A A list of symbols used

Symbols

Norms	
$\\|\cdot\\|_p$	ℓ_p-norm for sequences.
$\\|\cdot\\|_{\mathrm{F}}$	Frobenius norm for matrices.
$\\|\cdot\\|$	Spectral norm for matrices.
$d_{\mathrm{TV}}(\cdot,\cdot)$	Total variation distance between two distributions.
Sets	
$\mathbb{S}^{n-1}$	The n-dimensional unit sphere
Δ^{n-1}	Probability simplex of dimension $n-1$.
$\mathbb{\Delta}^{n\times(n-1)}$	Set of n-dimensional left-stochastic matrices.
$\mathbb{N}_0 \triangleq \{0,1,2,\ldots\}$	Set of nonnegative integers.
$\mathbb{N}_+ \triangleq \{1,2,3,\ldots\}$	Set of strictly positive integers.
$\mathbb{R} \triangleq (-\infty,\infty)$	Set of reals.
$[n]=\{1,2,\ldots,n\}$	Set of integers 1 through n.
$\mathrm{Perm}(n)$	All permutations of length n.
Asymptotics	
$f(n)=\omega(g(n))$	Small-omega notation. $\liminf_{n\to\infty} f(n)/g(n)=\infty$
$f(n)=\Omega(g(n))$	Big-omega notation. $\liminf_{n\to\infty} f(n)/g(n)>0$
$f(n)=O(g(n))$	Big-o notation. $\limsup_{n\to\infty} f(n)/g(n)<\infty$
$f(n)\sim g(n)$	Asymptotic equivalence. $\lim_{n\to\infty} f(n)/g(n)=1$
$f(n)=o(g(n))$	Little-o notation. $\limsup_{n\to\infty} f(n)/g(n)=0$
$X_n=\Omega_{\mathbb{P}}(a_n)$	Omega-p notation. $\forall_\varepsilon \exists_{\delta_\varepsilon,N_\varepsilon} : \mathbb{P}[\lvert X_n/a_n\rvert \le \delta_\varepsilon] \le \varepsilon \,\forall_{n>N_\varepsilon}$
$X_n=O_{\mathbb{P}}(a_n)$	Stochastic boundedness. $\forall_\varepsilon \exists_{\delta_\varepsilon,N_\varepsilon} : \mathbb{P}\left[\left\lvert\frac{X_n}{a_n}\right\rvert \ge \delta_\varepsilon\right] \le \varepsilon \,\forall_{n>N_\varepsilon}$
$X_n \asymp_{\mathbb{P}} (a_n)$	Equivalence-p notation. $\forall_\varepsilon \exists_{\delta_\varepsilon^-,\delta_\varepsilon^+,N_\varepsilon} : \mathbb{P}[\delta_\varepsilon^- \le \lvert X_n/a_n\rvert \le \delta_\varepsilon^+] \ge 1-\varepsilon \,\forall_{n>N_\varepsilon}$
$X_n=o_{\mathbb{P}}(a_n)$	Convergence in probability. $\mathbb{P}\left[\left\lvert\frac{X_n}{a_n}\right\rvert \ge \delta\right] \to 0 \,\forall_{\delta>0} \Leftrightarrow \forall_{\varepsilon,\delta}\exists_{N_{\varepsilon,\delta}} : \mathbb{P}\left[\left\lvert\frac{X_n}{a_n}\right\rvert \ge \delta\right] \le \varepsilon \,\forall_{n>N_{\varepsilon,\delta}}$
"an absolute constant"	A constant independent of n.
Block Markov Chain	
$n\in\mathbb{N}_+$	Number of states.
$K\in\mathbb{N}_+$	Number of clusters.
$\mathcal{V}_1,\ldots,\mathcal{V}_K \subseteq [n] \triangleq \mathcal{V}$	Clusters, and set of all states. $\mathcal{V}=\cup_{k=1}^K \mathcal{V}_k$, and $\mathcal{V}_a\cap\mathcal{V}_b=\emptyset\,\forall_{a\neq b}$
$\alpha_1,\ldots,\alpha_K\in(0,1)$	Relative sizes of the clusters. $\alpha_k \triangleq \lim_{n\to\infty} \lvert\mathcal{V}_k\rvert/n$.
$\alpha_{\min}>0,\alpha_{\max}$	Minimum and maximum relative sizes of the clusters.
$\sigma:[n]\to[K]$	Cluster assignment.
$\{X_t\}_{t\ge 0}$	Markov chain.
$p\in\mathbb{\Delta}^{K\times(K-1)}, P,Q\in\mathbb{\Delta}^{n\times(n-1)}$	Transition matrices. $P_{x,y} \triangleq \frac{p_{\sigma(x),\sigma(y)}}{\lvert\mathcal{V}_{\sigma(y)}\rvert-\mathbb{1}[\sigma(x)=\sigma(y)]}\mathbb{1}[x\neq y]\,\forall_{x,y\in\mathcal{V}}$
$\eta>1$	Assumed separability constant. $\exists_{1<\eta} : \max_{a,b,c}\{p_{b,a}/p_{c,a}, p_{a,b}/p_{a,c}\} \le \eta$
$\pi\in\Delta^{K-1}, \Pi\in\Delta^{n-1}$	(Limiting) Equilibrium distribution. $\Pi_x \triangleq \lim_{t\to\infty}\mathbb{P}[X_t=x]\,\forall_{x\in\mathcal{V}}$ $\pi_k \triangleq \lim_{n\to\infty}\sum_{x\in\mathcal{V}_k}\Pi_x\,\forall_{k\in[K]}$
$P^*\in\mathbb{\Delta}^{n\times(n-1)}$	Time-reversed transition matrix. $P^*_{x,y} \triangleq \frac{P_{x,y}}{\Pi_x}\Pi_y$
$\delta(P)\in(0,\infty)$	Dobrushin's ergodic coefficient. $\delta(P) \triangleq \frac{1}{2}\sup_{x,y\in\mathcal{V}}\sum_{z\in\mathcal{V}}\lvert P_{x,z}-P_{y,z}\rvert$

$0 \leq t_{\text{mix}}(\varepsilon) \leq -c_{\text{mix}} \ln \varepsilon$	Mixing time. $t_{\text{mix}}(\varepsilon) \triangleq \min\{t \geq 0 : \sup_{x \in \mathcal{V}} d_{\text{TV}}(P^t_{x,\cdot}, \Pi) \leq \varepsilon\}$ We prove the second inequality, with c_{mix} being an absolute constant.
γ_{ps}	Pseudo spectral gap. $\gamma_{\text{ps}} \triangleq \max_{i \geq 1} \frac{1-\lambda((P^*)^i P^i)}{i}$
Information bound	
$\mathbb{P}_P[X_0 = x_0, \ldots, X_T = x_T]$	Probability of a sample path. $\mathbb{P}_P[X_0 = x_0, \ldots, X_T = x_T] \triangleq \prod_{t=1}^T P_{x_{t-1},x_t}$
V^*	Vertex chosen uniformly at random from two clusters. $V^* \stackrel{d}{=} \text{Unif}(\mathcal{V}_a \cup \mathcal{V}_b), a, b \in [K], a \neq b$
Φ	Probability measure of the true model. I.e., under P and cluster assignments $\mathcal{V}_1, \ldots, \mathcal{V}_K$.
Ψ	Probability measure of the modified model. I.e., under Q which is a perturbation of P constructed after the random vertex V^* was moved into its own cluster.
$\mathcal{Q}$	Set of all possible change-of-measure parameters. $\mathcal{Q} \triangleq \left\{(q_{k,0}, q_{0,k})_{k=0,\ldots,K} \in (0,\infty) \middle\vert q_{0,0} = 0, \sum_{l=1}^K q_{0,l} = 1\right\}$
$\mathcal{Q}(k,l)$	Sets of change-of-measure parameters leading to confusion between assigning to either cluster k, l. $\mathcal{Q}(k,l) \triangleq \left\{q \in \mathcal{Q} \middle\vert I_k(q\|p) = I_l(q\|p)\right\} \neq \emptyset$, $k \neq l$
$L \triangleq \ln \frac{\mathbb{P}_Q[X_0, X_1, \ldots, X_T]}{\mathbb{P}_P[X_0, X_1, \ldots, X_T]}$	Log-likelihood ratio.
$I_c(q\|p)$	Leading order coefficient in an asymptotic expansion of the log-likelihood ratio. $I_c(q\|p) \triangleq \lim_{n\to\infty} \frac{n}{T} \mathbb{E}_Q[L \vert \sigma(V^*) = c]$
$I_{a,b}(q\|p)$	Deconditioned leading order coefficient in an asymptotic expanding of the log-likelihood ratio. $I_{a,b}(q\|p) \triangleq \lim_{n\to\infty} \frac{n}{T} \mathbb{E}_\Psi[L]$
$I_{a,b}(\alpha, p)$	Separation between cluster a and b. $I_{a,b}(\alpha,p) \triangleq \sum_{k=1}^K \frac{1}{\alpha_a}\left(\pi_a p_{a,k} \ln \frac{p_{a,k}}{p_{b,k}} + \pi_k p_{k,a} \ln \frac{p_{k,a}\alpha_b}{p_{k,b}\alpha_a}\right) + \left(\frac{\pi_b}{\alpha_b} - \frac{\pi_a}{\alpha_a}\right)$
$0 \leq J(\alpha,p) \leq I(\alpha,p) < \infty$	Information quantities. $J(\alpha,p) \triangleq \min_{k\neq l} \min_{q \in \mathcal{Q}(k,l)} I_{k,l}(q\|p)$ $I(\alpha,p) \triangleq \min_{a\neq b} I_{a,b})\alpha, p)$
Generic algorithm quantities	
$T \in \mathbb{N}_+$	Observation length.
$X_0, X_1, \ldots, X_T$	Sample path of our Markov chain.
$\hat{N} \in \mathbb{N}_0^{n\times n}$	Observation matrix. Each entry contains the number of times the Markov chain jumped from x to y, i.e., $\hat{N}_{x,y} \triangleq \sum_{t=0}^{T-1} \mathbb{1}[X_t = x, X_{t+1} = y] \,\forall_{x,y\in\mathcal{V}}$
$\hat{\mathcal{V}}_1, \ldots, \hat{\mathcal{V}}_K$	Approximated cluster assignments.
$\gamma^{\text{opt}} \in \min_{\gamma \in \text{Perm}(K)} \left\vert\bigcup_{k=1}^K \hat{\mathcal{V}}_{\gamma(k)} \backslash \mathcal{V}_k\right\vert$	Permutation that minimizes the overlap between the true clusters and a cluster assignment.
$\mathcal{E} \subset \mathcal{V}$	Set of misclassified vertices. $\mathcal{E} \triangleq \cup_{k=1}^K \hat{\mathcal{V}}_{\gamma^{\text{opt}}(k)} \backslash \mathcal{V}_k$
Regime terminologies	
"asymptotically accurate detection"	$\mathbb{E}_P[\vert\mathcal{E}\vert] = o(n)$
"asymptotically exact detection"	$\mathbb{E}_P[\vert\mathcal{E}\vert] = o(1)$
"dense regime"	$T = \omega(n \ln n)$
"critical regime"	$T \sim cn \ln n$ for some absolute constant $c > 0$
"sparse regime"	$\omega(n) = T = o(n \ln n)$
Spectral clustering algorithm	
$\Gamma \subseteq \mathcal{V}$	Truncated vertices. This set is obtained from $\mathcal{V}$ by removing the $\lfloor n \exp(-(T/n)\ln(T/n))\rfloor$ states with the highest numbers of visits in the observed sample path of length T.
$\hat{N}_\Gamma$	Truncated observation matrix. This matrix is obtained from $\hat{N}$ by setting all entries on the rows and columns corresponding to setates not in Γ to zero.
$U\Sigma V^{\text{T}}$	Singular value decomposition of $\hat{N}_\Gamma$.
$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$	Singular values of $\hat{N}_\Gamma$.
$\hat{R} \triangleq \sum_{k=1}^K \sigma_k U_{\cdot,k} V_{\cdot,k}{}^{\text{T}}$	Best rank-K approximation of $\hat{N}_\Gamma$.

$\mathcal{N}_x$	Neighborhood. $\mathcal{N}_x \triangleq \left\{y \in \mathcal{V} \middle\| \sqrt{\|\hat{R}_{x,\cdot} - \hat{R}_{y,\cdot}\|_2^2 + \|\hat{R}_{\cdot,x} - \hat{R}_{\cdot,y}\|_2^2} \leq \frac{1}{n} \cdot \left(\frac{T}{n}\right)^{3/2}\left(\ln\frac{T}{n}\right)^{4/3}\right\}$
$z_1^*, \ldots, z_K^* \in \mathcal{V}$	Iteratively constructed cluster centers.
Cluster improvement algorithm	
$\hat{p}, \hat{\pi}, \hat{\alpha}$	Approximated BMC parameters.
$u_x^{[t]}(c)$	Approximated difference between two log-likelihood functions. $u_x^{[t]}(c) \triangleq \left\{\sum_{k=1}^K \left(\hat{N}_{x,\hat{\mathcal{V}}_k^{[t]}} \ln \hat{p}_{c,k} + \hat{N}_{\hat{\mathcal{V}}_k^{[t]},x} \ln \frac{\hat{p}_{k,c}}{\hat{\alpha}_c}\right) - \frac{T}{n} \cdot \frac{\hat{\pi}_c}{\hat{\alpha}_c}\right\}$
$\mathcal{H} \subseteq \mathcal{V}$	Set of well-behaved vertices.