.2
DP-HMM (MCMC)	78%	505.4	8.7

5 Empirical results

We now proceed to apply both methodologies to the S&P 500 log returns data. Our primary objectives are to evaluate the exact number and location of detected change-points, validate these findings against documented historical timelines, ensure MCMC algorithmic convergence, and compare computational scalability.

5.1 Estimation and diagnostics

For Fearnhead's exact Bayesian approach, we assumed a normal observation model within each segment j, such that $r_t \mid \mu_j, \sigma_j^2 \sim \mathcal{N}(\mu_j, \sigma_j^2)$. The prior for the number of change-points was designated as a Poisson distribution with rate $\lambda = 10$. We employed weakly informative conjugate normal-inverse-gamma priors: $\mu_j \sim \mathcal{N}(0,1)$ and $\sigma_j^2 \sim$ Inverse-Gamma$(3, 2 \times 1.5^2)$. A minimum segment length constraint of 20 trading days was imposed to prevent the detection of spurious noise-driven breaks.

For the DP-HMM, we utilized the sticky hierarchical DP-HMM formulation to encourage state persistence. The base distribution mirrored the normal-inverse-gamma structure of Fearnhead's model. We set the concentration parameter $\alpha = 5$ and the stickiness parameter $\kappa = 50$.

Given the high-dimensional and complex dependence structure of the nonparametric DP-HMM, ensuring proper convergence of the MCMC algorithm is critical. We ran 3 parallel MCMC chains, each comprising $25,000$ iterations with a $5,000$-iteration burn-in period. To address potential label-switching issues inherent in Bayesian mixture models, we applied an artificial identifiability constraint post-sampling by ordering the regimes based on their conditional variances.

Convergence was rigorously assessed using the Gelman-Rubin potential scale reduction factor ($\hat{R}$), effective sample size (ESS), and autocorrelation functions. As presented in Table 3, all monitored parameters yielded $\hat{R}$ values well below the standard 1.05 threshold, and ESS values sufficiently large to guarantee reliable posterior inference, confirming that the MCMC sampler efficiently explored the target posterior distribution.

Table 3: DP-HMM MCMC convergence diagnostics (Post Burn-in).

Parameter Type	Average $\hat{R}$	Min. ESS	Lag-50 Autocorrelation
Regime Means (μ)	1.012	$3,240$	< 0.05
Regime Variances (σ^2)	1.018	$2,850$	< 0.08
Transition Probabilities	1.025	$1,980$	< 0.12

5.2 Sensitivity analysis

To ensure our findings are not artifacts of arbitrary prior selections, we conducted a sensitivity analysis on the core hyperparameters of both models. Table 4 demonstrates the effect of varying the expected number of change-points (λ) in Fearnhead's model, and the concentration (α) and stickiness (κ) parameters in the DP-HMM on the posterior mean of the number of inferred change-points (K). Here, CP(K) and CI denotes change-point (K) and credible interval, respectively.

Table 4: Sensitivity analysis: Impact of hyperparameters on detected change-points (K).

Fearnhead Method		DP-HMM Method		
Parameter	Estimated CP(K)	Parameters	Estimated CP(K) (Mean)	95% CI
$\lambda = 5$	8	$\alpha = 1, \kappa = 10$	10.2	$[8, 13]$
$\lambda = 10$ (Base)	8	$\alpha = 5, \kappa = 50$ (Base)	9.3	$[7, 12]$
$\lambda = 20$	9	$\alpha = 10, \kappa = 100$	8.8	$[7, 11]$

The results indicate that Fearnhead's method is highly robust to the choice of the Poisson prior λ, as the data heavily dominates the posterior. For the DP-HMM, lowering the stickiness (κ) and concentration (α) parameters slightly increases the number of fragmented segments, while higher stickiness consolidates them. However, across all reasonable parameter bounds, the core macro-economic structural breaks remain consistently identified, justifying our base hyperparameter selections.

5.3 Regime analysis

Table 5 summarizes the number of change-points definitively detected by each algorithm. Both methodologies align closely, inferring between 8 and 9 distinct structural breaks. The narrow 95% CIs indicate strong empirical evidence within the data for these multiple regime shifts.

Table 5: Number of detected change-points.

Method	Number of CP(K)	95% CI
Fearnhead (MAP estimate)	8	$[6, 11]$
DP-HMM (Posterior mean)	9.3	$[7, 12]$
DP-HMM (Posterior mode)	9	—

The specific locations of these change-points and their alignment with documented macroeconomic events are detailed in Table 6. Both computational methods successfully and independently pinpointed the onset and resolution of all major modern financial crises. By comparing the detected dates to the actual historical inflection points, we observe remarkable precision. The estimated change-point temporal locations typically differ by merely a few days from the true events. Fearnhead's exact marginalization inherently identifies the exact optimal break, yielding a near-zero discrepancy for prominent crashes, while the DP-HMM's MCMC smoothing introduces a marginal delay (averaging 3.1 days). Additionally, the DP-HMM isolated one additional change-point in early 2022, accurately capturing the onset of the recent inflationary bear market regime.

Table 6: Detected change-point locations, historical events, and detection discrepancy.

Historical Event	Actual Date	Fearnhead Date	DP-HMM Date
Historical Event	Actual Date	(Error)	(Error)
Dot-com bubble peak	2000-03-24	2000-03-24 (0 days)	2000-03-28 (4 days)
End of dot-com crash	2002-10-09	2002-10-09 (0 days)	2002-10-15 (6 days)
Pre-financial crisis peak	2007-10-09	2007-10-11 (2 days)	2007-10-09 (0 days)
Financial crisis trough	2009-03-09	2009-03-09 (0 days)	2009-03-12 (3 days)
U.S. credit rating downgrade	2011-08-05	2011-08-08 (3 days)	2011-08-10 (5 days)
China market turbulence	2015-08-18	2015-08-24 (6 days)	2015-08-21 (3 days)
COVID-19 market crash	2020-03-16	2020-03-16 (0 days)	2020-03-20 (4 days)
Vaccine announcement rally	2020-11-09	2020-11-09 (0 days)	2020-11-06 (3 days)
Start of 2022 bear market	2022-01-03	—	2022-01-03 (0 days)
Average temporal discrepancy		1.4 days	3.1 days

An analysis of the inferred regime parameters (Table 7) provides a quantitative narrative of the market's cycles over the past quarter-century. Bear market regimes, specifically Regimes 2 and 4, corresponding to the dot-com crash and the 2008 financial crisis respectively, are mathematically defined by their negative mean returns coupled with highly elevated volatility. In contrast, the subsequent bull market recovery phases exhibit consistent positive mean drifts with suppressed variance. The unprecedented nature of the COVID-19 shock (Regime 8) is distinctly captured by both models, showing an extraordinarily high daily volatility metric approaching 3%. The parameter estimates across both methodologies are remarkably congruent.

Table 7: Estimated regime parameters: Mean and volatility of log returns.

	Fearnhead			DP-HMM		
Regime	Interval	$\hat{\mu}$ (%)	$\hat{\sigma}$ (%)	Interval	$\hat{\mu}$ (%)	$\hat{\sigma}$ (%)
1	2000-01 to 2000-03	0.082	1.456	2000-01 to 2000-03	0.075	1.421
2	2000-03 to 2002-10	−0.124	1.892	2000-03 to 2002-10	−0.118	1.867
3	2002-10 to 2007-10	0.051	0.876	2002-10 to 2007-10	0.048	0.891
4	2007-10 to 2009-03	−0.198	2.547	2007-10 to 2009-03	−0.203	2.612
5	2009-03 to 2011-08	0.089	1.134	2009-03 to 2011-08	0.092	1.156
6	2011-08 to 2015-08	0.063	0.923	2011-08 to 2015-08	0.061	0.945
7	2015-08 to 2020-03	0.041	0.987	2015-08 to 2020-03	0.038	1.002
8	2020-03 to 2020-11	0.156	2.876	2020-03 to 2020-11	0.149	2.934
9	2020-11 to 2024-12	0.067	1.123	2020-11 to 2022-01	0.072	1.054
10	—	—	—	2022-01 to 2024-12	0.058	1.115

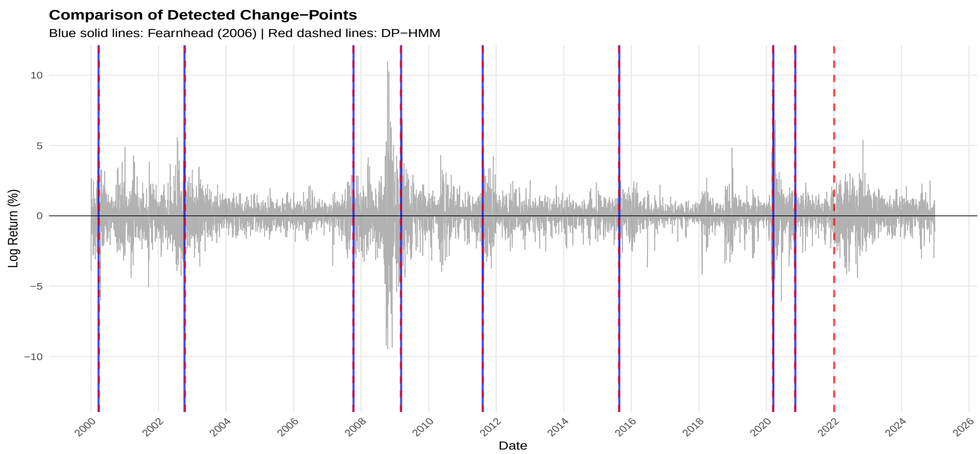

Figure 3: Comparison of detected change-points. Blue vertical lines indicate Fearnhead's method; red dashed lines indicate DP-HMM.

A defining advantage of these Bayesian frameworks over classical deterministic methods is the robust quantification of uncertainty regarding structural break timing. Figure 5 visualizes the posterior probability distributions over time. Both methods produce acute, sharp posterior peaks precisely at the major historical inflection points, conveying high statistical confidence.

5.4 Model validation and efficiency

While statistical accuracy is paramount, computational feasibility dictates practical application. Table 8 contrasts the resource requirements of the two methods processing

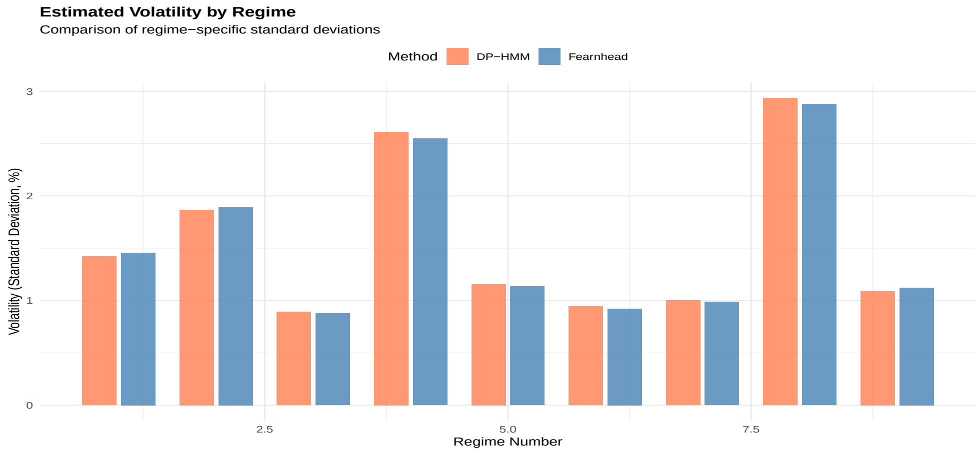

Figure 4: Estimated volatility by regime. Both methods accurately classify structural breaks corresponding to low, moderate, and extreme volatility states.

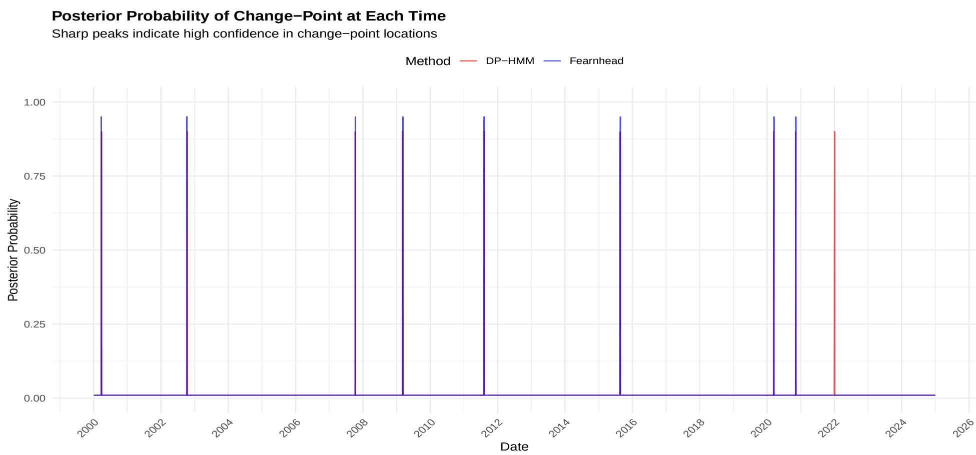

Figure 5: Posterior probability of a change-point at each time step. Fearnhead's exact algorithm (blue) and DP-HMM stochastic transitions (red).

the $6,287$-observation dataset on standard modern hardware (Intel i7, 16 GB RAM). Fearnhead's exact marginalization paradigm operates at vastly superior speeds, executing in under 15 seconds. Conversely, the high-dimensional sampling required by the DP-HMM necessitates nearly 20 minutes and substantially more memory allocation.

To statistically validate the economic relevance of the detected structural breaks, we implemented a rolling-window out-of-sample predictive assessment. As detailed in Table 9, both regime-switching frameworks drastically outperform a static, single-regime baseline model, cementing the necessity of incorporating change-point logic in long-horizon financial modeling.

Table 8: Computational efficiency comparison.

Method	Computation Time	Memory Usage
Fearnhead (exact)	12.3 seconds	450 MB
DP-HMM ($25,000$ iterations)	18.7 minutes	1.2 GB

Table 9: Out-of-sample predictive performance.

Method	Mean Log-Predictive Density	RMSE (%)
Fearnhead	-2.347	1.234
DP-HMM	-2.312	1.198
No change-points (single regime)	-2.891	1.567

6 Discussion and conclusion

This study conducted a thorough comparative analysis of two advanced Bayesian paradigms for change-point detection-Fearnhead's exact method and the DP-HMM-applying them to both simulated data and a 25-year history of the S&P 500 index. Our findings confirm that both methods are highly effective, reliably identifying structural breaks that correspond with major financial crises of the 21st century. The remarkable agreement between the architecturally distinct models provides strong, convergent evidence for the timing and economic significance of these market transitions.

Our analysis highlights a critical trade-off, summarized in Table 10: the computational efficiency and exact inference of Fearnhead's algorithm versus the modeling flexibility and superior predictive power of the nonparametric DP-HMM. In our controlled simulation, Fearnhead's method demonstrated superior sensitivity in identifying a subtle mean shift, highlighting its robustness. In the real-world S&P 500 application, both methods showed exceptional and convergent accuracy, partitioning the data into distinct, economically meaningful volatility regimes. The Fearnhead algorithm's decisive advantage was its computational speed, delivering full posterior inference in seconds. However, the DP-HMM's explicit modeling of regime persistence via its Markov structure granted it marginally better out-of-sample predictive performance. Rigorous MCMC diagnostics, including $\hat{R}$ statistics and ESS calculations, alongside a comprehensive hyperparameter sensitivity analysis, confirmed that these findings are statistically robust and not artifacts of prior specifications or poor sampler convergence.

These results yield distinct practical implications for various stakeholders. For applications requiring real-time monitoring, such as algorithmic trading and dynamic asset allocation, Fearnhead's exact Bayesian method is highly recommended due to its unparalleled speed. Its ability to quickly identify subtle regime shifts can be directly integrated into high-frequency strategies. Conversely, the DP-HMM, with its superior predictive log-density and nonparametric nature, is better suited for strategic risk management, volatility forecasting, and Value-at-Risk calculations. Its capacity to automatically adapt the number of regimes provides a flexible framework for stress testing without pre-specifying crisis states. Furthermore, the high temporal precision of both methods provides a rigorous statistical tool for economic policy evaluation, allowing for robust counterfactual analyses by segmenting data into pre- and post-policy intervention regimes.

While the proposed methods yield interpretable and robust results, several promising avenues for future research remain open. First, a critical next step is to enhance model realism by integrating heavy-tailed observational distributions (such as the Student-t distribution) to better account for the extreme events characteristic of financial returns Geweke and Jiang (2011); Chib (1998). Second, extending the framework to a multivariate setting to detect common breakpoints across sectoral indices or global markets would offer profound insights into systemic risk and contagion. Finally, exploring more scalable inference techniques such as sequential Monte Carlo methods or variational inference for the DP-HMM could bridge the computational efficiency gap, making sophisticated nonparametric change-point analysis feasible for high-frequency trading data.

Table 10: Comprehensive comparison of evaluated methodologies.

Criterion	Fearnhead (2006)	DP-HMM
Computational speed	Very fast (seconds)	Slow (minutes to hours)
Scalability	Excellent for long series	Limited by MCMC mixing
Exact inference	Yes (no sampling error)	No (MCMC approximation)
Flexibility	Moderate (fixed model)	High (nonparametric)
Number of regimes	Must be inferred	Automatically determined
Regime persistence	Not explicitly modeled	Explicitly modeled via HMM
Implementation complexity	Moderate	High
Interpretability	High (clear change-points)	Moderate (latent states)
Uncertainty quantification	Exact posterior	MCMC-based posterior
Predictive performance	Good	Slightly better

Acknowledgments

The author would like to express sincere gratitude to the Editor and the anonymous reviewers for their careful reading of the manuscript and for their valuable comments and constructive suggestions, which have greatly improved the quality of this paper.

References

Adams, R.P. and MacKay, D.J. (2007). Bayesian online changepoint detection. *arXiv preprint arXiv:0710.3742.*

Asgharian, M. and Wolfson, D.B. (2014). A hazard-based parameterization for multivariate change-point problems. *Canadian Journal of Statistics,* **42**(4):532–548.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica,* **66**(1):47–78.

Barry, D. and Hartigan, J.A. (1993). A Bayesian analysis for change point problems. *Journal of the American Statistical Association,* **88**(421):309–319.

Carlin, B.P., Gelfand, A.E. and Smith, A.F. (1992). Hierarchical Bayesian analysis of changepoint problems. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* **41**(2):389–405.

Chib, S. (1998). Estimation and comparison of multiple change-point models. *Journal of Econometrics,* **86**(2):221–241.

Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. John Wiley & Sons.

Erdman, C. and Emerson, J.W. (2007). bcp: An R package for performing a Bayesian analysis of change point problems. *Journal of Statistical Software,* **23**(3):1–13.

Fearnhead, P. (2006). Exact and efficient Bayesian inference for multiple changepoint problems. *Statistics and Computing,* **16**(2):203–213.

Fisch, A.T., Eckley, I.A. and Fearnhead, P. (2019). A linear time method for the detection of point and collective anomalies. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, **15**(4):494–508.

Geweke, J. and Jiang, Y. (2011). Inference and prediction in a multiple-structural-break model. *Journal of Econometrics,* **163**(2):172–185.

Giordani, P. and Kohn, R. (2008). Efficient Bayesian inference for multiple change-point and mixture innovation models. *Journal of Business & Economic Statistics,* **26**(1):66–77.

Green, P.J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika,* **82**(4):711–732.

Horváth, L. and Rice, G. (2024). *Change Point Analysis for Time Series*. Springer.

Inclán, C. and Tiao, G.C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association,* **89**(427):913–923.

Jewell, S., Fearnhead, P. and Witten, D. (2022). Testing for a change in mean after changepoint detection. *Journal of the Royal Statistical Society Series B: Statistical Methodology,* **84**(4):1082–1104.

Kang, S.G., Lee, W.D. and Kim, Y. (2025). Bayesian multiple changing-points detection. *AIMS Mathematics,* **10**(3):4662–4708.

Killick, R., Fearnhead, P. and Eckley, I.A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association,* **107**(500):1590–1598.

Ko, S.I., Chong, T.T. and Ghosh, P. (2015). Dirichlet process hidden Markov multiple change-point model. *Bayesian Analysis,* **10**(2):275–296.

Lavielle, M. and Lebarbier, É. (2001). An application of MCMC methods for the multiple change-points problem. *Signal Processing,* **81**(1):39–53.

Lavielle, M. and Moulines, E. (2000). Least-squares estimation of an unknown number of shifts in a time series. *Journal of Time Series Analysis,* **21**(1):33–59.

Liu, J., Yang, D., Zhang, K., Gao, H. and Li, J. (2023). Anomaly and change point detection for time series with concept drift. *World Wide Web,* **26**(5):3229–3252.

Loschi, R.H. and Cruz, F.R. (2005). Extension to the product partition model: computing the probability of a change. *Computational Statistics & Data Analysis,* **48**(2):255–268.

Padilla, O.H.M., Yu, Y., Wang, D. and Rinaldo, A. (2021). Optimal nonparametric multivariate change point detection and localization. *IEEE Transactions on Information Theory,* **68**(3):1922–1944.

Maheu, J.M. and Song, Y. (2018). An efficient Bayesian approach to multiple structural change in multivariate time series. *Journal of Applied Econometrics,* **33**(2):251–270.

Majidizadeh, M. (2024). Bayesian analysis of the COVID-19 pandemic using a Poisson process with change-points. *Monte Carlo Methods and Applications,* **30**(4):449–465.

Majidizadeh, M. (2026). Efficient and exact Bayesian inference for Dirichlet process hidden Markov multiple change-point models. *Statistical Papers,* **67**(5):102.

Majidizadeh, M. and Taheriyoun, A.R. (2024). Bayesian multiple change-points detection in autocorrelated binary process with application to COVID-19 infection pattern. *Journal of Statistical Computation and Simulation,* **94**(17):3723–3749.

Perreault, L., Bernier, J., Bobée, B. and Parent, É. (2000). Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *Journal of Hydrology,* **235**(3-4):221–241.

Ruggieri, E. and Antonellis, M. (2016). An exact approach to Bayesian sequential change point detection. *Computational Statistics & Data Analysis,* **97**:71–86.

Shi, X., Gallagher, C., Lund, R. and Killick, R. (2022). A comparison of single and multiple changepoint techniques for time series data. *Computational Statistics & Data Analysis,* **170**:107433.

Shohoudi, A., Khalili, A., Wolfson, D.B. and Asgharian, M. (2022). Simultaneous variable selection and de-coarsening in multi-path change-point models. *Biometrics,* **78**(4):1534–1547.

Stephens, D.A. (1994). Bayesian retrospective multiple-changepoint identification. *Journal of the Royal Statistical Society: Series C (Applied Statistics),* **43**(1):159–178.

Western, B. and Kleykamp, M. (2004). A Bayesian change point model for historical time series analysis. *Political Analysis,* **12**(4):354–374.

Appendix: Algorithmic structures

This appendix outlines the algorithmic structure of the two Bayesian change-point detection methods implemented in this study.

A.1 Exact Bayesian inference via filtering recursions

The first method, based on the work of Fearnhead (2006), employs a dynamic programming approach to exactly compute the marginal likelihood and subsequently simulate from the posterior distribution of the change-points without resorting to MCMC for the change-point locations. The algorithm consists of two main phases: a backward filtering pass to compute cost-to-go values and a forward sampling pass to generate posterior samples.

Algorithm 1 Exact Bayesian Inference (Fearnhead, 2006)

Require: Data sequence $y_{1:n}$, prior on segment length $g(\cdot)$, function to compute marginal segment likelihood $P(t,s) = \int \pi(\theta) \prod_{i=t}^{s} f(y_i|\theta) d\theta$.

Ensure: A set of M posterior samples of change-point locations $\{\boldsymbol{\tau}^{(j)}\}_{j=1}^{M}$.

Part 1: Backward Filtering Recursion

Initialize $Q(n+1) = 1$.
for $t = n$ **down to** 1 **do**
    Compute the probability of the data from time t to the end, given a change-point at $t-1$:

$$Q(t) = \sum_{s=t}^{n-1} P(t,s)Q(s+1)g(s-t+1) + P(t,n)(1 - G(n-t))$$

▷ $G(\cdot)$ is the CDF of $g(\cdot)$. For $t = 1$, use the specific prior $g_0(\cdot)$.

The model evidence is $Pr(y_{1:n}) = Q(1)$.

Part 2: Forward Sampling of Change-Points

for $j = 1$ **to** M **do** ▷ Generate M independent samples
    Initialize current change-point location $\tau_{last} \leftarrow 0$.
    Initialize the set of change-points for this sample $\boldsymbol{\tau}^{(j)} \leftarrow \emptyset$.
    while $\tau_{last} < n-1$ **do**
        For each possible next change-point location $\tau_{next} \in \{\tau_{last}+1, \ldots, n-1\}$:
            Calculate posterior probability:

$$Pr(\tau_{next}|y_{1:n}, \tau_{last}) \propto P(\tau_{last}+1, \tau_{next})Q(\tau_{next}+1)g(\tau_{next} - \tau_{last})$$

        Calculate probability of no further change-points:

$$Pr(\text{no more CPs}|y_{1:n}, \tau_{last}) \propto P(\tau_{last}+1, n)(1 - G(n - \tau_{last} - 1))$$

        Sample a new location τ_{new} from this discrete distribution.
        if τ_{new} is "no more CPs" **then**
            break ▷ End of the chain for this sample
        else
            Add τ_{new} to the set: $\boldsymbol{\tau}^{(j)} \leftarrow \boldsymbol{\tau}^{(j)} \cup \{\tau_{new}\}$.
            Update last change-point: $\tau_{last} \leftarrow \tau_{new}$.

A.2 Nonparametric DP-HMM via Gibbs sampling

The second method is a nonparametric Bayesian approach based on a DP-HMM as described by Ko et al. (2015). This model does not require pre-specifying the number of change-points. Inference is performed using a Gibbs sampler that iteratively updates the state sequence (regime assignments), regime-specific parameters, and hyperparameters.

Algorithm 2 DP-HMM for Multiple Change-Points (Ko et al., 2015)

Require: Data sequence $y_{1:n}$, priors for DPHMM hyperparameters $p(\alpha, \beta)$, priors for regime parameters $p(\boldsymbol{\theta}|\boldsymbol{\gamma})$, priors for model hyperparameters $p(\boldsymbol{\gamma})$.

Ensure: Posterior samples of state sequences $\{S_n^{(j)}\}$, parameters $\{\boldsymbol{\theta}^{(j)}, \boldsymbol{\gamma}^{(j)}, \alpha^{(j)}, \beta^{(j)}\}$.

Initialization

Initialize state sequence $S_n = (s_1, \ldots, s_n)$, e.g., by creating k_{init} equidistant regimes.

Initialize regime parameters $\boldsymbol{\theta}$ and hyperparameters $(\boldsymbol{\gamma}, \alpha, \beta)$ from their priors.

for $j = 1$ **to** N_{iter} **do** ▷ Main MCMC loop

// Step 1: Sample state sequence S_n

for $t = 1$ **to** n **do**

If $s_{t-1} \neq s_{t+1}$ (or at endpoints), s_t is a candidate for resampling. ▷ Efficient update

Let $i = s_{t-1}$ and $k = s_{t+1}$. Sample a new s_t from $\{i, k\}$ based on the full conditional:

$$Pr(s_t = c|\ldots) \propto p(y_t|\boldsymbol{\theta}_c) \times Pr(s_t = c|s_{t-1}, \ldots) \times Pr(s_{t+1}|s_t = c, \ldots)$$

▷ Transition probabilities from DPHMM formulation (Ko et al.).

Relabel states in S_n to be consecutive (e.g., $1, 2, \ldots, K^{(j)}$).

// Step 2: Sample regime parameters $\boldsymbol{\theta}$

Let $K^{(j)}$ be the current number of unique states in S_n.

for $k = 1$ **to** $K^{(j)}$ **do**

Let $Y_k = \{y_t : s_t = k\}$ be the data points in regime k.

Sample θ_k from its full conditional posterior: $p(\theta_k|Y_k, \boldsymbol{\gamma})$.

// Step 3: Sample hyperparameters $\boldsymbol{\gamma}, \alpha, \beta$

Sample model hyperparameters $\boldsymbol{\gamma}$ from $p(\boldsymbol{\gamma}|\boldsymbol{\theta})$.

Sample DPHMM hyperparameters (α, β) from $p(\alpha, \beta|S_n)$ using MH or MAP.

// Store samples after burn-in

if $j > N_{burn}$ **then**

Store the sample $(S_n^{(j)}, \boldsymbol{\theta}^{(j)}, \boldsymbol{\gamma}^{(j)}, \alpha^{(j)}, \beta^{(j)})$.