Supplementary Material for "Deterministic Sampling of Expensive Posteriors Using Minimum Energy Designs"

V. Roshan Joseph
Georgia Institute of Technology, Atlanta, GA
Email: roshan@gatech.edu

Dianpeng Wang
Beijing Institute of Technology

Li Gu
Amazon.com, Inc.

Shiji Lyu
Princeton University

Rui Tuo
Texas A& M University

1 Main Theorem

Let $\mathcal{X} = [0,1]^p$ and $q(x) : \mathcal{X} \to (0, +\infty)$ be a continuous function. Given a design $D = \{x_1, \ldots, x_n\} \subset \Omega, n \geq 2$, for $s > 0$, define

$$E_s(D) = \max_{i,j} \frac{q(x_i)q(x_j)}{d_s(x_i, x_j)}, \tag{1}$$

with

$$d_s(u, v) = \left(\sum_{i=1}^{p} |u_i - v_i|^s \right)^{1/s},$$

where $u = (u_1, \ldots, u_p)$ and $v = (v_1, \ldots, v_p)$. Consider the minimum energy design D_0 under d_s satisfying

$$E_s(D_0) = \min_{\substack{D \subset \Omega \\ \mathrm{card}(D) = n}} E_s(D), \tag{2}$$

where $\mathrm{card}(D)$ denotes the cardinality of the set D.

We now introduce the index of a design. Fix $0 < s < \infty$. For a design $D = \{x_1, \ldots, x_n\}$, define its index, denoted by $IN(D)$, to be the number of pairs $(x_k, x_l), 1 \leq k < l \leq n$, with the greatest value of $(q(x_i)q(x_j))/d_s(x_i, x_j)$ over all $i \neq j$, i.e.,

$$IN(D) = \mathrm{card} \left\{ (x_k, x_l) : 1 \leq k < l \leq n, \frac{q(x_k)q(x_l)}{d_s(x_k, x_l)} = \min_{1 \leq i < j \leq n} \frac{q(x_i)q(x_j)}{d_s(x_i, x_j)} \right\}.$$

We are particularly interested in the minimum energy designs with the smallest index, because such designs are more space-filling than regular minimum energy designs.

Theorem 1. Suppose q is Lipschitz continuous, i.e., $|q(x) - q(y)| \leq L\|x - y\|$, for $x, y \in \mathcal{X}$ and a constant $L > 0$, where $\|\cdot\|$ denotes the Euclidean distance. Let $D^* = \{x_1^*, \ldots, x_n^*\}$ be an n-point minimum energy design under d_s with the smallest index and $\mathscr{B}$ be the Borel σ-algebra of $\mathcal{X}$. Define the following probability measures on $(\mathcal{X}, \mathscr{B})$:

$$\mathcal{P}_n(A) = \frac{\text{card}\{x_i^* : 1 \leq i \leq n, x_i^* \in A\}}{n}, \text{ for any } A \in \mathscr{B}. \tag{3}$$

Then there exists a probability measure $\mathcal{P}$ such that $\mathcal{P}_n$ converges to $\mathcal{P}$ weakly for all fixed $s \in (0, +\infty)$ as $n \to \infty$. Moreover, $\mathcal{P}$ has a density f over $\mathcal{X}$ with $f(x) \propto 1/q^{2p}(x)$.

2 Comparison of Measure

The proof of Theorem 1 relies on some results in measure theory, stated by Theorem 2 and Theorem 3 below. First we introduce some necessary notation.

Let P be a probability measure on $(\mathcal{X}, \mathscr{B})$, satisfying $P(\partial\mathcal{X}) = 0$. Let $Cu(x, l)$ denotes the open cube centered at x with side length $2l$, i.e.,

$$Cu(x, l) := \left\{ y \in \mathbb{R}^p : \max_{1 \leq i \leq p} |y_i - x_i| \leq l \right\},$$

where x_i, y_i denote the ith entry of x and y respectively. We will use the following Condition 1 for P. Denote the set of interior points of $\mathcal{X}$ by $\mathcal{X}^\circ$.

Condition 1. Let $g : \mathcal{X} \to (0, +\infty)$ be a continuous function. For all $x_1, x_2 \in \mathcal{X}^\circ$,

$$\lim_{r \downarrow 0} \frac{P(Cu(x_1, \frac{r}{g(x_1)}))}{P(Cu(x_2, \frac{r}{g(x_2)}))} = 1. \tag{4}$$

The main aim of this section is to prove Theorem 2.

Theorem 2. Suppose P is a Borel probability measure on $\mathcal{X}$ with $P(\partial\mathcal{X}) = 0$ and P satisfies Condition 1. Then P has a density function f with respect to the Lebesgue measure m on $\mathcal{X}$, i.e., $P(E) = \int_E f \mathrm{d}m$ for all Borel set $E \subset \mathcal{X}$. Moreover, $f \propto g^p$, m-almost everywhere on $\mathcal{X}$.

We will first prove that P is absolutely continuous with respect to m and then find the density function. To prove the absolute continuity, we find that it is more convenient to work with the following Condition 2, which is weaker than Condition 1.

Condition 2. There exist positive constants c_1, c_2 and c_3, such that

$$\liminf_{r \downarrow 0} \frac{P(B(x_1, c_1 r))}{P(B(x_2, c_2 r))} \geq c_3, \tag{5}$$

where $B(x, R) = \{y \in \mathbb{R}^p : d_2(x, y) < R\}$ is the Euclidean open ball centered at x.

Theorem 3. Suppose P is a Borel probability measure on $\mathcal{X}$ with $P(\partial \mathcal{X}) = 0$ and P satisfies Condition 2. Then P has a density function f, with respect to the Lebesgue measure m on $\mathcal{X}$.

The proof of Theorem 2 is based on Theorem 3. The proof of Theorem 3 is accomplished in two steps: the first step, formalized as Lemma 2, is to compare P with m around a point $x \in \mathcal{X}^\circ$; the second step, given in Lemma 3, is to compare P with m on an arbitrary rectangular region in $\mathcal{X}$.

Because $\mathcal{X} \subset \mathbb{R}^p$, P can also be regarded as a probability measure on $\mathbb{R}^p$. For notational simplicity, for any Borel set $E \subset \mathbb{R}^d$, $P(E \cap \mathcal{X})$ will still be denoted as $P(E)$. For fixed $r > 0$, define

$$\phi_r(x) := P(B(x, r)).$$

We will need some measurability properties of ϕ_r later, which is ensured by Lemma 1.

Lemma 1. For fixed $r > 0$, and any Borel probability measure P on $\mathcal{X}$, $\phi_r(x)$ is lower semi-continuous, in the sense that $(\phi_r)^{-1}(\alpha, +\infty)$ is open for all $\alpha \in \mathbb{R}$.

Proof. Fix $\alpha \in \mathbb{R}, x \in \mathbb{R}^d$, with $\phi_r(x) > \alpha$. It suffices to show that there is an open ball centered at x and contained in $(\phi_r)^{-1}(\alpha, +\infty)$.

First we show

$$\lim_{\delta \downarrow 0} P(B(x, r - \delta)) = \phi_r(x), \tag{6}$$

or equivalently, for all sequences $\delta_n \downarrow 0$,

$$\lim_{n \to \infty} P(B(x, r - \delta_n)) = \phi_r(x).$$

For such a sequence, denote $E_n = B(x, r - \delta_n)$ and $E = B(x, r)$. Then E_n is an increasing sequence of Borel sets converging to E. Hence $\lim_{n \to 0} P(E_n) = P(E)$, i.e., $\lim_{n \to \infty} P(B(x, r - \delta_n)) = \phi_r(x)$, and (6) follows.

Consequently, there exists $\delta_0 > 0$, such that $P(B(x, r - \delta_0)) > \alpha$. Clearly $y \in B(x, \delta_0)$ implies $B(y, r) \supset B(x, r - \delta_0)$, and thus $\phi_r(y) > \alpha$. Therefore, $B(x, \delta_0) \subset (\phi_r)^{-1}(\alpha, +\infty)$. The desired result follows. $\square$

Lemma 2 gives a comparison between the magnitude of P and m around an interior point of $\mathcal{X}$ in a limiting sense.

Lemma 2. Under the conditions of Theorem 3, for all $x \in \mathcal{X}^\circ$, we have

$$\limsup_{r \downarrow 0} \frac{\phi_r(x)}{r^p} \leq M < +\infty,$$

where

$$M = \frac{4}{c_3}\Big(\frac{8c_1}{c_2}\Big)^p$$

is a constant.

Proof. We will prove the result by showing the contrary will lead to a contradiction. Suppose that there exists $x_0 \in \mathcal{X}^\circ$, such that

$$\limsup_{r \downarrow 0} \frac{\phi_r(x_0)}{r^p} > M.$$

Then there exists a sequence $r_n \downarrow 0$ with

$$\frac{\phi_{c_2 r_n}(x_0)}{(c_2 r_n)^p} > M, \tag{7}$$

for all n. Consider function

$$f_n(x) = \frac{\phi_{c_1 r_n}(x)}{\phi_{c_2 r_n}(x_0)}, \tag{8}$$

which is Borel measurable on $\mathbb{R}^p$ according to Lemma 1. Let

$$F_n(x) = \inf_{k \geq n} f_k(x),$$

then F_n is a sequence of Borel measurable functions converging to

$$F(x) = \liminf_{n \to +\infty} f_n(x).$$

Hence by Egorov's Theorem (see Stein and Shakarchi, 2009, p. 33), there exists a Borel set $E \subset \mathcal{X}^\circ$, with $m(E) > m(\mathcal{X}^\circ) - 2^{-(p+1)} = 1 - 2^{-(p+1)}$ and on which F_n converges to F uniformly. Noticed by (5), we have $F(x) \geq c_3$. Thus we can choose $N \in \mathbb{N}^+$, such that $r_N < 1/(4c_1)$ and for any $x \in E, f_N(x) \geq F_N(x) \geq c_3/2$. Besides, there exists a positive integer l, such that $l < 1/(4c_1 r_N) \leq l + 1$.

We then partition $\mathcal{X}$ into l^p cubes, each of which has the form:

$$\prod_{i=1}^p [\frac{k_i}{l}, \frac{k_i+1}{l}], k_i \in \mathbb{Z}, 0 \leq k_i < l.$$

Denote these cubes by $Q_1, ..., Q_{l^p}$, and let $\mathcal{T}(Q_j)$ to be the cube with the same center as Q_j and half of the side length of Q_j, that is,

$$\mathcal{T}(\prod_{i=1}^{p}[\frac{k_i}{l}, \frac{k_i+1}{l}]) = \prod_{i=1}^{p}[\frac{k_i+1/4}{l}, \frac{k_i+3/4}{l}].$$

Clearly $\mathcal{T}(Q_j), j \in \mathcal{B}$ are disjoint, each with Lebesgue measure $(2l)^{-p}$. Thus the set $\mathcal{A} = \{1 \leq j \leq l^d p : \mathcal{T}(Q_j) \subset \mathcal{X} - E\}$ satisfies

$$(2l)^{-p}\text{card}(\mathcal{A}) \leq m(\mathcal{X} - E) < 2^{-(p+1)},$$

which implies $\text{card}(\mathcal{A}) < l^p/2$.

Let $\mathcal{B} = \{1 \leq j \leq l^d : j \notin \mathcal{A}\}$, then $\text{card}(\mathcal{B}) > l^p/2$. For each $j \in \mathcal{B}$, $\mathcal{T}(Q_j)$ intersects E. Pick $x_j \in E \cap \mathcal{T}(Q_j)$, then $f_N(x_j) \geq c_3/2$, which, together with (7) and (8), yields

$$P(B(x_j, c_1 r_N)) = P(B(x_0, c_2 r_N))f_N(x_j) > \frac{1}{2}Mc_3(c_2 r_N)^p. \tag{9}$$

It can be seen that $l < 1/(4c_1 r_N)$ from the choice of l. Thus

$$B(x_j, c_1 r_N) \subset B(x_j, \frac{1}{4l}).$$

Noting that the distance between $\mathcal{T}(Q_j)$ and the complement of Q_j° is exactly $1/(4l)$, and $x_j \in \mathcal{T}(Q_j)$, we have $B(x_j, 1/(4l)) \subset Q_j^\circ$. Hence $B(x_j, 1/(4l)), j \in \mathcal{B}$, are disjoint, and consequently $B(x_j, c_1 r_N), j \in \mathcal{B}$, are disjoint, which, together with (9) implies

$$\begin{aligned} P(\mathcal{X}) \geq \sum_{j \in \mathcal{B}} P(B(x_j, c_1 r_N)) &> \frac{1}{2}Mc_3(c_2 r_N)^p \text{card}(\mathcal{B}) \\ &> \frac{1}{2}\frac{4}{c_3}\Big(\frac{8c_1}{c_2}\Big)^p c_3 \Big(\frac{c_2}{4c_1(l+1)}\Big)^p \frac{1}{2} l^p \geq 1. \end{aligned}$$

This leads to a contradiction because P is a probability measure. □

Next we compare P with m on a cube with arbitrary size. The result is given in Lemma 3, which directly leads to the absolute continuity. The proof follows by a similar argument in that of Lemma 2. Specifically, we will partition the cubes into sufficiently small parts and approximate them with balls. Under this consideration we may use expressions parallel to that we have used before.

Lemma 3. Under the conditions of Theorem 3, for all cubes $Q = \prod_{i=1}^{p}[a_i, a_i + \alpha] \subset \mathcal{X}$,we have $P(Q) \leq Km(Q)$, here $K = 2M(2\sqrt{d})^p$ is a constant, M is the constant appeared in Lemma 2.

Proof. Fix $\epsilon > 0$. It suffices to show $P(Q) \leq Km(Q) + \epsilon$.

Define auxiliary function $\phi_r : \mathcal{X} \to (0, +\infty]$ as

$$\psi_r(x) = \sup_{0<R\leq r} \frac{\phi_R(x)}{R^p}.$$

In the light of Lemma 1, ψ_r is the superior of a family of lower-semi continuous functions. Hence it is lower-semi continuous as well, and thus is Borel measurable.

It is easy to see that

$$\psi_{1/n}(x) \to \limsup_{r\downarrow 0} \frac{\phi_r(x)}{r^p} =: \psi(x),$$

as $n \to \infty$, for all $x \in \mathbb{R}^p$, and in particular for all $x \in \mathcal{X}^\circ$. Therefore $\psi(x)$ is Borel measurable. Note that $P(\mathcal{X}^\circ) = P(\mathcal{X}) = 1$, because $P(\partial\mathcal{X}) = 0$. We apply Egorov's Theorem (see Stein and Shakarchi, 2009, p. 33) for the probability measure P and find that there exists a Borel set $E \subset \mathcal{X}^\circ$, with $P(E) > P(\mathcal{X}^\circ) - \epsilon = 1 - \epsilon$ and on which $\psi_{1/n}$ converges to ψ uniformly.

Choose $N \in \mathbb{N}^+$, such that $N\sqrt{p}\alpha \geq 1$ and for any $x \in E$, $|\psi_{1/N} - \psi|(x) \leq M$. Obviously, there exists an integer $l \geq 2$, such that $l - 1 \leq N\sqrt{p}\alpha < l$.

We then partition Q into l^p cubes, each of which has the form:

$$\prod_{i=1}^{p} [a_i + \frac{k_i}{l}\alpha, a_i + \frac{k_i + 1}{l}\alpha], k_i \in \mathbb{Z}, 0 \leq k_i < l.$$

Denote these cubes by $Q_1, ..., Q_{l^p}$.

Let $\mathcal{A} = \{1 \leq j \leq l^p : Q_j \subset \mathcal{X} - E\}$ and $\mathcal{B} = \{1 \leq j \leq l^p : j \notin \mathcal{A}\}$. One has

$$Q = \bigcup_{j\in\mathcal{A}} Q_j \cup \bigcup_{j\in\mathcal{B}} Q_j \subset (\mathcal{X} - E) \cup \bigcup_{j\in\mathcal{B}} Q_j.$$

Hence

$$P(Q) \leq P(\mathcal{X} - E) + \sum_{j\in\mathcal{B}} P(Q_j) < \epsilon + \sum_{j\in\mathcal{B}} P(Q_j). \tag{10}$$

Now we estimate $P(Q_j), j \in \mathcal{B}$. Pick $x_j \in Q_j \cap E$. Then

$$\phi_{1/N}(x_j) \leq \frac{1}{N^p}\psi_{1/N}(x_j) \leq \frac{1}{N^p}(\psi(x_j) + M) \leq \frac{2M}{N^p}. \tag{11}$$

It can be verified that the diameter of Q_j is $\sqrt{p}\alpha/l < 1/N$ by the choice of l. Hence

$$Q_j \subset \overline{B(x_j, \sqrt{p}\alpha/l)} \subset B(x_j, \frac{1}{N}),$$

which, together with (11), yields

$$P(Q_j) \le P(B(x_j, \frac{1}{N})) = \phi_{1/N}(x_j) \le \frac{2M}{N^p}. \tag{12}$$

Combining (12) with (10), we find

$$P(Q) \le \frac{2Ml^p}{N^p} + \epsilon \le 2Ml^p\Big(\frac{\sqrt{p}\alpha}{l-1}\Big)^p + \epsilon \le K\alpha^p + \epsilon = Km(Q) + \epsilon,$$

which is the desired result. □

In the light of Lemma 3 we get the following Corollary 1 immediately.

Corollary 1. $P(E) \le Km(E)$ holds for all Borel sets $E \subset \mathcal{X}$.

Proof. We have proved that the desired inequality holds for all left-open right-closed cubes

$$\prod_{i=1}^{p}(a_i, a_i + \alpha] \subset \mathcal{X}.$$

The remainder of this proof follows from a standard monotone class argument (Kallenberg, 2006) and the fact that $P(\partial\mathcal{X}) = 0$. □

Now we complete our technical preparations and are ready to prove the main theorems in this section.

Proof of Theorem 3. By Corollary 1, P is absolutely continuous with respect to the restriction of Lebesgue measure m on the Borel algebra $\mathscr{B}$ of $\mathcal{X}$. By Radon-Nikodym's Theorem (see Stein and Shakarchi, 2009, p. 290), the density function f exists, and is Borel measurable on $\mathcal{X}$. □

Lemma 4 below is a direct consequence of the translation and dilation invariance of Lebesgue measure. Define $B_s(x,r) := \{y \in \mathbb{R}^p : d_s(x,y) < r\}$ for $0 < s < +\infty$.

Lemma 4. Let m be the Lebesgue measure on $\mathbb{R}^p$. Set $v_c = m(Cu(0,1)), v = m(B(0,1))$ and $v_s = m(B_s(0,1))$. Then $m(Cu(x,r)) = v_c r^p, m(B(x,r)) = vr^p, m(B_s(x,r)) = v_s r^p$.

We now present the proof of Theorem 2.

Proof of Theorem 2. First notice, obviously we have

$$Cu(x,R) \subset B(x, \sqrt{p}R) \subset Cu(x, \sqrt{p}R).$$

Next, because $\mathcal{X}$ is compact, g attains its maximum and minimum on $\mathcal{X}$, denoted by C and c respectively.

Thus $Cu(x, \frac{r}{g(x)}) \subset B(x, \frac{r}{g(x)/\sqrt{p}}) \subset B(x, \frac{r}{c/\sqrt{p}})$. Similarly, $Cu(x, \frac{r}{g(x)}) \supset B(x, \frac{r}{C})$. Invoking (4), we have

$$1 = \lim_{r \downarrow 0} \frac{P(Cu(x_1, \frac{r}{g(x_1)}))}{P(Cu(x_2, \frac{r}{g(x_2)}))} \leq \liminf_{r \downarrow 0} \frac{P(B(x_1, \frac{r}{c/\sqrt{p}}))}{P(B(x_2, \frac{r}{C}))},$$

which is Condition 2 with $c_1 = (c/\sqrt{p})^{-1}, c_2 = C^{-1}, c_3 = 1$. Thus the existence of the density function f is ensured by Theorem 3. We extend f to $\mathbb{R}^p$ with $f = 0$ outside $\mathcal{X}$. Then $f \in L^1(\mathbb{R}^p)$.

Let

$$L_f = \left\{ x \in \mathcal{X} : \lim_{r \downarrow 0} \frac{1}{m(B(x, r))} \int_{B(x,r)} |f - f(x)| \mathrm{d}m = 0 \right\}$$

be the Lebesgue set of f. Then $m(\mathbb{R}^p - L_f) = 0$. See Stein and Shakarchi (2009), p. 106.

Hence for all $x \in L_f$, using the fact that $B(x, r) \supset Cu(x, r/\sqrt{p})$,

$$\begin{aligned} 0 &= \lim_{r \downarrow 0} \frac{1}{m(B(x, r))} \int_{B(x,r)} |f - f(x)| \mathrm{d}m \\ &\geq \limsup_{r \downarrow 0} \frac{v_c p^{-p/2}/v}{m(Cu(x, r/\sqrt{d}))} \int_{Cu(x, r/\sqrt{d}))} |f - f(x)| \mathrm{d}m \\ &= \limsup_{r \downarrow 0} \frac{v_c p^{-p/2}/v}{m(Cu(x, r))} \int_{Cu(x,r)} |f - f(x)| \mathrm{d}m \geq 0, \end{aligned}$$

which implies

$$\lim_{r \downarrow 0} \frac{1}{m(Cu(x, r))} \int_{Cu(x,r)} |f - f(x)| \mathrm{d}m = 0.$$

As a consequence, for all $x_1, x_2 \in L_f \cap \mathcal{X}^\circ$, we have

$$\lim_{r \downarrow 0} \frac{1}{v_c(\frac{r}{g(x_i)})^p} \int_{Cu(x_i, \frac{r}{g(x_i)})} f \mathrm{d}m = f(x_i), i = 1, 2.$$

and

$$1 = \lim_{r \downarrow 0} \frac{P(Cu(x_1, \frac{r}{g(x_1)}))}{P(Cu(x_2, \frac{r}{g(x_2)}))} = \lim_{r \downarrow 0} \frac{\int_{Cu(x_1, \frac{r}{g(x_1)})} f \mathrm{d}m}{\int_{Cu(x_2, \frac{r}{g(x_2)})} f \mathrm{d}m}.$$

An elementary calculation shows that $(fg^{-p})(x_1) = (fg^{-p})(x_2)$. Thus fg^{-p} keeps as a constant on $L_f \cap \mathcal{X}^\circ$. Hence $f \propto g^p$, m-almost everywhere on $\mathcal{X}$. $\square$

3 Characteristics of Minimum Energy Designs

To prove the asymptotic result given in Theorem 1, it is necessary to exploit some properties of the minimum energy designs with finite sample size. First we introduce some necessary notation. Fix $s \in (0, +\infty)$. Define the energy of two distinct points $x, x' \in \mathcal{X}$ by

$$E_{s0}(x, x') := q(x)q(x')/d_s(x, x').$$

It is easily seen that $E_s(D)$ in (1) is the maximum energy values among all pairs of points from the design D. For two non-empty sets of disjoint scattered points G, H, define the energy between G, H by

$$e_s^*(G, H) := \sup\{(q(x)q(x'))/d_s(x, x') : x \in G, x' \in H\}.$$

Proposition 1 shows how the energy function can be calculated using the subsets of the design. It can be proved in a straightforward manner.

Proposition 1. Let D_1, D_2 be two sets of disjoint design points over $\mathcal{X}$, each having at least two points. Then

$$E_s(D_1 \cup D_2) = \max\{E_s(D_1), E_s(D_2), e_s^*(D_1, D_2)\}.$$

Besides, if $E_s(D_2) < E_s(D_1)$ and $e_s^*(D_1, D_2) < E_s(D_1)$, then $IN(D_1 \cup D_2) = IN(D_1)$.

Given a design D, we call $x \in D$ a *critical point*, if there exists $x' \in D$ with $x' \neq x$ such that $E_{s0}(x, x') = E_s(D)$. Lemma 5 describes an important property of the minimum energy designs with the smallest index.

Lemma 5. Suppose D is an n-point minimum energy design with the smallest index and D' is an n'-point design with $n' > n$ and $E_s^*(D') = E_s^*(D)$. Then $IN(D') > IN(D)$ holds.

Proof. Suppose $E_s^*(D) = E_s^*(D')$ and $IN(D) \geq IN(D')$. By deleting a critical point from D', we obtain an $(n' - 1)$-point design D'_{-1} with $E_s^*(D'_{-1}) \leq E_s^*(D')$. Then we should have $E_s^*(D'_{-1}) = E_s^*(D')$. If this is not true, we can find an n-point design D'' with $E_s^*(D'') \leq E_s^*(D'_{-1}) < E_s^*(D') = E_s^*(D)$ by deleting any $n' - n - 1$ points from D'_{-1}, which contradicts the minimum energy property of D given by (2).

Now consider two cases of $IN(D')$. If $IN(D') = 1$, there is only one pair of points having the minimum energy. Thus $E_*(D'_{-1}) > E_*(D')$, which has been proved to be impossible. For D' with $IN(D') \geq 2$, by the definition of critical points, we have $IN(D'_{-1}) < IN(D')$. By repeating this scheme, we can obtain an n-point design $\tilde{D}'$ with $E_*(\tilde{D}') = E_*(D') = E_*(D)$ and $IN(\tilde{D}') < IN(D') \leq IN(D)$, which is a contradiction because D has the smallest index among all minimum energy designs of n points. $\square$

For two nonempty subset $A, B \subset \mathbb{R}^p$, define

$$d_s(A,B) = \inf_{x\in A, x'\in B} d_s(x,x').$$

Denote the closure of a set A as $\overline{A}$. Obviously $d_s(A,B) = d_s(\overline{A},\overline{B})$ for any nonempty sets $A, B \subset \mathbb{R}^p$. Lemma 6 shows a simple but useful result.

Lemma 6. Let $A = Cu(x_0, l_1), B = \mathbb{R}^p \setminus Cu(x_0, l_2)$ with $l_2 > l_1$ and $x_0 \in \mathbb{R}^p$. Then

$$d_s(A,B) = l_2 - l_1.$$

Proof. Note that $a = (x_{01} + l_1, x_{02}, \dots, x_{0p}) \in \overline{A}, b = (x_{01} + l_2, x_{02}, \dots, x_{0p}) \in B$, where x_{0i} denotes the ith entry of x_0. Thus

$$d_s(A,B) \le d_s(a,b) = l_2 - l_1.$$

On the other hand, for any $x = (x_1, \dots, x_p) \in A, x' = (x'_1, \dots, x'_p) \in B$, it is easily seen that there exist $i_0 \in \{1, \dots, d\}$ so that $|x_{i_0} - x'_{i_0}| \ge l_2 - l_1$. This implies

$$d_s(x,x') = \left(\sum_{i=1}^{p} |x_i - x'_i|^s \right)^{1/s} \ge |x_{i0} - x'_{i0}| \ge l_2 - l_1,$$

which yields $d_s(A,B) \ge l_2 - l_1$. In summary we obtain $d_s(A,B) = l_2 - l_1$. □

It can be seen that a minimum energy design becomes a maximin distance design if $q(x) \equiv 1$ and $s = 2$. This gives us an intuition that minimum energy designs are not too far from space-filling designs (Santner et al., 2003). We account for this space-filling property for two special regions in Theorem 4. The results are useful in a "peeling argument" in the proof of Theorem 1.

Theorem 4. Let $D = \{x_1, \dots, x_n\}$ be a minimum energy design over $\mathcal{X}$ under d_s with charge function $q(x)$. Suppose $0 < \underline{q} \le \overline{q} < +\infty$ for all $x \in \mathcal{X}$ and D has the smallest index among all such designs. Let $e_s = E_s(D)$. Then the following statements are true.

(i) There exists a constant C_0 depending only on $\underline{q}, \overline{q}$ and p, such that for all $n \ge C_0$,

$$\underline{q}^2 (n v_s)^{1/p}/4 < e_s < 2\overline{q}^2 n^{1/p}.$$

(ii) Let $Cu(a,l) \subset \mathcal{X}$. For $0 < \epsilon \le l/2$, let $N_1 = \mathrm{card}(D \cap \overline{Cu(a,l) \setminus Cu(a,l-\epsilon)})$. Then there exists a constant C_1 depending only on $\underline{q}, \overline{q}$ and p, such that for all $n \ge C_1$ we have

$$N_1 \le p 8^p v_s^{-1} \underline{q}^{-2p} (e_s l)^{p-1} \max(\epsilon e_s, \overline{q}^2),$$

where v_s is defined in Lemma 4.

(iii) For any $Cu(a,l) \subset \mathcal{X}$, let $N_2 = \text{card}(D \cap \overline{Cu(a,l)})$. Then there exists a constant C_2 depending only on $\underline{q}, \overline{q}$ and p, such that for all $n > C_2 l^{-p}$ we have

$$N_2 \geq (e_s l/\overline{q}^2)^p.$$

Proof. For any design $D_0 \subset \mathcal{X}$, define the minimum distance

$$d_{s0}(D_0) = \min\{d_s(x_i, x_j) : x_i, x_j \in D_0, x_i \neq x_j\}.$$

By the definition of $\underline{q}$ and $\overline{q}$, for any $x_i \neq x_j$ we have

$$\underline{q}^2/d_s(x_i,x_j) \leq q(x_i)q(x_j)/d_s(x_i,x_j) \leq \overline{q}^2/d_s(x_i,x_j).$$

This implies

$$\max_{i \neq j} \underline{q}^2/d_s(x_i,x_j) \leq \max_{i \neq j} q(x_i)q(x_j)/d_s(x_i,x_j) \leq \max_{i \neq j} \overline{q}^2/d_s(x_i,x_j),$$

which is

$$\underline{q}^2/d_{s0}(D_0) \leq E_s(D_0) \leq \overline{q}^2/d_{s0}(D_0). \tag{13}$$

We will repeatedly use (13) in the derivations below. We will use the notation $d_1 = d_{s0}(D)$ throughout the proof of this theorem.

First we find a lower bound for e_s. As in Lemma 4 we denote $B_s(x,r) = \{x' \in \mathbb{R}^p : d_s(x,x') < r\}$ for $x \in \mathbb{R}^p, r > 0$. Let $B_i = B_s(x_i, d_0/2)$. The definition of d_1 implies that B_i's are disjoint. Besides, the union of B_i's is covered by the cube $[-d_0/2, 1+d_0/2]^d$. Therefore, the volume (i.e., Lebesgue measure) of $\cup B_i$ is less than the volume of the cube. Since the B_i's have the same volume, the volume of $\cup B_i$ equals the volume of each B_i times n. We Invoke Lemma 4 to obtain

$$n v_s (d_1/2)^p \leq (1+d_1)^p. \tag{14}$$

Thus

$$1/d_1 \geq (n v_s)^{1/p}/2 - 1,$$

which, together with (13), implies

$$e_s \geq \underline{q}^2 \left((n v_s)^{1/p}/2 - 1 \right).$$

Choose C_0 to be large enough so that $n v_s > 4^p$. Then we have

$$e_s \geq \underline{q}^2 (n v_s)^{1/p}/4. \tag{15}$$

Next we prove the upper bound. Since D is a minimum energy design under d_s, $e_s = E_s(D)$ is no greater than $E_s(D')$ for any n-point design D'. Note that a lattice design over $[0,1]^p$ with m levels along each axis has m^p design points and minimum distance $1/m$. Thus n points can form a (fractional) lattice design with minimum distance $1/\lceil n^{1/p}\rceil$, where $\lceil a\rceil$ denotes the minimum integer which is no less than a. Let D' be such a design. Thus $d_{s0}(D') \geq 1/\lceil n^{1/p}\rceil > (n^{1/p}+1)^{-1}$. By (13),

$$\overline{q}^2(n^{1/p}+1) > \overline{q}^s/d_{s0}(D') \geq E_s(D') \geq E_s(D) = e_s,$$

which, together with (15) implies the desired result in (i).

The proof of (ii) follows from a similar argument as that for (15). From (13) and (15), we have

$$d_1/l \leq \overline{q}^2/(e_s l) < 4\overline{q}^2\underline{q}^2(v_s C_1)^{-1/p}.$$

Thus we can choose a sufficiently large C_1 so that $d_1/2 \leq \overline{q}^2/(2e_s) < l/2 < l-\epsilon$. As before, we let $B_i = B_s(x_i, d_1/2)$. By the definition of d_1, B_i's are disjoint. Let $\Omega = \overline{Cu(a,l) \setminus Cu(a,l-\epsilon)}$. Lemma 6 implies that $\cup_{x_i\in\Omega} B_i$ is covered by

$$\overline{Cu(x, l+d_1/2) \setminus Cu(x,(l-\epsilon)-d_1/2)}. \tag{16}$$

Thus we apply an argument similar to (14) to obtain that

$$N_1 v_s (d_1/2)^p \leq (2l+d_1)^p - (2(l-\epsilon)-d_1)^p,$$

which, together with (13), yields

$$N_1 v_s \underline{q}^{2p} \leq N_1 v_s (d_1 e_s)^p \leq (4le_s + 2\overline{q}^2)^p - (4(l-\epsilon)e_s - 2\underline{q}^2)^p. \tag{17}$$

Note the we can bound the right hand side of (17) with

$$\begin{aligned}(4le_s + 2\overline{q}^2)^p - (4(l-\epsilon)e_s - 2\underline{q}^2)^p &\leq d(4\epsilon e_s + 2\overline{q}^2 + 2\underline{q}^2)(4le_s + 2\overline{q}^2)^{p-1} \\ &\leq 4d(\epsilon e_s + \overline{q}^2)(4le_s + 2\overline{q}^2)^{p-1}.\end{aligned} \tag{18}$$

where the first inequality follows by applying the mean value theorem to $f(x) = x^p$. From (15), for sufficiently large n,

$$e_s l \geq (\underline{q}^2 v_s^{1/p}/4)(n^{1/p} l) > \underline{q}^2 (C_1 v_s)^{1/p}/4. \tag{19}$$

This implies that we can choose C_1 large enough so that $2le_s > \overline{q}^2$. Then by applying the inequality $|A+B| \leq 2\max(|A|,|B|)$ to (18), we prove (ii).

We now prove (iii). Let $d_0 := d_1\overline{q}^2/\underline{q}^2$. First we will prove that under the conditions of (iii),

$$N_2 \geq \left(\frac{2e_s(l-d_0-\delta)}{\overline{q}^s} - 1\right)^p, \tag{20}$$

for all sufficiently small δ.

Because $n^{1/p}l > C_2^{1/p}$, we first force C_2 to be large enough so that we can apply (15) to obtain $e_s l > \underline{q}^2(v_s C_2)^{1/p}/4$. Thus we can choose a sufficiently large C_2 so that the following inequality holds

$$2e_s l\underline{q}^2 - 2\overline{q}^4 > 3\overline{q}^6/\underline{q}^2. \tag{21}$$

Combining (21) and (13) we obtain

$$2e_s(l-d_0) \geq 2e_s l - 2\overline{q}^4/\underline{q}^2 > 3\overline{q}^6/\underline{q}^4 \geq 3\overline{q}^2. \tag{22}$$

From (22) it is easily seen that for sufficiently small δ,

$$\left\lceil \frac{2e_s(l-d_0-\delta)}{\overline{q}^2} \right\rceil - 1 \geq 2. \tag{23}$$

Denote the lattice design over $\overline{Cu(x, l-d_0-\delta)}$ with

$$\left\lceil \frac{2e_s(l-d_0-\delta)}{\overline{q}^2} \right\rceil - 1$$

levels by D'. According to (23), D' contains at least 2^p points. Then

$$E_s(D') < \frac{\overline{q}^2}{2(l-d_0-\delta)/\left\{\left\lceil \frac{2e_s(l-d_0-\delta)}{\overline{q}^2} \right\rceil - 1\right\}} < e_s. \tag{24}$$

Define a new design D^* by replacing the points of D in $\overline{Cu(x,l)}$ with D', i.e.,

$$D^* = (D \setminus \overline{Cu(x,l)}) \cup D'.$$

Let $D_1 = D \setminus \overline{Cu(x,l)}$. Note that

$$\begin{aligned} E_s(D_1) &\leq E_s(D) = e_s, \qquad (25)\\ e_s^*(D_1, D') &\leq \frac{\overline{q}^2}{d_s\left(Cu(x, l-d_0-\delta), \mathbb{R}^p \setminus Cu(x,l)\right)} \\ &= \frac{\overline{q}^2}{d_0+\delta} < \frac{\overline{q}^2}{d_0} = \frac{\underline{q}^2}{d_1} \leq e_s, \qquad (26) \end{aligned}$$

where in formula (26), the first inequality follows from the definition of e_s^*; the first equality follows from Lemma 6; the last inequality follows from (13). Combining (25), (24), (26) and apply Proposition 1, we obtain $E_s(D^*) \leq e_s$ and $IN(D^*) = IN(D_1) \leq IN(D)$.

Besides, by the definition of D' we have

$$\operatorname{card}(D') = \left(\left\lceil \frac{2e_s(l-d_0-\delta)}{\overline{q}^2} \right\rceil - 1\right)^p \geq \left(\frac{2e_s(l-d_0-\delta)}{\overline{q}^2} - 1\right)^p. \tag{27}$$

Now suppose (20) is false. By (27) we have $\mathrm{card}(D') > \mathrm{card}(D)$. But this contradicts Lemma 5, because D is a minimum energy design with the smallest index. Hence we have proved (20), which, together with (13), yields

$$N_2 \geq \left(\frac{2e_s(l-d_0-\delta)}{\overline{q}^2} - 1\right)^p \geq \left(\frac{2e_s(l-\delta)}{\overline{q}^2} - \frac{2\overline{q}^2}{\underline{q}^2} - 1\right)^p.$$

By letting $\delta \downarrow 0$, we obtain

$$N_2 \geq \left(\frac{2e_s l}{\overline{q}^2} - \frac{2\overline{q}^4}{\underline{q}^2} - 1\right)^p.$$

Applying an argument similar to (19), we can choose a large enough C_2 so that $e_s l/\overline{q}^2 > 2\overline{q}^4/\underline{q}^2 + 1$. The desired result of (iii) then follows.

□

4 Proof of Theorem 1

We are now ready to prove our main theorem.

Proof of Theorem 1. Since we have fixed s, we abbreviate $\mathcal{P}_n^s$ to $\mathcal{P}_n$ with no ambiguity.

By Prohorov's Theorem (Dudley, 2002), there exists a subsequence of $\{\mathcal{P}_n\}$, denoted as $\{\mathcal{P}_{k_n}\}$, which converges weakly to a probability measure $\mathcal{P}$. Thus it suffices to prove that every convergent subsequence of $\{\mathcal{P}_{k_n}\}$ tends to the uniform distribution. For notational simplicity, we assume $\mathcal{P}_n \xrightarrow{w} \mathcal{P}$, so that the goal is to prove that $\mathcal{P}$ has density f with $f \propto q^{-2p}$. Let D_n be the design corresponding to $\mathcal{P}_n$.

First we prove $\mathcal{P}(\partial\mathcal{X}) = 0$. Let $a_0 = (1/2, \ldots, 1/2) \in \mathbb{R}^p$ and $\mathcal{X}_\epsilon = \mathcal{X} \setminus \overline{Cu(a_0, 1/2-\epsilon)}$ for $0 < \epsilon < 1/4$. Now apply Theorem 4 for a fixed ϵ to find constants c_1, c_2 and N depending only on $\epsilon, \underline{q}, \overline{q}$ and p, so that for all $n > N$,

$$\mathcal{P}_n(\mathcal{X}_\epsilon) = \frac{\mathrm{card}(D_n \cap \mathcal{X}_\epsilon)}{n} \leq c_1(n^{-1}E_s^p(D_n)\epsilon) \leq c_2\epsilon,$$

where the first inequality follows from (ii) of Theorem 4 and the second inequality follows from (i) of Theorem 4. Noting that $\mathcal{X}_\epsilon$ is an open set in $\mathcal{X}$, we have

$$\mathcal{P}(\mathcal{X}_\epsilon) \leq \liminf_{n\to\infty} \mathcal{P}_n(\mathcal{X}_\epsilon) \leq c_2\epsilon,$$

where the first inequality is a consequence of the weak convergence. See Klenke (2013) for details. Hence we conclude that $\mathcal{P}(\partial\mathcal{X}) = \lim_{\epsilon\downarrow 0}\mathcal{P}(\mathcal{X}_\epsilon) \leq 0$.

Next we prove that

$$\liminf_{l\downarrow 0} \frac{\mathcal{P}\left(\overline{Cu(y, lq^2(y))}\right)}{\mathcal{P}(Cu(x, lq^2(x)))} \geq 1, \tag{28}$$

for any $x, y \in \mathcal{X}^\circ$. Otherwise, suppose

$$\limsup_{l \downarrow 0} \frac{\mathcal{P}\left(\overline{Cu(a_2, lq^2(a_2))}\right)}{\mathcal{P}(Cu(a_1, lq^2(a_1)))} \leq 1 - \delta, \tag{29}$$

for some $a_1, a_2 \in \mathcal{X}^\circ$ and $\delta > 0$. It follows from $\mathcal{P}_n \xrightarrow{w} \mathcal{P}$ that

$$\begin{aligned} \liminf_{n\to\infty} \mathcal{P}_n(Cu(a_1, lq^2(a_1))) &\geq \mathcal{P}(Cu(a_1, lq^2(a_1))), &(30)\\ \limsup_{n\to\infty} \mathcal{P}_n\left(\overline{Cu(a_1, lq^2(a_1))}\right) &\leq \mathcal{P}\left(\overline{Cu(a_1, lq^2(a_1))}\right), &(31) \end{aligned}$$

for sufficiently small l. See Klenke (2013). Combining (29), (30) and (31), for any small l and $N > 0$ we can find $n > N$ such that

$$1 - \delta/2 \geq \frac{\mathcal{P}_n\left(\overline{Cu(a_2, lq^2(a_2))}\right)}{\mathcal{P}_n(Cu(a_1, lq^2(a_1)))} \geq \frac{\mathcal{P}_n\left(\overline{Cu(a_2, lq^2(a_2))}\right)}{\mathcal{P}_n\left(\overline{Cu(a_1, lq^2(a_1))}\right)} = \frac{\text{card}\,(D_{n2}(l))}{\text{card}\,(D_{n1}(l))}, \tag{32}$$

where $D_{ni}(l) = D_n \cap \overline{Cu(a_i, lq^2(a_i))}$ for $i = 1, 2$. This suggests that $D_{n1}(l)$ contains much more points than $D_{n2}(l)$. Similar with (iii) of Theorem 4, we will construct a "better" design under the minimum energy criterion, which leads to a contradiction. The idea is to replace the design points in $\overline{Cu(a_2, lq^2(a_2))}$ with these in $\overline{Cu(a_1, lq^2(a_1))}$ after a "peeling" operation.

Set $\underline{q} = \inf_{x\in\mathcal{X}} q(x)$ and $\overline{q} = \sup_{x\in\mathcal{X}} q(x)$. Let l and m be constants to be determined later. We first assume that l is sufficiently small so that

$$\overline{Cu(a_i, lq^2(a_i))} \subset \mathcal{X}^\circ, \tag{33}$$

holds for $i = 1, 2$. Define

$$\epsilon = 3\sqrt{p}L\overline{q}l^2. \tag{34}$$

We remind that L is the Lipschitz constant of q. Obviously $\epsilon < l$ for sufficiently small l. Define $D^*_{m1}(l) = D_{m1}(l) \cap Cu(a_1, lq^2(a_1) - 2\epsilon)$. By Theorem 4

$$\begin{aligned} &\frac{\text{card}(D_{m1}(l) \setminus D^*_{m1}(l))}{\text{card}(D_{m1}(l))} \leq O\left((E_s(D_m)l)^{-1} \max(2\epsilon E_s(D_m), \overline{q}^2)\right) \\ = \;& \max(O(\epsilon/l), O((E_s(D_m)l)^{-1})) = \max(O(l), O((E_s(D_m)l)^{-1})), \end{aligned} \tag{35}$$

provided that (33) holds, $\epsilon < l$ and

$$m > Cl^p, \tag{36}$$

for some constant C depending on $q, \overline{q}$ and p only. According to (35) and (36), we can find constants $l_0 > 0$, and for each $0 < l < l_0$ there exists $M_l > 0$, such that

$$\frac{\mathrm{card}(D_{m1}(l) \setminus D^*_{m1}(l))}{\mathrm{card}(D_{m1}(l))} \leq \frac{\delta/2}{1-\delta/2}.$$

holds for all $m > M_l$. Now we choose $l < l_0$ so that

$$\sqrt{p}L\overline{q}l < 2, \tag{37}$$

and (32) is possible to hold for a large n. Next we choose $m > M_l$ such that (32) and

$$\overline{q}^2/(3\sqrt{p}q^2(a_2)Ll^2) < E_s(D_n) \tag{38}$$

hold. Note that (32) and (35) imply that $D^*_{1m}(l)$ contains more points than $D_{2m}(l)$.

For simplicity, in the derivations below we will use the sloppy notation D^*_1, D_2, D instead of the precise notation $D^*_{1m}(l), D_{2m}(l), D_m$ respectively. Now define the affine transformation

$$\begin{aligned} T: \quad \overline{Cu(a_1, q^2(a_1)(l-2\epsilon))} &\rightarrow \overline{Cu(a_2, q^2(a_2)(l-\epsilon))}, \\ a_1 + u &\mapsto a_2 + \frac{q^2(a_2)(l-\epsilon)}{q^2(a_1)(l-2\epsilon)}u, \end{aligned}$$

for $u \in \overline{Cu(0, q^2(a_1)(l-2\epsilon))}$. Define a new set of design points by

$$D^* = (D \setminus D_2) \cup T(D^*_1).$$

Because $T(D^*_1)$ maintains the same number of points as D^*_1, D^* contains fewer points than D. Now we bound $E_s(D^*)$ using Proposition 1. Obviously,

$$\begin{aligned} E_s(D \setminus D_2) &\leq E_s(D), && (39) \\ e^*_s(D \setminus D_2, T(D^*_1)) &\leq \frac{\overline{q}^2}{d_s(Cu(a_2, q^2(a_2)(l-\epsilon), \mathcal{X} \setminus Cu(a_2, lq^2(a_2)))} \\ &\leq \overline{q}^2/(q^2(a_2)\epsilon) = \overline{q}^2/(3\sqrt{p}Lq^2(a_2)l^2) < E_s(D), && (40) \end{aligned}$$

where in (40), the second inequality follows from Lemma 6 and the last inequality follows from (38). Besides, we bound $E_s(T(D_1^*))$ by

$$
\begin{aligned}
E_s(T(D_1^*)) &= \max_{x_i,x_j\in D_1^*, x_i\neq x_j} \frac{q(T(x_i))q(T(x_j))}{d_s(T(x_i),T(x_j))} \\
&= \max_{x_i,x_j\in D_1^*, x_i\neq x_j} \left\{ \frac{q(x_i)q(x_j)}{d_s(x_i,x_j)} \cdot \frac{q^2(a_1)(l-2\epsilon)q(T(x_i))q(T(x_j))}{q^2(a_2)(l-\epsilon)q(x_i)q(x_j)} \right\} \\
&\leq E_s(D_1^*) \max_{x,y\in D_1^*} \frac{q^2(a_1)(l-2\epsilon)q(T(x_i))q(T(x_j))}{q^2(a_2)(l-\epsilon)q(x_i)q(x_j)} \\
&\leq E_s(D_1^*) \frac{q^2(a_1)(l-2\epsilon)(q(a_2)+\sqrt{p}Lq^2(a_2)(l-\epsilon)/2)^2}{q^2(a_2)(l-\epsilon)(q(a_1)-\sqrt{p}Lq^2(a_1)(l-2\epsilon)/2)^2} \\
&< E_s(D_1^*) \frac{q^2(a_1)(l-2\epsilon)(q(a_2)+\sqrt{p}Lq^2(a_2)l/2)^2}{q^2(a_2)(l-\epsilon)(q(a_1)-\sqrt{p}Lq^2(a_1)l/2)^2} \qquad (41) \\
&< E_s(D_1^*) \leq E_s(D_1), \qquad (42)
\end{aligned}
$$

where the second equality follows from the dilation invariance of d_s; the second inequality follows from the Lipschitz property of q and the fact that $q(a_1)-\sqrt{p}Lq^2(a_1)l > 0$ according to (37); the fourth inequality follows from some elementary calculations together with (34) and (37). We verify the calculations in the fourth inequality of (42) in Appedix A. Combining (39), (40) and (42) and invoking Proposition 1 we conclude that $E_s(D^*) \leq E_s(D)$ and $IN(D^*) \leq IN(D)$. Besides, we have shown that $\mathrm{card}(D^*) < \mathrm{card}(D)$. But this contradicts Lemma 5, because D is a minimum energy design with the smallest index. Therefore, (28) is true.

The remainder of this proof follows from the comparison-of-measure argument introduced in Section 2. By (28), it is easy to verify that Condition 2 in Section 2 holds for $c_1 = 2\overline{q}\sqrt{p}, c_2 = \underline{q}, c_3 = 1$. Besides, we have proved that $\mathcal{P}(\partial\mathcal{X}) = 0$. Therefore, we can apply Theorem 3 to conclude that $\mathcal{P}$ is absolutely continuous with respect to the Lebesgue measure on $\mathcal{X}$. As a consequence, we can remove the closure operation in (28) and obtain

$$\liminf_{l\downarrow 0} \frac{\mathcal{P}\left(Cu(y,lq^2(y))\right)}{\mathcal{P}(Cu(x,lq^2(x)))} \geq 1, \qquad (43)$$

for any $x,y\in\mathcal{X}^\circ$. By the symmetricity between x and y, (43) implies

$$\lim_{l\downarrow 0} \frac{\mathcal{P}\left(Cu(y,lq^2(y))\right)}{\mathcal{P}(Cu(x,lq^2(x)))} = 1, \qquad (44)$$

for any $x,y\in\mathcal{X}^\circ$, which is Condition 1. Then the desired result is a consequence of Theorem 2. □

Appendix

A Verification of Energy Bound

We check here that the factor in (41) satisfies

$$\frac{q^2(a_1)(l-2\epsilon)(q(a_2)+\sqrt{p}Lq^2(a_2)l/2)^2}{q^2(a_2)(l-\epsilon)(q(a_1)-\sqrt{p}Lq^2(a_1)l/2)^2} < 1. \tag{45}$$

To this end, we consider the following inequality with unknown x

$$\frac{q^2(a_1)(l-2x)(q(a_2)+\sqrt{p}Lq^2(a_2)l/2)^2}{q^2(a_2)(l-x)(q(a_1)-\sqrt{p}Lq^2(a_1)l/2)^2} < 1. \tag{46}$$

The solution to (46) is

$$\begin{aligned} x &> \frac{l\left\{(2+\sqrt{p}Lq(a_2)l)^2-(2-\sqrt{p}Lq(a_1)l)^2\right\}}{2(2+\sqrt{p}Lq(a_2)l)^2-(2-\sqrt{p}Lq(a_1)l)^2} \\ &= \frac{l^2\left\{4\sqrt{p}L(q(a_1)+q(a_2))+pL^2l(q^2(a_1)+q^2(a_2)))\right\}}{2(2+\sqrt{p}Lq(a_2)l)^2-(2-\sqrt{p}Lq(a_1)l)^2} \end{aligned} \tag{47}$$

Using (38), we can bound the right hand side of (47) by

$$\frac{l^2(12\sqrt{p}L\overline{q})}{2\cdot 2^2-2^2} = 3\sqrt{p}L\overline{q}l^2 = \epsilon.$$

This implies that $x=\epsilon$ can ensure (46), which proves (45).

References

Dudley, R. M. (2002). *Real analysis and probability*, Volume 74. Cambridge University Press.

Kallenberg, O. (2006). *Foundations of modern probability.* Springer Science & Business Media.

Klenke, A. (2013). *Probability theory: a comprehensive course.* Springer Science & Business Media.

Santner, T., B. Williams, and W. Notz (2003). *The Design and Analysis of Computer Experiments.* Springer Verlag.

Stein, E. M. and R. Shakarchi (2009). *Real analysis: measure theory, integration, and Hilbert spaces.* Princeton University Press.