505B Final Solutions[1]

1. Question 1

Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion. Let $\{Y(t)\}_{t\geq 0}$, and let $\{Z(t)\}_{t\geq 0}$ be two (coupled) Itô processes. So, $\sigma_1, \mu_1, \sigma_2, \mu_2 \colon \mathbf{R}^2 \to \mathbf{R}$ are continuous functions such that

$$dY(t) = \sigma_1(t)dB(t) + \mu_1(t)dt, \qquad \forall\, t \geq 0.$$
$$dZ(t) = \sigma_2(t)dB(t) + \mu_2(t)dt, \qquad \forall\, t \geq 0.$$

Show that the following "product rule" holds for all $t > 0$:

$$d(YZ)(t) = Y(t)dZ(t) + Z(t)dY(t) + \sigma_1(t, Y(t))\sigma_2(t, Z(t))dt.$$

(Hint: apply Itô's formula to Y^2, Z^2 and $(Y+Z)^2$ each separately and then use the equality $YZ = (1/2)((Y+Z)^2 - Y^2 - Z^2)$.)

Solution. Itô's formula and Lemma 8.33

$$dY^2(t) = 2Y(t)dY(t) + d[Y]_t = 2Y(t)dY(t) + \sigma_1^2(t)dt$$
$$dZ^2(t) = 2Z(t)dZ(t) + d[Z]_t = 2Z(t)dZ(t) + \sigma_2^2(t)dt$$

Note that

$$d(Y+Z)(t) = dY(t) + dZ(t) = \Big(\sigma_1(t) + \sigma_2(t)\Big)dB(t) + \Big(\mu_1(t) + \mu_2(t)\Big)dt, \qquad \forall\, t \geq 0.$$

So, $Y + Z$ is an Itô process, and again by Itô's formula and Lemma 8.33

$$\begin{aligned} d(Y+Z)^2(t) &= 2(Y(t) + Z(t))d(Y+Z)(t) + d[Y+Z]_t \\ &= 2(Y+Z)(t)(dY(t) + dZ(t)) + (\sigma_1(t) + \sigma_2(t))^2 dt. \end{aligned}$$

So

$$\begin{aligned} &2d(YZ)(t) = d(Y+Z)^2(t) - dY^2(t) - dZ^2(t) \\ &= 2Y(t)dZ(t) + 2Z(t)dY(t) + \Big((\sigma_1(t) + \sigma_2(t))^2 - \sigma_1^2(t) - sigma_2^2(t)\Big)dt \end{aligned}$$

2. Question 2

Let $\{Y(t)\}_{t\geq 0}$, and let $\{Z(t)\}_{t\geq 0}$ be two Itô processes. Then, we define the **covariation** process $\{[Y,Z]_t\}_{t\geq 0}$ so that, for any $b > 0$, $[Y,Z]_b$ is the limit in probability as $n \to \infty$ of

$$\sum_{i=0}^{n-1} (Y(b(i+1)/n) - Y(bi/n))(Z(b(i+1)/n) - Z(bi/n)),$$

if this limit exists. (In particular, $[Y,Y]_t = [Y]_t$ for all $t \geq 0$, where $[Y]_t$ denotes the quadratic variation of Y at time t.)

- Show that, for any $t > 0$, with probability one we have
$$[Y+Z]_t = [Y]_t + [Z]_t + 2[Y,Z]_t.$$
- Show that, for any $t > 0$, with probability one we have
$$\big|[Y,Z]_t\big| \leq \sqrt{[Y]_t[Z]_t}.$$

(This is a special case of the so-called Kunita-Watanabe inequality.)

[1]May 17, 2021,

- Using the arithmetic mean-geometric mean inequality, conclude that, for any $t > 0$, with probability one we have

$$[Y + Z]_t \leq 2\big([Y]_t + [Z]_t\big), \qquad \forall\, t > 0.$$

Solution. For any $b > 0$, $n \geq 1$ and $0 \leq i \leq n-1$, we have

$$\begin{aligned}
&\Big((Y+Z)(b(i+1)/n) - (Y+Z)(bi/n)\Big)\Big((Y+Z)(b(i+1)/n) - (Y+Z)(bi/n)\Big)\\
&\quad = (Y(b(i+1)/n) - Y(bi/n))(Y(b(i+1)/n) - Y(bi/n))\\
&\qquad + (Z(b(i+1)/n) - Z(bi/n))(Z(b(i+1)/n) - Z(bi/n))\\
&\qquad + 2(Y(b(i+1)/n) - Y(bi/n))(Z(b(i+1)/n) - Z(bi/n)).
\end{aligned}$$

Summing over $0 \leq i \leq n-1$, and then letting $n \to \infty$ implies that $[Y+Z]_b = [Y]_b + [Z]_b + 2[Y,Z]_b$ for any $b > 0$.

From the Cauchy-Schwarz inequality (for discrete sequences of real numbers), for any $b > 0$ and $n \geq 1$,

$$\begin{aligned}
&\Big(\sum_{i=0}^{n-1}(Y(b(i+1)/n) - Y(bi/n))(Z(b(i+1)/n) - Z(bi/n))\Big)^2\\
&\leq \sum_{i=0}^{n-1}(Y(b(i+1)/n) - Y(bi/n))^2 \sum_{j=0}^{n-1}(Z(b(j+1)/n) - Z(bj/n))^2.
\end{aligned}$$

Letting $n \to \infty$ implies that $[Y,Z]_b^2 \leq [Y]_b[Z]_b$ for any $b > 0$. The AMGM inequality them implies that

$$[Y,Z]_b \leq \sqrt{[Y]_b[Z]_b} \leq (1/2)([Y]_b + [Z]_b) \qquad (*).$$

Combining the above,

$$[Y+Z]_b = [Y]_b + [Z]_b + 2[Y,Z]_b \overset{(*)}{\leq} 2\big([Y]_b + [Z]_b\big).$$

3. Question 3

Let $\{B(t)\}_{t\geq 0}$ be a Brownian motion in $\mathbf{R}^n$ (so that $B(t) = (B_1(t), \ldots, B_n(t))$ where $\{B(t)\}_{t\geq 0}, \ldots, \{B_n(t)\}_{t\geq 0}$ are n independent one-dimensional Brownian motions). For any $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$, denote $\|x\| := (x_1^2 + \cdots + x_n^2)^{1/2}$. Denote the open unit ball in $\mathbf{R}^n$ as

$$D := \{x = (x_1, \ldots, x_n) \in \mathbf{R}^n \colon\ \|x\| < 1\}.$$

For any $x \in D$, we use the notation $\mathbf{E}_x$ to denote that the Brownian motion is started at x (so that $B(0) = x$). Define a stopping time

$$T := \inf\{t > 0 \colon B(t) \in \partial D\}.$$

- This problem describes the distribution of the exit time of Brownian motion from the unit ball D. Let $f \colon \mathbf{R}^n \to \mathbf{R}$ be an infinitely differentiable function. Let μ denote the probability measure that is uniformly distributed in ∂D. Show that,

$$\mathbf{E}_x f(B(T)) = \int_{\partial D} \frac{1 - \|x\|^2}{\|x - y\|^n} f(y)\, d\mu(y), \qquad \forall\, x \in D.$$

(Hint: let $k_y(x) := \frac{1-\|x\|^2}{\|x-y\|^n}$ for any $x, y \in D$ with $x \neq y$. You can freely use the fact that $\Delta k_y(x) = 0$, where Δ is the Laplacian on $\mathbf{R}^n$. Define

$$v(x) := \begin{cases} \int_{\partial D} \frac{1-\|x\|^2}{\|x-y\|^n} f(y) d\mu(y) & , \text{if } x \in D \\ f(x) & , \text{if } x \in \partial D. \end{cases}$$

Is it true that $\Delta v(x) = 0$? You do not need to justify moving the Laplacian inside the integral. Moreover, you may freely use that a solution of the Dirichlet problem is unique.)

- Let $0 < a < b < \infty$. Let U denote the annulus

  $$U := \{x \in \mathbf{R}^n : a < \|x\| < b\}.$$

  Define $u : \mathbf{R}^n \setminus \{0\} \to \mathbf{R}$ by

  $$u(x) := \begin{cases} \|x\| & , \text{if } n = 1 \\ \log \|x\| & , \text{if } n = 2 \\ \|x\|^{2-n} & , \text{if } n \geq 3. \end{cases}$$

  Verify that $\Delta u(x) = 0$ for all $x \neq 0$. Define

  $$T_a := \inf\{t > 0 : \|B(t)\| = a\}, \qquad T_b := \inf\{t > 0 : \|B(t)\| = b\}.$$

  Show that, for any $x \in U$,

  $$\mathbf{P}_x(T_a < T_b) = \frac{u(b, 0, \ldots, 0) - u(x)}{u(b, 0, \ldots, 0) - u(a, 0, \ldots, 0)}.$$

  (Hint: Since u is harmonic, how is $u(x)$ related to $\mathbf{E}_x u(B(T_a \wedge T_b))$?)
- Let $x \in \mathbf{R}^n$ with $\|x\| > a > 0$. Conclude that

  $$\mathbf{P}_x(T_a < \infty) = \begin{cases} 1 & , \text{if } n \leq 2 \\ (a/\|x\|)^{n-2} & , \text{if } n \geq 3. \end{cases}$$

  (This is analogous to our recurrience/transience results for the simple random walk on $\mathbf{Z}^n$.)

Solution.

For any $x \in \overline{D}$, we know from Theorem 8.40 in the notes that

$$g(x) := \mathbf{E}_x f(B(T))$$

solves the Dirichlet problem, i.e. $\Delta g(x) = 0$ for all $x \in D$ and $g(x) = f(x)$ for all $x \in \partial D$. So, by uniqueness of solution of the Dirichlet problem, since $g(x) = v(x)$ for all $x \in \partial D$, it suffices to verify that $\Delta v(x) = 0$ for all $x \in D$. We can justify this identity by moving the derivatives inside the integral:

$$\Delta v(x) = \int_{\partial D} \Delta_x \frac{1 - \|x\|^2}{\|x - y\|^n} f(y) d\mu(y) = \int_{\partial D} 0 d\mu(y) = 0.$$

Now, since u is harmonic, we have from Theorem 8.40 that

$$\begin{aligned} u(x) &= \mathbf{E}_x u(B(T_a \wedge T_b)) = u(a, 0, \ldots, 0)\mathbf{P}(T_a < T_b) + u(b, 0, \ldots, 0)\mathbf{P}(T_a > T_b) \\ &= u(a, 0, \ldots, 0)\mathbf{P}(T_a < T_b) + u(b, 0, \ldots, 0)[1 - \mathbf{P}(T_a < T_b)]. \end{aligned}$$

Solving for $\mathbf{P}(T_a < T_b)$ gives

$$\mathbf{P}(T_a < T_b) = \frac{u(x) - u(b,0,\ldots,0)}{u(a,0,\ldots,0) - u(b,0,\ldots,0)}.$$

Letting $b \to \infty$, we have $T_b \to \infty$ as well, so that

$$\mathbf{P}(T_a < T_b) = \lim_{b\to\infty} \frac{u(x) - u(b,0,\ldots,0)}{u(a,0,\ldots,0) - u(b,0,\ldots,0)}.$$

When $n = 1$, we have

$$\lim_{b\to\infty} \frac{u(x) - u(b)}{u(a) - u(b)} = \lim_{b\to\infty} \frac{|x| - |b|}{|a| - |b|} = 1.$$

When $n = 2$, we have

$$\lim_{b\to\infty} \frac{u(x) - u(b,0)}{u(a,0) - u(b,0)} = \lim_{b\to\infty} \frac{\log\|x\| - \log\|b\|}{\log\|a\| - \log\|b\|} = 1.$$

When $n \geq 3$, we have When $n = 1$, we have

$$\lim_{b\to\infty} \frac{u(x) - u(b,0,\ldots,0)}{u(a,0,\ldots,0) - u(b,0,\ldots,0)} = \lim_{b\to\infty} \frac{\|x\|^{2-n} - \|b\|^{2-n}}{\|a\|^{2-n} - \|b\|^{2-n}} = \frac{\|x\|^{2-n}}{\|a\|^{2-n}}.$$

4. Question 4

Suppose we have n finite, irreducible transition matrices $P_1, \ldots, P_n$ with n corresponding state spaces $\Omega_1, \ldots, \Omega_n$ and stationary distributions $\pi_1, \ldots, \pi_n$. Define $\Omega := \Omega_1 \times \cdots \times \Omega_n$. Let $w = (w_1, \ldots, w_n)$ be a probability distribution on $\{1, \ldots, n\}$. Consider the (discrete-time) Markov chain on Ω that, at each step, selects coordinate $1 \leq j \leq n$ with probability w_j, and then changes the state of the chain only in coordinate j according to P_j. The transition matrix P for this chain is then

$$P(x,y) := \sum_{j=1}^{n} w_j P_j(x_j, y_j) \prod_{\substack{k\in\{1,\ldots,n\}:\\ k\neq j}} 1_{\{x_k = y_k\}}, \quad \forall\, x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \Omega.$$

(You do not have to prove this.) For each $1 \leq j \leq n$, let $f_j \colon \Omega_j \to \mathbf{R}$, and define the tensor product function $f_1 \otimes f_2 \otimes \cdots \otimes f_n \colon \Omega \to \mathbf{R}$ by

$$(f_1 \otimes f_2 \otimes \cdots \otimes f_n)(x_1, \ldots, x_n) := \prod_{j=1}^{n} f_j(x_j), \qquad \forall\, x = (x_1, \ldots, x_n) \in \Omega.$$

Let $\pi := \pi_1 \otimes \cdots \otimes \pi_n$ (so that π is a probability distribution on Ω, where we consider each π_j to be a function on Ω_j in order to use the tensor product definition). Then π is stationary for P (you do not have to prove this). Assume that, for any $1 \leq j \leq n$, the transition matrix P_j has an eigenfunction $f_j \in \mathbf{R}^{\Omega_j}$ with eigenvalue λ_j.

- Show that the function $f := f_1 \otimes \cdots \otimes f_n$ is an eigenfunction of P with eigenvalue $\sum_{j=1}^{n} w_j \lambda_j$.
- Assume that, for each $1 \leq j \leq n$, $\mathcal{B}_j$ is an orthonormal basis for $\mathbf{R}^{\Omega_j}$ with respect to the inner product $\langle \cdot, \cdot \rangle_{\pi_j}$. Show that

$$\mathcal{B} := \{f_1 \otimes f_2 \otimes \cdots \otimes f_n \colon f_j \in \mathcal{B}_j,\ \forall\, 1 \leq j \leq n\}$$

is an orthonormal basis with respect to the inner product $\langle \cdot, \cdot \rangle_{\pi}$.

- If P_j has spectral gap γ_j for all $1 \leq j \leq n$, show that P has spectral gap

$$\gamma := \min_{1 \leq j \leq n} w_j \gamma_j.$$

- For any $1 \leq j \leq n$, let $\Omega_j := \{-1, 1\}$ and let $P_j(a, b) = 1$ for all $a, b \in \{-1, 1\}$ with $a \neq b$ and $P_j(a, b) = 0$ otherwise. Let $w := (1/n, \ldots, 1/n)$. Then P on Ω corresponds to the simple random walk on the discrete hypercube $\{-1, 1\}^n$. (You do not have to prove this.) Conclude that the spectral gap for the Markov chain P on Ω is

$$\gamma = 2/n.$$

- Using the result from Exam 2, then give a bound on the mixing time of the Markov chain P on $\Omega = \{-1, 1\}^n$.

Solution. By assumption $P_j f_j(x_j) = \lambda_j f_j(x_j)$, i.e.

$$\sum_{y_j \in \Omega_j} P_j(x_j, y_j) f_j(y_j) = \lambda_j f_j(x_j). \qquad (*)$$

Now, fix $x = (x_1, \ldots, x_n) \in \Omega$. Then

$$\begin{aligned}
P(f_1 \otimes \cdots \otimes f_n)(x) &= \sum_{y \in \Omega} P(x, y)(f_1 \otimes \cdots \otimes f_n)(y) \\
&= \sum_{y \in \Omega} \sum_{j=1}^{n} w_j P_j(x_j, y_j) \prod_{\substack{k \in \{1, \ldots, n\}: \\ k \neq j}} 1_{\{x_k = y_k\}} \prod_{i=1}^{n} f_i(y_i) \\
&= \sum_{j=1}^{n} w_j \sum_{y_j \in \Omega_j} P_j(x_j, y_j) \sum_{y_\ell \in \Omega_\ell \ \forall \ell \in \{1, \ldots, n\} \setminus \{j\}} \prod_{\substack{k \in \{1, \ldots, n\}: \\ k \neq j}} 1_{\{x_k = y_k\}} \prod_{i=1}^{n} f_i(y_i) \\
&= \sum_{j=1}^{n} w_j \sum_{y_j \in \Omega_j} P_j(x_j, y_j) \prod_{i=1}^{n} f_i(x_i) = \sum_{j=1}^{n} w_j \sum_{y_j \in \Omega_j} P_j(x_j, y_j) f_j(y_j) \prod_{\substack{k \in \{1, \ldots, n\}: \\ k \neq j}} f_k(x_k) \\
&\overset{(*)}{=} \sum_{j=1}^{n} w_j \lambda_j f_j(x_j) \prod_{\substack{k \in \{1, \ldots, n\}: \\ k \neq j}} f_k(x_k) = \sum_{j=1}^{n} w_j \lambda_j \prod_{k=1}^{n} f_k(x_k) = \sum_{j=1}^{n} w_j \lambda_j (f_1 \otimes \cdots \otimes f_n)(x).
\end{aligned}$$

By definition of π, if $g = g_1 \otimes \cdots \otimes g_n \in \mathcal{B}$, then

$$\langle f, g \rangle_\pi = \prod_{i=1}^{n} \langle f_i, g_i \rangle_{\pi_i}.$$

This equality implies: if $f, g \in \mathcal{B}$ with $f \neq g$, then $\langle f, g \rangle_\pi = 0$ (since at least one term in the product on the right is zero) and $\langle f, f \rangle_\pi = 1$ (since the term on the right is then a product of ones). So, $\mathcal{B}$ forms an orthonormal set. To see that $\mathcal{B}$ is a basis, note that the set of all functions on Ω has dimension $\prod_{i=1}^{n} |\Omega_i|$, and $\mathcal{B}$ has the same number of elements. Therefore, $\mathcal{B}$ is a basis.

The set of all eigenvalues of P is

$$\Lambda := \{\sum_{j=1}^{n} w_j\lambda_j \colon \lambda_j \text{ is an eigenvalue of } P_j \ forall\ 1 \leq j \leq n\}.$$

Since $\lambda_j \leq 1$ for all $1 \leq j \leq n$, the largest element of Λ is $\sum_{j=1}^{n} w_j = 1$. The next largest element of Λ then correspond to all choices of λ_j being equal to one, except one (the choice of j with the smallest (weighted) spectral gap). Let $1 \leq j' \leq n$ such that $w_{j'}\gamma_{j'} = \min_{1\leq j\leq n} w_j\gamma_j$. That is, the second largest eigenvalue of P is

$$w_{j'}[1-\gamma_{j'}] + \sum_{j\in\{1,\ldots,n\}\colon j\neq j'} w_j = 1 - \gamma_{j'}w_{j'}.$$

So, P has spectral gap equal to 1 minus this number, i.e.

$$\gamma = w_{j'}\gamma_{j'} = \min_{1\leq j\leq n} w_j\gamma_j.$$

By assumption, the eigenvalues of $P_j = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ are 1 and -1, so $\gamma_j = 2$ for all $1 \leq j \leq n$. With $w = (1/n, \ldots, 1/n)$, the previous problem shows that $\gamma = (1/n)(2) = 2/n$.

Finally, the result from exam 2 shows that $\gamma = O(n^2)$.

5. Question 5

This problem is a continuation of the mixing time problem from Exam 2. Suppose we have a finite, irreducible Markov chain with state space Ω, transition matrix P, and stationary distribution π. For any $0 < p < \infty$, and for any $f \in \mathbf{R}^{\Omega}$, define

$$\|f\|_{p,\pi} := \Big(\sum_{x\in\Omega} |f(x)|^p \, \pi(x)\Big)^{1/p}, \qquad \mathbf{E}_{\pi} f := \sum_{x\in\Omega} f(x)\pi(x).$$

Another way to bound the mixing time is to give a bound on the logarithmic-Sobolev constant of the chain (or log-Sobolev constant). Define $\alpha \geq 0$ so that $1/\alpha$ is the smallest constant $c > 0$ such that, for all $f \in \mathbf{R}^{\Omega}$ with $\|f\|_{2,\pi} \neq 0$,

$$\sum_{x\in\Omega} |f(x)|^2 \Big[\log\Big(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\Big)\Big]\pi(x) \leq c \cdot \frac{1}{2}\sum_{x,y\in\Omega} |f(x)-f(y)|^2 \, P(x,y)\pi(x). \qquad (*)$$

(If no such c exists, $\alpha := 0$.) (Note that $(*)$ is dilation invariant, i.e. if $f \in \mathbf{R}^{\Omega}$ satisfies $(*)$, then tf also satisfies $(*)$ for all $t > 0$.) You can freely use the following bound:

$$\max_{x\in\Omega} \|H_t(x,\cdot) - \pi(\cdot)\|_{\mathrm{TV}} \leq \sqrt{\frac{1}{2}\log\Big(\frac{1}{\min_{y\in\Omega}\pi(y)}\Big)}\, e^{-\alpha t}, \qquad \forall\, t > 0.$$

- Give an upper bound on the mixing time of a Markov chain in terms of the log-Sobolev constant α.
- Additionally assume the Markov chain is reversible. Show that the spectral gap γ can be equivalently defined so that $1/\gamma$ is the smallest constant $b > 0$ such that, for all $f \in \mathbf{R}^{\Omega}$,

$$\frac{1}{2}\sum_{x,y\in\Omega} |f(x)-f(y)|^2 \, \pi(x)\pi(y) \leq b \cdot \frac{1}{2}\sum_{x,y\in\Omega} |f(x)-f(y)|^2 \, P(x,y)\pi(x).$$

(If no such b exists, $\gamma := 0$.) (Hint: the left side is the variance of f.)

- Additionally assume the Markov chain is reversible. Show that $2\alpha \leq \gamma$. (Hint: consider $f = 1 + \varepsilon g$ in the definition of the log-Sobolev constant, and let $\varepsilon \to 0^+$.)
- Let us re-use the notation of the previous problem to define $P_1, \ldots, P_n$, $\Omega_1, \ldots, \Omega_n$, $\pi_1, \ldots, \pi_n$, Ω, w and P. Show: if P_j has logarithmic Sobolev constant α_j for all $1 \leq j \leq n$, then P has logarithmic Sobolev constant

$$\alpha := \min_{1 \leq j \leq n} w_j \alpha_j.$$

(Hint: just consider the case $n = 2$. Let $f\colon \Omega_1 \times \Omega_2 \to \mathbf{R}$. Define $F\colon \Omega_2 \to \mathbf{R}$ by $F(x_2) := \sqrt{\sum_{x_1 \in \Omega_1} |f(x_1, x_2)|^2 \pi_1(x_1)}$, $\forall\, x_2 \in \Omega_2$. Observe that

$$\sum_{x \in \Omega} |f(x)|^2 \Big[\log\Big(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\Big)\Big]\pi(x) = \sum_{x_2 \in \Omega_2} |F(x_2)|^2 \log\Big(\frac{|F(x_2)|^2}{\|F\|_{2,\pi_2}^2}\Big)\pi_2(x_2)$$
$$+ \sum_{(x_1,x_2) \in \Omega_1 \times \Omega_2} |f(x_1, x_2)|^2 \Big[\log\Big(\frac{|f(x_1, x_2)|^2}{|F(x_2)|^2}\Big)\Big]\pi(x_1, x_2).$$

Apply the definition of α_2 to the first term, and apply the definition of α_1 to the second term. Then, observe that $|F(a) - F(b)| \leq \|f(\cdot, a) - f(\cdot, b)\|_{2,\pi_1}$ for all $a, b \in \Omega_2$ to get a bound on the first term in terms of $f(x_1, \cdot)$.)

- For any $1 \leq j \leq n$, let $\Omega_j := \{-1, 1\}$ and let $P_j(a, b) = 1$ for all $a, b \in \{-1, 1\}$ with $a \neq b$ and $P_j(a, b) = 0$ otherwise. Let $w := (1/n, \ldots, 1/n)$. Then P on Ω corresponds to the simple random walk on the discrete hypercube $\{-1, 1\}^n$. (You do not have to prove this.) You can freely use that $\alpha_j = 1$ for all $1 \leq j \leq n$. Conclude that the log-Sobolev constant for this Markov chain P on Ω is

$$\alpha = 1/n.$$

How does the resulting bound on the mixing time of this Markov chain compare with the bound from the previous problem?

Solution. Expanding out the right side, using reversibility ($\pi(x)P(x, y) = \pi(y)P(y, x)$), and that the sum of any row of P is 1,

$$\frac{1}{2}\sum_{x,y \in \Omega} |f(x) - f(y)|^2 P(x, y)\pi(x) = \sum_{x,y \in \Omega} |f(x)|^2 P(x, y)\pi(x) - \sum_{x,y \in \Omega} f(x)f(y)P(x, y)\pi(x)$$
$$= \sum_{x \in \Omega} |f(x)|^2 \pi(x) - \sum_{x \in \Omega} f(x)(Pf)(x)\pi(x).$$

So, we need to show that, for all $f \in \mathbf{R}^{\Omega}$,

$$\mathrm{Var}_\pi f \leq \frac{1}{\gamma}\Big(\sum_{x \in \Omega} |f(x)|^2 \pi(x) - \sum_{x \in \Omega} f(x)(Pf)(x)\pi(x)\Big),$$

with equality when f is the eigenfunction of P with the second largest eigenvalue. Let $f \in \mathbf{R}^{\Omega}$ and let $f_1, \ldots, f_\Omega$ be eigenfunctions of P with eigenvalues $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n \geq -1$.

Then $f = \sum_{i=1}^{|\Omega|} c_i f_i$ where $c_i = \langle f, f_i \rangle_\pi$ for all $1 \le i \le |\Omega|$. Then

$$\mathrm{Var}_\pi f = \sum_{i=2}^{|\Omega|} c_i^2,$$

$$\sum_{x\in\Omega} |f(x)|^2 \pi(x) - \sum_{x\in\Omega} f(x)(Pf)(x)\pi(x) = \sum_{i=1}^{|\Omega|} c_i^2 - \sum_{i=1}^{|\Omega|} \lambda_i c_i^2 = \sum_{i=1}^{|\Omega|}(1-\lambda_i)c_i^2 = \sum_{i=2}^{|\Omega|}(1-\lambda_i)c_i^2.$$

The penultimate equality used $\lambda_1 = 1$. Since $\lambda_2 \ge \lambda_i$ for all $2 \le i \le |\Omega|$, we have $1 - \lambda_i \ge 1 - \lambda_2 = \gamma$ also, i.e. $\frac{1}{\gamma}(1-\lambda_i) \ge 1$ for all $2 \le i \le n$, so that

$$\frac{1}{\gamma}\sum_{i=2}^{|\Omega|}(1-\lambda_i)c_i^2 \ge \sum_{i=2}^{|\Omega|} c_i^2,$$

And equality holds when $c_i = 0$ for all $3 \le i \le n$. That is,

$$\mathrm{Var}_\pi f \le \frac{1}{\gamma}\Big(\sum_{x\in\Omega} |f(x)|^2 \pi(x) - \sum_{x\in\Omega} f(x)(Pf)(x)\pi(x)\Big),$$

with equality when $f = f_2$.

Now, consider $f = 1 + \varepsilon g$. The right side of $(*)$ is

$$\frac{1}{2}\sum_{x,y\in\Omega} |f(x)-f(y)|^2 P(x,y)\pi(x) = \varepsilon^2 \frac{1}{2}\sum_{x,y\in\Omega} |g(x)-g(y)|^2 P(x,y)\pi(x)$$

We now examine the left side of $(*)$. We have

$$|f(x)|^2 = (1+\varepsilon g(x))^2 = 1 + 2\varepsilon g(x) + \varepsilon^2 (g(x))^2.$$
$$\|f\|_{2,\pi}^2 = \|1+\varepsilon g\|_{2,\pi}^2 = 1 + 2\varepsilon \mathbf{E}_\pi g + \varepsilon^2 \|g\|_{2,\pi}^2.$$
$$\frac{|f(x)|^2}{\|f\|_{2,\pi}^2} = \frac{1 + 2\varepsilon g(x) + \varepsilon^2 (g(x))^2}{1 + 2\varepsilon \mathbf{E}_\pi g + \varepsilon^2 \|g\|_{2,\pi}^2}.$$

Then, using $\log(1 + a\varepsilon + b\varepsilon^2) = a\varepsilon + b\varepsilon^2 - (a\varepsilon + b\varepsilon^2)^2/2 + O(\varepsilon^3) = a\varepsilon + \varepsilon^2(b - a^2/2) + O(\varepsilon^3)$

$$\begin{aligned}\log\Big(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\Big) &= \log\Big(1 + 2\varepsilon g(x) + \varepsilon^2 (g(x))^2\Big) - \log\Big(1 + 2\varepsilon \mathbf{E}_\pi g + \varepsilon^2 \|g\|_{2,\pi}^2\Big)\\ &= 2\varepsilon(g(x) - \mathbf{E}_\pi g) + \varepsilon^2((g(x))^2 - 2(g(x))^2 + 2(\mathbf{E}_\pi g)^2 - \|g\|_{2,\pi}^2) + O(\varepsilon^3)\\ &= 2\varepsilon(g(x) - \mathbf{E}_\pi g) + \varepsilon^2(-(g(x))^2 + 2(\mathbf{E}_\pi g)^2 - \|g\|_{2,\pi}^2) + O(\varepsilon^3).\end{aligned}$$

Therefore,

$$\begin{aligned}&|f(x)|^2\Big[\log\Big(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\Big)\Big]\\ &= \Big(1 + 2\varepsilon g(x) + \varepsilon^2 (g(x))^2\Big)\Big(2\varepsilon(g(x) - \mathbf{E}_\pi g) + \varepsilon^2(-(g(x))^2 + 2(\mathbf{E}_\pi g)^2 - \|g\|_{2,\pi}^2) + O(\varepsilon^3)\Big)\\ &= 2\varepsilon(g(x) - \mathbf{E}_\pi g) + \varepsilon^2\Big(-(g(x))^2 + 2(\mathbf{E}_\pi g)^2 - \|g\|_{2,\pi}^2 + 4(g(x))^2 - 4g(x)\mathbf{E}_\pi g\Big)\\ &= 2\varepsilon(g(x) - \mathbf{E}_\pi g) + \varepsilon^2\Big(3(g(x))^2 + 2(\mathbf{E}_\pi g)^2 - \|g\|_{2,\pi}^2 - 4g(x)\mathbf{E}_\pi g\Big) + O(\varepsilon^3)\end{aligned}$$

And

$$\sum_{x\in\Omega}|f(x)|^2\left[\log\Big(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\Big)\right]\pi(x)=\varepsilon^2\Big(3\|g\|_{2,\pi}^2+2(\mathbf{E}_\pi g)^2-\|g\|_{2,\pi}^2-4(\mathbf{E}_\pi g)^2\Big)+O(\varepsilon^3)$$
$$=\varepsilon^2 2(\|g\|_{2,\pi}^2-\mathbf{E}_\pi g)^2+O(\varepsilon^3)=\varepsilon^2 2\mathrm{Var}(g)+O(\varepsilon^3)$$

So, letting $\varepsilon\to 0^+$ in the log-Sobolev inequality, we arrive at

$$2\mathrm{Var}(g)\leq\frac{1}{\alpha}\frac{1}{2}\sum_{x,y\in\Omega}|g(x)-g(y)|^2P(x,y)\pi(x).$$

From the previous part of the problem, we conclude that $\frac{1}{2\alpha}\geq\frac{1}{\gamma}$, i.e. $\gamma\geq 2\alpha$.

For the next part of the problem, note that it suffices to prove the case $n=2$ and then iterate. As in the hint, let $F(x_2):=\sqrt{\sum_{x_1\in\Omega_1}|f(x_1,x_2)|^2\pi_1(x_1)}$, $\forall\ x_2\in\Omega_2$, and observe that

$$\sum_{x\in\Omega}|f(x)|^2\left[\log\Big(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\Big)\right]\pi(x)=\sum_{x_2\in\Omega_2}|F(x_2)|^2\log\Big(\frac{|F(x_2)|^2}{\|F\|_{2,\pi_2}^2}\Big)\pi_2(x_2)$$
$$+\sum_{(x_1,x_2)\in\Omega_1\times\Omega_2}|f(x_1,x_2)|^2\left[\log\Big(\frac{|f(x_1,x_2)|^2}{|F(x_2)|^2}\Big)\right]\pi(x_1,x_2).$$

We then apply the definition of α_2 to the first term, and apply the definition of α_1 to the second term (noting that $|F(x_2)|^2=\|f(\cdot,x_2)\|_{2,\pi_1}^2$) to get

$$\sum_{x\in\Omega}|f(x)|^2\left[\log\Big(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\Big)\right]\pi(x)\leq\frac{1}{\alpha_2}\frac{1}{2}\sum_{x_2,y_2\in\Omega_2}|F(x_2)-F(y_2)|^2P_2(x_2,y_2)\pi_2(x_2)$$
$$+\frac{1}{\alpha_1}\frac{1}{2}\sum_{x_2\in\Omega_2}\Big(\sum_{x_1,y_1\in\Omega_1}|f(x_1,x_2)-f(y_1,x_2)|^2P_1(x_1,y_2)\pi_1(x_1)\Big)\pi_2(x_2).$$

Also as suggested in the hint, observe that for all $a,b\in\Omega_2$, by the (reverse) triangle inequality,

$$|F(a)-F(b)|=\Big|\|f(\cdot,a)\|_{2,\pi_1}-\|f(\cdot,b)\|_{2,\pi_1}\Big|\leq\|f(\cdot,a)-f(\cdot,b)\|_{2,\pi_1}.$$

So, we can apply this inequality to the first term above to get

$$\sum_{x\in\Omega}|f(x)|^2\left[\log\Big(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\Big)\right]\pi(x)\leq\frac{1}{\alpha_2}\frac{1}{2}\sum_{x_2,y_2\in\Omega_2}\|f(\cdot,x_2)-f(\cdot,y_2)\|_{2,\pi_1}^2P_2(x_2,y_2)\pi_2(x_2)$$
$$+\frac{1}{\alpha_1}\frac{1}{2}\sum_{x_2\in\Omega_2}\Big(\sum_{x_1,y_1\in\Omega_1}|f(x_1,x_2)-f(y_1,x_2)|^2P_1(x_1,y_2)\pi_1(x_1)\Big)\pi_2(x_2)$$
$$=\frac{1}{w_2\alpha_2}\frac{1}{2}\sum_{x_1\in\Omega_1}\Big(\sum_{x_2,y_2\in\Omega_2}|f(x_1,x_2)-f(x_1,y_2)|^2w_2P_2(x_2,y_2)\pi_2(x_2)\Big)\pi_1(x_1)$$
$$+\frac{1}{w_1\alpha_1}\frac{1}{2}\sum_{x_2\in\Omega_2}\Big(\sum_{x_1,y_1\in\Omega_1}|f(x_1,x_2)-f(y_1,x_2)|^2w_1P_1(x_1,y_2)\pi_1(x_1)\Big)\pi_2(x_2).$$

By definition of P, we then have

$$\sum_{x\in\Omega} |f(x)|^2 \left[\log\left(\frac{|f(x)|^2}{\|f\|_{2,\pi}^2}\right)\right]\pi(x) \le \max\left(\frac{1}{w_2\alpha_2}, \frac{1}{w_1\alpha_1}\right)\frac{1}{2}\sum_{x,y\in\Omega} |f(x)-f(y)|^2 P(x,y)\pi(x).$$

The result follows.

Finally, the mixing time bound we get from this problem is $O(n\log n)$, which is much smaller than the $O(n^2)$ bound we got from the previous problem.

More specifically, using the inequality

$$\max_{x\in\Omega} \|H_t(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}} \le \sqrt{\frac{1}{2}\log\left(\frac{1}{\min_{y\in\Omega}\pi(y)}\right)}e^{-\alpha t}, \qquad \forall\, t>0,$$

to get a mixing time bound, it suffices to solve for t such that

$$\sqrt{\frac{1}{2}\log\left(\frac{1}{\min_{y\in\Omega}\pi(y)}\right)}e^{-\alpha t} \le 1/4.$$

Solving for t, we get $e^{-2\alpha t} \le 1/8\log\left(\frac{1}{\min_{y\in\Omega}\pi(y)}\right)$, i.e. $-2\alpha t \le \log\left(1/8\log\left(\frac{1}{\min_{y\in\Omega}\pi(y)}\right)\right)$, i.e.

$$t \ge \frac{1}{2\alpha}\log\left(8\log\left(\frac{1}{\min_{y\in\Omega}\pi(y)}\right)\right).$$

We conclude that

$$t_{\mathrm{mix}} \le \frac{1}{2\alpha}\log\left(8\log\left(\frac{1}{\min_{y\in\Omega}\pi(y)}\right)\right).$$

Substituting in $\pi(y) = 2^{-n}$ for all $y\in\Omega$ and $\alpha = 1/n$, we get

$$t_{\mathrm{mix}} \le \frac{n}{2}\log\left(8\log 2^n\right) = \frac{n}{2}\log(8n\log 2).$$