Convergence of Laplacian spectra from random samples

Zuoqiang Shi*†

May 12, 2017

Abstract

Eigenvectors and eigenvalues of discrete Laplacians are often used for manifold learning and nonlinear dimensionality reduction. Graph Laplacian is one widely used discrete laplacian on point cloud. It was previously proved by Belkin and Niyogi [4] that the eigenvectors and eigenvalues of the graph Laplacian converge to the eigenfunctions and eigenvalues of the Laplace-Beltrami operator of the manifold in the limit of infinitely many data points sampled independently from the uniform distribution over the manifold. Recently, we introduced Point Integral method (PIM) [15] to solve elliptic equations and corresponding eigenvalue problem on point clouds. In this paper, we prove that the eigenvectors and eigenvalues obtained by PIM converge in the limit of infinitely many random samples. Moreover, one estimate of the rate of the convergence is given.

keywords: Graph Laplacian; Laplacian spectra; random samples; convergence rate.

1 Introduction

In the past decade, data science plays more and more important role in sciences, engineering and our daily lives. Among varieties of data analysis methods and models, manifold model attracts more and more attentions. In the manifold model, data is represented as a point cloud, which is defined as a collection of points that are embedded in a high dimensional Euclidean space. It is assumed that the point cloud samples a smooth manifold. Thus, the structure of the manifold are very useful to understand the data. On the other hand, research in mathematics shows that the Laplace-Beltrami operator is one of the most important object associated to Riemannian manifolds. Its eigenvalue and eigenfunctions encode all intrinsic geometry of the manifolds. To reveal the structure of the underlying manifold sampled by the data, many discrete counterparts of LBO are developed. The eigenvalues and eigenvectors of the discrete Laplace-Beltrami operators are widely used in many fields, including machine learning, data analysis, computer graphics and computer vision, and geometric modeling and processing [2, 6, 18, 16]. Then, one question is that if the eigenvalues and eigenvectors of these discrete operators converge to the eigenvalues and eigenfunctions of their continuous counterpart, Laplace-Beltrami operator as the point cloud converges to the manifold. This is essential to understand these discrete operators and algorithms associated to them.

The convergence between the graph Laplacian and the Laplace-Beltrami operator has been studied extensively in the literature [13, 3, 4, 9, 8, 7, 21, 23]. In the presence of no boundary and the sample points are uniformly distributed, Belkin and Niyogi [4] showed that the spectra of the normalized graph Laplacian converges to the spectra of Laplace-Beltrami operator. When there is a boundary, it was observed in [13, 5] that the integral Laplace operator L_t is dominated by the first order derivative and thus fails to be true

*Yau Mathematical Sciences Center, Tsinghua University, Beijing, China, 100084. *Email: zqshi@math.tsinghua.edu.cn.*

†This research was supported by NSFC Grant 11371220 and 11671005.

Laplacian near the boundary. Recently, Singer and Wu [22] showed the spectral convergence in the presence of the Neumann boundary. In this paper, we study this problem from another point of view. We study the solution operators of graph Laplacian and Laplace-Beltrami operator. Based on the convergence between the solutions operators, we get more delicate estimate of the convergence, include the convergence rate.

In this paper, we assume that the data points, $X_n = \{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$, are sampled independently over the manifold $\mathcal{M}$ from a probability distribution $p(\mathbf{x})$. On the sample points, we consider following discrete eigenvalue problem.

$$\frac{1}{t}\sum_{j=1}^{n} R\left(\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{4t}\right)(u_i - u_j) = \lambda \sum_{j=1}^{n} \bar{R}\left(\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{4t}\right) u_j, \tag{1.1}$$

where $R : \mathbb{R}^+ \to \mathbb{R}^+$ is a kernel function satisfies some conditions (see Assumption 1). , $\bar{R}(r) = \int_r^{+\infty} R(s)\mathrm{d}s$. The purpose of this paper is to study the behavior of discrete eigenvalue problem (1.1) as $n \to \infty$ and $t \to 0$. We show that when $n \to \infty$ and $t \to 0$, the spectral of (1.1) converge to the spectra of Laplace-Beltrami operator,

$$\begin{cases} -\frac{1}{p^2(\mathbf{x})}\,\mathrm{div}\left(p^2(\mathbf{x})\nabla u(\mathbf{x})\right) &= \lambda u(\mathbf{x}), \ \mathbf{x} \in \mathcal{M}, \\ \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) &= 0, \qquad \mathbf{x} \in \partial\mathcal{M}. \end{cases} \tag{1.2}$$

where $\mathbf{n}$ is the out normal vector of $\mathcal{M}$.

The analysis in this paper consists of two parts by introducing an intermediate integral operator.

$$\frac{1}{t}\int_{\mathcal{M}} R\left(\frac{\|\mathbf{x} - \mathbf{y}\|^2}{4t}\right)(u(\mathbf{x}) - u(\mathbf{y}))p(\mathbf{y})\mathrm{d}\mathbf{y} = \lambda \int_{\mathcal{M}} \bar{R}\left(\frac{\|\mathbf{x} - \mathbf{y}\|^2}{4t}\right) u(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}, \quad \mathbf{x} \in \mathcal{M}. \tag{1.3}$$

The convergence between the integral operator, (1.3), and graph Laplacian (1.1) has been well studied in different settings [11, 12, 19, 10, 24, 25]. Under the assumption of smoothness of the kernel function R, the integral operator becomes a compact operator in some suitable space. Then, with the help of the perturbation theory of compact operator, the spectral convergence can be proved. In this paper, we also use this approach. First, the convergence between the integral operator and graph Laplacian is obtained by using standard estimate in the empirical process theory. Then, perturbation theory of compact operator is invoked to prove the convergence of spectra. The second convergence, from integral operator to Laplace-Beltrami operator, is more involved. Laplace-Beltrami operator is not compact, so we consider the solution operator (inverse operator of Laplace-Beltrami in some sense) which is known to be compact. The convergence between the solutions operators is more difficult. Fortunately, the convergence between the solution operators in H^1 has been proved in our previous paper. With the help of this strong convergence, we can even get the rate of the spectral convergence.

The main contribution of this paper is of two fold. First, we prove the convergence for general compact manifolds, with or without boundary. Previous studies mainly focus on the manifolds without boundary. Our analysis works for general compact manifolds. Secondly, we get the rate of spectral convergence, although may not be optimal. The rate comes from the convergence of the solution operators in H^1, which allows us to do more delicate estimate.

The rest of the paper is organized as follows. We define some necessary notations and operators in Section 2. In Section 3 and 4, the main theorem is stated and proved respectively. Some technical results are proved in Section 5. Finally, conclusions and remarks are given in Section 6.

2 Notations and Preliminaries

We start with defining three solution operators, $T, T_t, T_{t,n}$ under the assumption that $\mathcal{M} \in C^\infty$, $\partial\mathcal{M} \in C^\infty$.

- $T : L^2(\mathcal{M}) \to H^2(\mathcal{M})$ is the solution operator of the problem (2.1), i.e., $u = T(f)$ with $\int_{\mathcal{M}} u(\mathbf{x})p(\mathbf{x})\mathrm{d}\mathbf{x} = 0$ is the solution of the following problem:

$$\begin{cases} -\frac{1}{p^2(\mathbf{x})}\ \mathrm{div}\left(p^2(\mathbf{x})\nabla u(\mathbf{x})\right) &= f(\mathbf{x}) - \bar{f},\ \mathbf{x} \in \mathcal{M}, \\ \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) &= 0, \qquad\quad \mathbf{x} \in \partial\mathcal{M}. \end{cases} \tag{2.1}$$

  where $\mathbf{n}$ is the out normal vector of $\mathcal{M}$, $\bar{f}$ is a constant such that $\int_{\mathcal{M}}(f(\mathbf{x}) - \bar{f})p^2(\mathbf{x})\mathrm{d}\mathbf{x} = 0$. Here, we further assume that $p(\mathbf{x}) \in C^1(\mathcal{M})$ and $\min_{\mathbf{x}\in\mathcal{M}} p(\mathbf{x}) > 0$, $\max_{\mathbf{x}\in\mathcal{M}} p(\mathbf{x}) < \infty$.

- $T_t : L^2(\mathcal{M}) \to L^2(\mathcal{M})$ is the solution operator of following integral equation (2.2), i.e. $u = T_t(f)$ with $\int_{\mathcal{M}} u(\mathbf{x})p(\mathbf{x})\mathrm{d}\mathbf{x} = 0$ solves the following integral equation

$$\frac{1}{t}\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x}) - u(\mathbf{y}))p(\mathbf{y})\mathrm{d}\mathbf{y} = \int_{\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})(f(\mathbf{y}) - \bar{f}_t)p(\mathbf{y})\mathrm{d}\mathbf{y}. \tag{2.2}$$

  where

$$R_t(\mathbf{x},\mathbf{y}) = \frac{1}{(4\pi t)^{k/2}} R\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right), \quad \bar{R}_t(\mathbf{x},\mathbf{y}) = \frac{1}{(4\pi t)^{k/2}} \bar{R}\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right).$$

  $R \in C^2$ and $\bar{f}_t$ is a constant such that $\int_{\mathcal{M}}\int_{\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})(f(\mathbf{y}) - \bar{f}_t)p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{y}\mathrm{d}\mathbf{x} = 0$

- $T_{t,n} : C(\mathcal{M}) \to C(\mathcal{M})$ is defined as follows.

$$T_{t,n}(f)(\mathbf{x}) = \frac{1}{n\, w_{t,n}(\mathbf{x})}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j)u_j + \frac{t}{n\, w_{t,n}(\mathbf{x})}\sum_{j=1}^{n} \bar{R}_t(\mathbf{x},\mathbf{x}_j)(f(\mathbf{x}_j) - \bar{f}_{t,n}) \tag{2.3}$$

  where $w_{t,n}(\mathbf{x}) = \frac{1}{n}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j)$ and $\mathbf{u} = (u_1, \cdots, u_n)^t$ with $\sum_{i=1}^{n} u_i = 0$ solves following linear system,

$$\frac{1}{nt}\sum_{j=1}^{n} R_t(\mathbf{x}_i,\mathbf{x}_j)(u_i - u_j) = \frac{1}{n}\sum_{j=1}^{n} \bar{R}_t(\mathbf{x}_i,\mathbf{x}_j)(f(\mathbf{x}_j) - \bar{f}_{t,n}), \tag{2.4}$$

  $\bar{f}_{t,n}$ is the constant that satisfies $\sum_{i,j=1}^{n} \bar{R}_t(\mathbf{x}_i,\mathbf{x}_j)(f(\mathbf{x}_j) - \bar{f}_{t,n}) = 0$.

Using the definition of T, T_t and $T_{t,n}$, it is easy to show that the eigen problems $Tu = \lambda u$, $T_{t,n}(u) = \lambda u$ is equivalent to the eigen problems (1.2) and (1.1) respectively. Namely their eigenvalues are reciprocal to each other and they share the same eigenspaces.

Proposition 2.1. *Let $\theta(u)$ denote the restriction of u to the sample points P, i.e., $\theta(u) = (u(\mathbf{x}_1), \cdots, u(\mathbf{x}_n))^t$.*

1. *If a function u is an eigenfunction of $T_{t,n}$ with the eigenvalue λ, then the vector $\theta(u)$ is an eigenvector of the eigenproblem (1.1) with eigenvalue $1/\lambda$.*

2. *If a vector $\mathbf{u}$ is an eigenvector of the eigenproblem (1.1) with the eigenvalue $\lambda \neq 0$, then $I_\lambda(\mathbf{u})$ is an eigenfunction of $T_{t,n}$ with eigenvalue $1/\lambda$, where*

$$I_\lambda(\mathbf{u})(\mathbf{x}) = \frac{\sum_{p_j\in P} R_t(\mathbf{x},\mathbf{x}_j)u_j V_j - \lambda t\sum_{p_j\in P} \bar{R}_t(\mathbf{x},\mathbf{x}_j)u_j V_j}{\sum_{p_j\in P} R_t(\mathbf{x},\mathbf{x}_j)V_j}.$$

3. *A function u is the eigenfunction of the eigenproblem (1.2) with the eigenvalue $\lambda \neq 0$ if and only if the function u is an eigenfunction of T with the eigenvalue $1/\lambda$.*

Using the above proposition, we only need to prove the eigenvalues and the eigenfunctions of $T_{t,n}$ converge to the eigenvalues and the eigenfunctions of T.

The advantage of using the solution operators is that they are all compact operators.

Proposition 2.2.

1. *For any $t > 0$, $n > 0$, T, T_t are compact operators on $H^1(\mathcal{M})$ into $H^1(\mathcal{M})$; $T_t, T_{t,n}$ are compact operators on $C^1(\mathcal{M})$ into $C^1(\mathcal{M})$.*
2. *All eigenvalues of $T, T_t, T_{t,n}$ are real numbers. All generalized eigenvectors of $T, T_t, T_{t,n}$ are eigenvectors.*

Proof. First, it is well known that T is compact operator. $T_{t,n}$ is actually finite dimensional operator, so it is also compact. To show the compactness of T_t, we need the following formula,

$$T_t u = \frac{1}{w_t(\mathbf{x})} \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y}) T_t u(\mathbf{y}) \mathrm{d}\mathbf{y} + \frac{t}{w_t(\mathbf{x})} \int_{\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y}) u(\mathbf{y}) \mathrm{d}\mathbf{y}, \quad \forall u \in H^1(\mathcal{M}).$$

In this paper, we assume that the kernel function $R \in C^2$. Then, direct calculation gives that that $T_t u \in C^2$. This implies the compactness of T_t both in H^1 and C^1.

For the operator T, the conclusion (2) is well known. The proof of T_t and $T_{t,n}$ are very similar, so here we only present the proof for T_t.

Let λ be an eigenvalue of T_t and u is corresponding eigenfunction, then

$$L_t T_t u = \lambda L_t u$$

where

$$L_t f(\mathbf{x}) = \frac{1}{t} \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y})(f(\mathbf{x}) - f(\mathbf{y})) p(\mathbf{y}) \mathrm{d}\mathbf{y}. \tag{2.5}$$

It is easy to see that

$$L_t(T_t u)(\mathbf{x}) = \int_{\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y}) u(\mathbf{y}) p(\mathbf{y}) \mathrm{d}\mathbf{y} \tag{2.6}$$

Then, we get

$$\lambda = \frac{\int_M \int_{\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y}) u^*(\mathbf{x}) u(\mathbf{y}) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y}}{\int_{\mathcal{M}} u^*(\mathbf{x})(L_t u)(\mathbf{x}) \mathrm{d}\mathbf{x}}$$

where u^* is the complex conjugate of u. From the symmetry of L_t and $\bar{R}(\mathbf{x}, \mathbf{y})$, it is easy to show that $\lambda \in \mathbb{R}$.

Now, we turn to study the eigenfunctions. Let u be a generalized eigenfunction of T_t with multiplicity $m > 1$ associate with eigenvalue λ. Let $v = (T_t - \lambda)^{m-1} u$, $w = (T_t - \lambda)^{m-2} u$, then v is an eigenfunction of T_t and

$$T_t v = \lambda v, \quad (T_t - \lambda) w = v$$

By applying L_t on both sides of above two equations, we have

$$\begin{aligned} \lambda L_t v &= L_t(T_t v) = \int_{\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y}) v(\mathbf{y}) \mathrm{d}\mathbf{y}, \\ L_t v &= L_t(T_t w) - \lambda L_t w = \int_{\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y}) w(\mathbf{y}) \mathrm{d}\mathbf{y} - \lambda L_t w \end{aligned}$$

Using above two equations and the fact that L_t is symmetric, we get

$$\begin{aligned} 0 &= \left\langle w, \lambda L_t v - \int_{\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y}) v(\mathbf{y}) \mathrm{d}\mathbf{y} \right\rangle_{\mathcal{M}} \\ &= \left\langle \lambda L_t w - \int_{\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y}) w(\mathbf{y}) \mathrm{d}\mathbf{y}, v \right\rangle_{\mathcal{M}} \\ &= \langle L_t v, v \rangle_{\mathcal{M}} \geq C \left\| v \right\|_2^2 \end{aligned}$$

which implies that $(T_t - \lambda)^{m-1} u = v = 0$. This proves that u is a generalized eigenfunction of T_t with multiplicity $m-1$. Repeating this process, we can show that u is actually an eigenfunction of T_t. □

3 Main Result

Before stating the main theorem, we summarize the assumptions in this paper as follows:

Assumption 1.

- *Assumptions on the manifold: $\mathcal{M}$ is k-dimensional compact and C^∞ smooth manifold isometrically embedded in a Euclidean space $\mathbb{R}^d$.*
- *Assumptions on the sample points: $X = \{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ are sampled independently over the manifold $\mathcal{M}$ distribution $p(\mathbf{x}) \in C^1(\mathcal{M})$ and $\min_{\mathbf{x}\in\mathcal{M}} p(\mathbf{x}) > 0$, $\max_{\mathbf{x}\in\mathcal{M}} p(\mathbf{x}) < \infty$.*
- *Assumptions on the kernel function $R(r)$:*
  - *(a) $R \in C^2(\mathbb{R}^+)$;*
  - *(b) $R(r) \geq 0$ and $R(r) = 0$ for $\forall r > 1$;*
  - *(c) $\exists \delta_0 > 0$ so that $R(r) \geq \delta_0$ for $0 \leq r \leq \frac{1}{2}$.*

The main result in this paper is stated with the help of the Riesz spectral projection. Let X be a complex Banach space and $L : X \to X$ be a compact linear operator. The resolvent set $\rho(L)$ is given by the complex numbers $z \in \mathbb{C}$ such that $z - L$ is bijective. The spectrum of L is $\sigma(L) = \mathbb{C} \backslash \rho(L)$. It is well known that $\sigma(L)$ is a countable set with no limit points other than zero. All non-zero value s in $\sigma(L)$ are eigenvalues. If λ is a nonzero eigenvalue of L, the ascent multiplicity α of $\lambda - L$ is the smallest integer such that $\ker(\lambda - L)^\alpha = \ker(\lambda - L)^{\alpha+1}$.

Given a closed smooth curve $\Gamma \subset \rho(L)$ which encloses the eigenvalue λ and no other elements of $\sigma(L)$, the Riesz spectral projection associated with λ is defined by

$$E(\lambda, L) = \frac{1}{2\pi i} \int_\Gamma (z - L)^{-1} \mathrm{d}z, \tag{3.1}$$

where $i = \sqrt{-1}$ is the unit imaginary.

Theorem 3.1. *Under the assumptions in Assumption 1, let λ_i be the ith largest eigenvalue of T (same eigenvalue is repeated according to its multiplicity) with multiplicity α_i and $\phi_i^k, k = 1, \cdots, \alpha_i$ be the linear independent eigenfunctions corresponding to λ_i. Let $\lambda_i^{t,n}$ be the ith largest eigenvalue of $T_{t,n}$. With probability at least $1 - 1/n$, there exists a constant $C_1 > 0, C_2 > 0$ depend on $\mathcal{M}$, kernel function R, distribution p and spectra of T, such that*

$$|\lambda_i^{t,n} - \lambda_i| \leq C_1 \left(t^{1/2} + \frac{\log n + |\log t| + 1}{t^{k+3}\sqrt{n}} \right),$$

and

$$\|\phi_i^k - E(\sigma_i^{t,n}, T_{t,n})\phi_i^k\|_{H^1(\mathcal{M})} \leq C_2 \left(t^{1/2} + \frac{\log n + |\log t| + 1}{t^{k+2}\sqrt{n}} \right),$$

as long as n large enough. Here $\sigma_i^{t,n} = \{\lambda_j^{t,n} \in \sigma(T_{t,n}) : j \in I_i\}$ and $I_i = \{j \in \mathbb{N} : \lambda_j = \lambda_i\}$.

Remark 3.1. *In above theorem, $\sqrt{t}$ and $\sqrt{n}$ seem to be optimal. However, the factor multiply $\sqrt{n}$, t^{k+2} and t^{k+3}, are not optimal. We believe that it can be improved by obtaining better a prior estimate of the integral equation* (1.3)*. Now, we only get L^2 estimate. In the spectra convergence analysis, we need C^1 estimate. In this paper, the regularity is lifted by using the regularity of the kernel function. The trade off is that a factor $t^{-k/4}$ emerges which reduces the rate of convergence. If we can get a prior estimate of the integral equation in C^1, $t^{-k/4}$ can be removed.*

4 Proof of the Main Theorem (Theorem 3.1)

The proof of Theorem 3.1 consists of three parts, which are given in Section 4.1, 4,2 and 4.3 respectively. The first part is to establish the connection between the difference of the eigenvalues and eigenfunctions and the difference of solution operators $\|T - T_t\|_{H^1}$ and $\|T_t - T_{t,n}\|_{C^1}$ (Theorem 4.3). This is achieved by using the results in the perturbation theory of compact operators.

In the second part, we estimate the difference of operators $T - T_t$ and $T_t - T_{t,n}$ in H^1 and C^1 norm respectively, Theorem 4.4 and 4.5. This is the most difficult part. Comparing with the pointwise convergence which was proved in previous works, convergence in H^1 or C^1 is stronger, hence more difficult to prove.

Finally, we use the the theory of the Glivenko-Cantelli class in statistical learning to estimate the error in the Monte-Carlo integration. The key ingredient in this part is to estimate the covering number of some function classes.

Here, we list some notations which will be used in the proof. Some of them have been defined in previous sections. We also list them here for the convenience of readers.

- k: dimension of the underlying manifold; d: dimension of the ambient Euclidean space; n: number of sample points.
- C: positive constant independent on t and sample points X_n. We abuse the notation to denote all the constants independent on t and sample points X_n by C. It may be different in different places.
- $C_t = \frac{1}{(4\pi t)^{k/2}}$ is the normalize constant of kernel function R.
- $p(\mathbf{x})$: probability distribution function.
- R: kernel function. $\bar{R}(r) = \int_r^\infty R(s)\mathrm{d}s$.
- $R_t(\mathbf{x}, \mathbf{y}) = \frac{1}{(4\pi t)^{k/2}} R\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right), \quad \bar{R}_t(\mathbf{x}, \mathbf{y}) = \frac{1}{(4\pi t)^{k/2}} \bar{R}\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right).$
- $L_t f(\mathbf{x}) = \frac{1}{t} \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y})(f(\mathbf{x}) - f(\mathbf{y}))p(\mathbf{y})\mathrm{d}\mathbf{y}$.
- $L_{t,n} f(\mathbf{x}) = \frac{1}{nt} \sum_{j=1}^n R_t(\mathbf{x}, \mathbf{x}_j)(f(\mathbf{x}) - f(\mathbf{x}_j))$.
- $w_t(\mathbf{x}) = \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}, \quad w_{t,n}(\mathbf{x}) = \frac{1}{n(4\pi t)^{k/2}} \sum_{j=1}^n R\left(\frac{|\mathbf{x}-\mathbf{x}_j|^2}{4t}\right)$
- $w_{\min}, w_{\max}$: $w_{\min} = \inf_{t>0} \min_{\mathbf{x}\in\mathcal{M}} w_t(\mathbf{x})$, $w_{\max} = \sup_{t>0} \max_{\mathbf{x}\in\mathcal{M}} w_t(\mathbf{x})$. Under the assumption in Assumption 1, we can show that $0 < w_{\min}, w_{\max} < \infty$.

- $p(f) = \int_{\mathcal{M}} f(\mathbf{x}) p(\mathbf{x}) \mathrm{d}\mathbf{x}, \quad p_n(f) = \frac{1}{n} \sum_{i=1}^{n} f(\mathbf{x}_i)$.
- $\mathcal{R}_t = \left\{ R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M} \right\}$
- $\overline{\mathcal{R}}_t = \left\{ \bar{R}\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M} \right\}$
- $\mathcal{D}_t = \left\{ \nabla_{\mathbf{x}} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M} \right\}$
- $\mathcal{R}_t \cdot \mathcal{K}_{t,n} = \left\{ \frac{1}{w_{t,n}(\mathbf{y})} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) R\left(\frac{|\mathbf{z}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M},\ \mathbf{z} \in \mathcal{M} \right\}$
- $\overline{\mathcal{R}}_t \cdot \mathcal{K}_{t,n} = \left\{ \frac{1}{w_{t,n}(\mathbf{y})} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) \bar{R}\left(\frac{|\mathbf{z}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M},\ \mathbf{z} \in \mathcal{M} \right\}$
- $\overline{\mathcal{R}}_t \cdot \overline{\mathcal{K}}_{t,n} = \left\{ \frac{1}{w_{t,n}(\mathbf{y})} \bar{R}\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) \bar{R}\left(\frac{|\mathbf{z}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M},\ \mathbf{z} \in \mathcal{M} \right\}$
- $\mathcal{D}_t \cdot \mathcal{K}_{t,n} = \left\{ \frac{\sqrt{t}}{w_{t,n}(\mathbf{y})} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) \nabla_{\mathbf{z}} R\left(\frac{|\mathbf{z}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M},\ \mathbf{z} \in \mathcal{M} \right\}$
- $\overline{\mathcal{D}}_t \cdot \mathcal{K}_{t,n} = \left\{ \frac{\sqrt{t}}{w_{t,n}(\mathbf{y})} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) \nabla_{\mathbf{z}} \bar{R}\left(\frac{|\mathbf{z}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M},\ \mathbf{z} \in \mathcal{M} \right\}$
- $\overline{\mathcal{D}}_t \cdot \overline{\mathcal{K}}_{t,n} = \left\{ \frac{\sqrt{t}}{w_{t,n}(\mathbf{y})} \bar{R}\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) \nabla_{\mathbf{z}} \bar{R}\left(\frac{|\mathbf{z}-\mathbf{y}|^2}{4t}\right) : \mathbf{x} \in \mathcal{M},\ \mathbf{z} \in \mathcal{M} \right\}$

4.1 Perturbation results of Solution Operators

First, we need two theorems in [1] regarding the perturbation of the compact operators.

Theorem 4.1. *([1]) Let $(X, \|\cdot\|_X)$ be an arbitrary Banach space. Let S and T be compact linear operators on X into X. Let $z \in \rho(T)$. Assume*

$$\|(T-S)S\|_X \leq \frac{|z|}{\|(z-T)^{-1}\|_X}. \tag{4.1}$$

Then $z \in \rho(S)$ and $(z-S)^{-1}$ has the bound

$$\|(z-S)^{-1}\|_X \leq \frac{1 + \|S\|_X \|(z-T)^{-1}\|_X}{|z| - \|(z-T)^{-1}\|_X \|(T-S)S\|_X}. \tag{4.2}$$

Theorem 4.2. *([1]) Let $(X, \|\cdot\|_X)$ be an arbitrary Banach space. Let S and T be compact linear operators on X into X. Let $z_0 \in \mathbb{C}$, $z_0 \neq 0$ and let $\epsilon > 0$ be less than $|z_0|$, denote the circumference $|z - z_0| = \epsilon$ by Γ and assume $\Gamma \subset \rho(T)$. Denote the interior of Γ by U. Let $\sigma_T = U \cap \sigma(T) \neq \emptyset$. $\sigma_S = U \cap \sigma(S)$. Let $E(\sigma_S, S)$ and $E(\sigma_T, T)$ be the corresponding spectral projections of S for σ_S and T for σ_T, i.e.*

$$E(\sigma_S, S) = \frac{1}{2\pi i} \int_{\Gamma} (z-S)^{-1} \mathrm{d}z, \quad E(\sigma_T, T) = \frac{1}{2\pi i} \int_{\Gamma} (z-T)^{-1} \mathrm{d}z. \tag{4.3}$$

Assume

$$\|(T-S)S\|_X \leq \min_{z \in \Gamma} \frac{|z|}{\|(z-T)^{-1}\|_X} \tag{4.4}$$

Then, we have

(1). Dimension $E(\sigma_S, S)X = E(\sigma_T, T)X$, thereby σ_S is nonempty and of the same multiplicity as σ_T.

(2). For every $x \in X$,

$$\|E(\sigma_T, T)x - E(\sigma_S, S)x\|_X \leq \frac{M\epsilon}{c_0} \left(\|(T-S)x\|_X + \|x\|_X \|(T-S)S\|_X \right).$$

where $M = \max_{z\in\Gamma} \|(z-T)^{-1}\|_X$, $c_0 = \min_{z\in\Gamma} |z|$.

To apply above two theorems, we need some estimates of T and T_t which are summarized in three lemmas below.

Lemma 4.1. *Let T be the solution operator of the Neumann problem (2.1) and $z \in \rho(T)$, then*

$$\|(z-T)^{-1}\|_{H^1(\mathcal{M})} \leq \max_{n\in\mathbb{N}} \frac{1}{|z-\lambda_n|},$$

where $\{\lambda_n\}_{n\in\mathbb{N}}$ is the set of eigenvalues of T.

Proof. Suppose $\phi_j,\ j \in \mathbb{N}$ be the normalized eigenfunction of T corresponding to $\lambda_j,\ j \in \mathbb{N}$. Then it is well known that $\{\phi_j\}_{j\in\mathbb{N}}$ is a orthonormal basis of $H^1(\mathcal{M})$.

For any $x \in H^1(\mathcal{M})$, $z \in \rho(T)$, first we can expand x over $\{\phi_j\}_{j\in\mathbb{N}}$ to obtain

$$x = \sum_{j=1}^{\infty} c_j \phi_j. \tag{4.5}$$

Then, we have

$$\begin{aligned} \|(z-T)x\|_{H^1} &= \left\| \sum_{j=1}^{\infty} c_j (z-T)\phi_j \right\|_{H^1} = \left\| \sum_{j=1}^{\infty} c_j (z-\lambda_j)\phi_j \right\|_{H^1} \\ &= \left(\sum_{j=1}^{\infty} c_j^2 |z-\lambda_j|^2 \right)^{1/2} \geq \min_{n\in\mathbb{N}} |z-\lambda_n| \left(\sum_{j=1}^{\infty} c_j^2 \right)^{1/2} \\ &= \min_{n\in\mathbb{N}} |z-\lambda_n| \|x\|_{H^1} \end{aligned} \tag{4.6}$$

□

Lemma 4.2. *Let T_t be the solution operator of the integral equation (2.2). For any $z \in \mathbb{C}\backslash \bigcup_{n\in\mathbb{N}} B(\lambda_n, r_0)$ with $r_0 > \|T - T_t\|_{H^1}$, then*

$$\|(z-T_t)^{-1}\|_{C^1} \leq \max\left\{ \frac{2|\mathcal{M}|}{|z|t^{(k+2)/4}} \left(\min_{n\in\mathbb{N}} |z-\lambda_n| - \|T-T_t\|_{H^1} \right)^{-1}, \frac{2}{|z|} \right\}.$$

Proof. For any $x \in H^1(\mathcal{M})$,

$$\begin{aligned} \|(z-T_t)x\|_{H^1} &\geq \|(z-T)x\|_{H^1} - \|(T-T_t)x\|_{H^1} \\ &\geq \left(\min_{n\in\mathbb{N}} |z-\lambda_n| - \|T-T_t\|_{H^1} \right) \|x\|_{H^1} \end{aligned} \tag{4.7}$$

Then $(z-T_t)^{-1}$ exists and

$$\|(z-T_t)^{-1}\|_{H^1} \leq \left(\min_{n\in\mathbb{N}} |z-\lambda_n| - \|T-T_t\|_{H^1} \right)^{-1}. \tag{4.8}$$

For any $u \in C^1(\mathcal{M})$,

$$\|(z-T_t)^{-1}u\|_{H^1} \leq \left(\min_{n\in\mathbb{N}}|z-\lambda_n| - \|T-T_t\|_{H^1}\right)^{-1} |\mathcal{M}|\|u\|_{C^1} \tag{4.9}$$

where $|\mathcal{M}|$ is the volume of the manifold $\mathcal{M}$.

On the other hand, let $v = (z-T_t)^{-1}u$ which means $v = (u+T_tv)/z$

$$\begin{aligned}
\|v\|_{C^1} &\leq \frac{1}{|z|}\left(\|u\|_{C^1} + \|T_tv\|_{C^1}\right) \\
&\leq \frac{1}{|z|}\left(\|u\|_{C^1} + t^{-(k+2)/4}\|v\|_{L^2}\right) \\
&\leq \frac{1}{|z|}\left(\frac{|\mathcal{M}|}{t^{(k+2)/4}}\left(\min_{n\in\mathbb{N}}|z-\lambda_n| - \|T-T_t\|_{H^1}\right)^{-1} + 1\right)\|u\|_{C^1}
\end{aligned}$$

which proves that

$$\|(z-T_t)^{-1}\|_{C^1} \leq \max\left(\frac{2|\mathcal{M}|}{|z|t^{(k+2)/4}}\left(\min_{n\in\mathbb{N}}|z-\lambda_n| - \|T-T_t\|_{H^1}\right)^{-1}, \frac{2}{|z|}\right) \tag{4.10}$$

□

Lemma 4.3. *Let T_t be the solution operator of the integral equation (2.2) and λ_n be eigenvalues of T, then*

$$\sigma(T_t) \subset \bigcup_{n\in\mathbb{N}} B\left(\lambda_n, 2\|T-T_t\|_{H^1(\mathcal{M})}\right).$$

Proof. Let $r_0 = \|T-T_t\|_{H^1(\mathcal{M})}$, $\mathcal{A} = \mathbb{C}\backslash\bigcup_{n\in\mathbb{N}} B(\lambda_n, 2r_0)$. For any $z \in \mathcal{A}$, using Lemma 4.1, we have

$$\|(z-T)^{-1}\|_{H^1(\mathcal{M})} \leq \max_{n\in\mathbb{N}} \frac{1}{|z-\lambda_n|} \leq \frac{1}{2r_0}$$

which implies that

$$\|T-T_t\|_{H^1(\mathcal{M})} = r_0 \leq \frac{1}{2\|(z-T)^{-1}\|_{H^1(\mathcal{M})}}.$$

Then using Theorem 4.1, we have $z \in \rho(T_t)$.

Since z is arbitrary in $\mathcal{A}$, we get $\mathcal{A} \subset \rho(T_t)$. This means that

$$\sigma(T_t) = \mathbb{C}\backslash\rho(T_t) \subset \mathbb{C}\backslash\mathcal{A} = \bigcup_{n\in\mathbb{N}} B(\lambda_n, 2\|T-T_t\|_{H^1(\mathcal{M})}).$$

□

Now, we get the main theorem in this subsection.

Theorem 4.3. *Let λ_m be the mth largest eigenvalue of T with multiplicity α_m and $\phi_m^k, k = 1, \cdots, \alpha_m$ be the eigenfunctions corresponding to λ_m. Let $\lambda_m^{t,n}$ be the mth largest eigenvalue of $T_{t,n}$. Let $\gamma_m = \min_{j\leq m, \lambda_j\neq\lambda_{j+1}} |\lambda_j - \lambda_{j+1}|$ and*

$$\begin{aligned}
&\|(T_{t,n}-T_t)T_{t,n}\|_{C^1} \leq \min\{\frac{t}{2}, \frac{\gamma_m t^{k/4+3/2}}{24}, \frac{(|\lambda_m|-\gamma_m/3)^2 t^{(k+2)/4}\gamma_m}{12}, \frac{(|\lambda_m|-\gamma_m/3)^2}{2}\}, \\
&\|T-T_t\|_{H^1(\mathcal{M})} \leq \gamma_m/12, \quad \|(T-T_t)T_t\|_{H^1(\mathcal{M})} \leq (|\lambda_m|-\gamma_m/3)\gamma_m/3
\end{aligned}$$

Then there exists a constant C_1, C_2 depend on $\mathcal{M}$, the kernel function R, γ_m and λ_m, such that

$$|\lambda_m^{t,n} - \lambda_m| \leq \frac{2}{t^{k/4+3/2}} \|(T_{t,n} - T_t)T_{t,n}\|_{C^1} + \|T - T_t\|_{H^1(\mathcal{M})}$$

and

$$\|\phi_m^k - E(\sigma_m^{t,n}, T_{t,n})\phi_m^k\|_{H^1(\mathcal{M})} \leq C(\|(T-T_t)\phi_m^k\|_{H^1} + \|(T-T_t)T_t\|_{H^1}) + \frac{C}{t^{(k+2)/4}}(\|(T_t - T_{t,n})\phi_m^k\|_{C^1} + \|(T_t - T_{t,n})T_{t,n}\|_{C^1})$$

Here $\sigma_m^{t,n} = \{\lambda_j^{t,n} \in \sigma(T_{t,n}) : j \in I_m\}$ and $I_m = \{j \in \mathbb{N} : \lambda_j = \lambda_m\}$.

Proof. Let $r_1 = \frac{2}{t^{k/4+3/2}} \|(T_{t,n} - T_t)T_{t,n}\|_{C^1} + \|T - T_t\|_{H^1(\mathcal{M})}$, $\mathcal{A} = \mathbb{C} \backslash \left[\bigcup_{n\in\mathbb{N}} B(\lambda_n, r_1) \bigcup B(0, t^{1/2})\right]$.

For any $z \in \mathcal{A}$, using Theorem 4.1, we have

$$\begin{aligned} \|(z - T_t)^{-1}\|_{C^1} &\leq \frac{2|\mathcal{M}|}{|z|t^{(k+2)/4}} \left(\min_{n\in\mathbb{N}} |z - \lambda_n| - \|T - T_t\|_{H^1}\right)^{-1} \\ &\leq \frac{2|\mathcal{M}|}{t^{k/4+1}} (r_1 - \|T - T_t\|_{H^1})^{-1} \\ &= \frac{t^{1/2}}{\|(T_{t,n} - T_t)T_{t,n}\|_{C^1}} \leq \frac{|z|}{\|(T_{t,n} - T_t)T_{t,n}\|_{C^1}} \end{aligned}$$

or

$$\|(z - T_t)^{-1}\|_{C^1} \leq \frac{2}{|z|} \leq \frac{2}{t^{1/2}} \leq \frac{\sqrt{t}}{\|(T_{t,n} - T_t)T_{t,n}\|_{C^1}} \leq \frac{|z|}{\|(T_{t,n} - T_t)T_{t,n}\|_{C^1}}.$$

Here, we use the assumption that $\|(T_{t,n} - T_t)T_{t,n}\|_{C^1} \leq t/2$.

Above two inequalies together imply that

$$\|(T_{t,n} - T_t)T_{t,n}\|_{C^1} \leq \frac{|z|}{\|(z - T_t)^{-1}\|_{C^1}}.$$

Then using Theorem 4.1, we have $z \in \rho(T_{t,n})$.

Since z is arbitrary in $\mathcal{A}$, $\mathcal{A} \subset \rho(T_{t,n})$. This means that

$$\sigma(T_{t,n}) = \mathbb{C} \backslash \rho(T_{t,n}) \subset \mathbb{C} \backslash \mathcal{A} = \bigcup_{n\in\mathbb{N}} B(\lambda_n, r_1) \bigcup B(0, t^{1/2}). \tag{4.11}$$

Moreover, using Theorem 4.3 and the definition of r_1, we have

$$\sigma(T_t) \subset \bigcup_{n\in\mathbb{N}} B(\lambda_n, 2r_1). \tag{4.12}$$

For any fixed eigenvalue $\lambda_m \in \sigma(T)$, let $\gamma_m = \min_{j\leq m} |\lambda_j - \lambda_{j+1}|$. Using the structure of $\sigma(T)$, we know that $\gamma_m > 0$. Notice that

$$\frac{2}{t^{k/4+3/2}} \|(T_{t,n} - T_t)T_{t,n}\|_{C^1} \leq \gamma_m/12, \quad \|T - T_t\|_{H^1(\mathcal{M})} \leq \gamma_m/12,$$

which gives $r_1 < \gamma_m/6$.

Let $\Gamma_j = \{z \in \mathbb{C} : |z - \lambda_j| = \gamma_j/3\}$, U_j be the aera enclosed by Γ_j. Let

$$\sigma_{t,j} = \sigma(T_t) \bigcap U_j, \quad \sigma_{t,n,j} = \sigma(T_{t,n}) \bigcap U_j.$$

Using the definition of Γ_j, we know that for any $j \leq m$, $\Gamma_j \subset \rho(T), \rho(T_t)$ and $\rho(T_{t,n})$.

In order to apply Theorem 4.2, we need to verify the condition

$$\|(T-T_t)T_t\|_{H^1} \;\le\; \min_{z\in\Gamma_j}\frac{|z|}{\|(z-T)^{-1}\|_{H^1}}, \tag{4.13}$$

$$\|(T_t-T_{t,n})T_{t,n}\|_{C^1} \;\le\; \min_{z\in\Gamma_j}\frac{|z|}{\|(z-T_t)^{-1}\|_{C^1}}. \tag{4.14}$$

Using Lemma 4.1 and the choice of Γ_j, we have

$$\min_{z\in\Gamma_m}\frac{|z|}{\|(z-T)^{-1}\|_{H^1}} \ge \frac{\min_{z\in\Gamma_m}|z|}{\max_{z\in\Gamma_m}\|(z-T)^{-1}\|_{H^1}} \ge (|\lambda_m|-\gamma_m/3)\min_{z\in\Gamma_m,n\in\mathbb{N}}|z-\lambda_m| = (|\lambda_m|-\gamma_m/3)\gamma_m/3.$$

Then, using the assumption that $\|(T-T_t)T_t\|_{H^1(\mathcal{M})} \le (|\lambda_m|-\gamma_m/3)\gamma_m/3$, condition (4.13) is true..

Using Lemma 4.2, we have

$$\begin{aligned}\min_{z\in\Gamma_m}\frac{|z|}{\|(z-T_t)^{-1}\|_{C^1}} &\ge \frac{\min_{z\in\Gamma_m}|z|}{\max_{z\in\Gamma_m}\|(z-T_t)^{-1}\|_{C^1}}\\ &\ge \frac{(|\lambda_m|-\gamma_m/3)^2t^{(k+2)/4}}{2}\left(\min_{z\in\Gamma_m,n\in\mathbb{N}}|z-\lambda_m|-\|T-T_t\|_{H^1}\right)\\ &\ge \frac{(|\lambda_m|-\gamma_m/3)^2t^{(k+2)/4}\gamma_m}{12}.\end{aligned} \tag{4.15}$$

or

$$\min_{z\in\Gamma_m}\frac{|z|}{\|(z-T_t)^{-1}\|_{C^1}} \;\ge\; \frac{\min_{z\in\Gamma_m}|z|}{\max_{z\in\Gamma_m}\|(z-T_t)^{-1}\|_{C^1}} \ge \frac{(|\lambda_m|-\gamma_m/3)^2}{2}. \tag{4.16}$$

To get the last inequality of (4.15), we use the assumption that $\|T-T_t\|_{H^1}\le\gamma/6$ and $\min_{z\in\Gamma_m,n\in\mathbb{N}}|z-\lambda_m| = \gamma_m/3$.

Using the assumption that $\|(T-T_{t,n})T_{t,n}\|_{C^1(\mathcal{M})} \le \min\{\frac{(|\lambda_m|-\gamma_m/3)^2t^{(k+2)/4}\gamma_m}{12}, \frac{(|\lambda_m|-\gamma_m/3)^2}{2}\}$, condition (4.14) is satisfied.

Then using Theorem 4.2, we have

$$\dim(E(\lambda_m,T)) = \dim(E(\sigma_{t,m},T_t)) = \dim(E(\sigma_{t,n,m},T_{t,n})). \tag{4.17}$$

It follows from (4.11) that

$$|\lambda_m^{t,n}-\lambda_m| \le r_1 = \frac{2}{t^{k/4+3/2}}\|(T_{t,n}-T_t)T_{t,n}\|_{C^1} + \|T-T_t\|_{H^1(\mathcal{M})}. \tag{4.18}$$

The convergence of eigenspace is also given by Theorem 4.2. For any $x\in E(\lambda_m,T)$, $\|x\|_{C^1}=1$,

$$\|x-E(\sigma_{t,m},T_t)x\|_{H^1} \le \frac{\max_{z\in\Gamma_m}\|(z-T)^{-1}\|_{H^1}\gamma_m/3}{\min_{z\in\Gamma_m}|z|}(\|(T-T_t)x\|_{H^1} + \|(T-T_t)T_t\|_{H^1}\|x\|_{H^1}).$$

Using Lemma 4.1, we know that

$$\max_{z\in\Gamma_m}\|(z-T)^{-1}\|_{H^1} \le \max_{j\in\mathbb{N}}\frac{1}{|z-\lambda_j|} \le \frac{3}{2\gamma_m},$$

and $\min_{z\in\Gamma_m}|z| = |\lambda_m|-\gamma_m/3$. This implies that from Theorems 4.4,

$$\|x-E(\sigma_{t,m},T_t)x\|_{H^1} \le C(\|(T-T_t)x\|_{H^1} + \|(T-T_t)T_t\|_{H^1}\|x\|_{H^1}). \tag{4.19}$$

Regarding the convergence from $T_{t,n}$ to T_t, using Theorem 4.2 again, we have

$$\|E(\sigma_{t,m},T_t)x - E(\sigma_{t,n,m},T_{t,n})x\|_{C^1} \le \frac{\gamma_m\max_{z\in\Gamma_m}\|(z-T_t)^{-1}\|_{C^1}}{3\min_{z\in\Gamma_m}|z|}\left(\|(T_t-T_{t,n})x\|_{C^1} + \|(T_t-T_{t,n})T_{t,n}\|_{C^1}\right). \tag{4.20}$$

Using Lemma 4.2, we know that

$$\max_{z\in\Gamma_m}\|(z-T_t)^{-1}\|_{C^1} \le \max_{z\in\Gamma_m}\left\{\frac{2}{|z|t^{(k+2)/4}}\left(\min_{j\in\mathbb{N}}|z-\lambda_j|-\|T-T_t\|_{H^1}\right)^{-1},\frac{2}{|z|}\right\}$$
$$\le \max\left\{\frac{12}{\gamma_m(|\lambda_m|-\gamma_m/3)t^{(k+2)/4}},\frac{2}{|\lambda_m|-\gamma_m/3}\right\}. \tag{4.21}$$

To get the last inequality, we use that $\|T-T_t\|_{H^1}\le\gamma_m/6$ and $|z-\lambda_m|=\gamma_m/3$, $|z|\ge|\lambda_m-\gamma/3|$ for $z\in\Gamma_m$.

Then the proof is completed by (4.18), (4.19), (4.20) and (4.21).

□

4.2 Convergence of Solution Operators

Now, we estimate $\|T-T_t\|_{H^1}$ and $\|T_t-T_{t,n}\|_{C^1}$ respectively. The uniform bound of $\|T-T_t\|_{H^1}$ has been obtained in [14].

Theorem 4.4. *([14]) Under the assumptions in Assumption 1, there exists a constant $C>0$ only depends on $\mathcal{M}$ and the kernel function R, such that*

$$\|T-T_t\|_{H^1}\le Ct^{1/2},\quad \|T_t\|_{H^1}\le C.$$

Regarding $\|T_t-T_{t,n}\|_{C^1}$, we have following upper bound.

Theorem 4.5. *Under the assumptions in Assumption 1 and*

$$C_t\sup_{f\in\mathcal{R}_{t'}\cup\mathcal{R}_t\cup\mathcal{R}_{8t}}|p(f)-p_n(f)|\le w_{\min}/2, \tag{4.22}$$

$$C_t\sup_{f\in\mathcal{K}_{t',n}\cup\mathcal{K}_{t',n}\cdot\mathcal{K}_{t',n}}|p(f)-p_n(f)|\le\frac{\delta^2}{2\max\{w_{\max}+w_{\min}/2,2/w_{\min}\}}, \tag{4.23}$$

where $\delta=\frac{w_{\min}}{4w_{\max}+3w_{\min}}$, $t'=t/18$. There exists a constant C only depends on $\mathcal{M}$ and kernel function R, such that

$$\|(T_{t,n}-T_t)T_{t,n}\|_{C^1}\le\frac{Ch_0}{t^{3k/4+3/2}},\quad \|(T_{t,n}-T_t)f\|_{C^1}\le\frac{Ch(f)}{t^{3k/4+3/2}}.$$

where

$$\begin{aligned} h_0 &= \sup_{g\in\mathcal{R}_t\cdot\mathcal{K}_{t,n}\cup\mathcal{R}_t}|p_n(g)-p(g)|+t\sup_{g\in\mathcal{D}_t\cup\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{R}}_t\cup\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t\cup\mathcal{K}_{t,n}\cdot\mathcal{D}_t}|p_n(g)-p(g)| \\ &\quad +t^2\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{D}}_t}|p_n(g)-p(g)|+t^3\sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{D}}_t}|p_n(g)-p(g)|, \end{aligned} \tag{4.24}$$

$$\begin{aligned} h(f) &= \sup_{g\in\mathcal{R}_t\cdot\mathcal{K}_{t,n}\cup\mathcal{R}_t}|p_n(g)-p(g)|+t\sup_{g\in\mathcal{D}_t\cup f\cdot\overline{\mathcal{R}}_t\cup\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t\cup\mathcal{K}_{t,n}\cdot\mathcal{D}_t}|p_n(g)-p(g)| \\ &\quad +t^2\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{D}}_t}|p_n(g)-p(g)|+t^3\sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{D}}_t}|p_n(g)-p(g)|, \end{aligned} \tag{4.25}$$

The proof of this theorem can be found in Section 5.

4.3 Entropy bound

In this subsection, we will verify the assumptions (4.22), (4.23) in Theorem 4.5 and give upper bounds of h_0 and $h(f)$ defined in (4.24) and (4.25). To achieve these goals, we invoke a powerful theorem in empirical process theory.

Theorem 4.6. *(Theorem 2.3 in [17]) Let F be a class of functions from $\mathcal{M}$ to $[-1,1]$ and set μ to be a probability measure on $\mathcal{M}$. Let $(\mathbf{x}_i)_{i=1}^{\infty}$ be independent random variables distributed according to μ. For every $\epsilon > 0$ and any $n \geq 8/\epsilon^2$,*

$$\mathbb{P}\left(\sup_{f\in F} |\frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_i) - \int_{\mathcal{M}} f(\mathbf{x})\mu(\mathbf{x})\mathrm{d}\mathbf{x}| > \epsilon\right) \leq 8\mathbb{E}_{\mu}[N(\epsilon/8, F, L_1(\mu_n))]\exp(-n\epsilon^2/128) \tag{4.26}$$

In above theorem, $N(\epsilon, F, L_p(\mu_n)$ denotes the covering numbers of F at scale ϵ with respect to the $L_p(\mu_n)$ norm. μ_n is the empirical measure supported on one sample of $(\mathbf{x}_i)_{i=1}^{\infty}$. Let (Y, d) be a metric space and set $F \subset Y$. For every $\epsilon > 0$, denote by $N(\epsilon, F, d)$ the minimal number of open balls (with respect to the metric d) needed to cover F. That is, the minimal cardinality of the set $\{y_1, \cdots, y_m\} \subset Y$ with the property that every $f \in F$ has is some y_i such that $d(f, y_i) < \epsilon$. The set $\{y_1, \cdots, y_m\}$ is called an ϵ-cover of F . The logarithm of the covering numbers is called the entropy of the set. For every sample $\{x_1, \cdots, x_n\}$ let μ_n be the empirical measure supported on that sample. For $1 \leq p < \infty$ and a function f, $\|f\|_{L_p(\mu_n)} = \left(\frac{1}{n}\sum_{i=1}^{n} |f(x_i)|^p\right)^{1/p}$ and $\|f\|_{\infty} = \max_{1\leq i\leq n} |f(x_i)|$.

Notice that

$$L_1(\mu_n) \leq L_{\infty}(\mu_n) \leq L_{\infty}$$

where $\|f\|_{L_\infty} = \max_{\mathbf{x}\in\mathcal{M}} |f(\mathbf{x})|$. We get one immediate corollary of Theorem 4.6.

Corollary 4.1. *Let F be a class of functions from $\mathcal{M}$ to $[-1,1]$ and set μ to be a probability measure on $\mathcal{M}$. Let $(\mathbf{x}_i)_{i=1}^{\infty}$ be independent random variables distributed according to μ. For every $\epsilon > 0$ and any $n \geq 8/\epsilon^2$,*

$$\mathbb{P}\left(\sup_{f\in F} |\frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_i) - \int_{\mathcal{M}} f(\mathbf{x})\mu(\mathbf{x})\mathrm{d}\mathbf{x}| > \epsilon\right) \leq 8N(\epsilon/8, F, L_\infty)\exp(-n\epsilon^2/128) \tag{4.27}$$

where $N(\epsilon, F, L_\infty)$ be the covering numbers of F at scale ϵ with respect to the L_∞ norm

Then, we get an upper bound of $\sup_{f\in F} |\frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_i) - \int_{\mathcal{M}} f(\mathbf{x})\mu(\mathbf{x})\mathrm{d}\mathbf{x}|$.

Corollary 4.2. *Let F be a class of functions from $\mathcal{M}$ to $[-1,1]$. Let $(\mathbf{x}_i)_{i=1}^{\infty}$ be independent random variables distributed according to p, where p is the probability distribution in Assumption 1. Then with probability at least $1-\delta$,*

$$\sup_{f\in F} |p(f) - p_n(f)| \leq \sqrt{\frac{128}{n}\left(\ln N(\sqrt{\frac{2}{n}}, F, L_\infty) + \ln\frac{8}{\delta}\right)},$$

where

$$p(f) = \int_{\mathcal{M}} f(\mathbf{x})p(\mathbf{x})\mathrm{d}\mathbf{x}, \quad p_n(f) = \frac{1}{n}\sum_{i=1}^{n} f(\mathbf{x}_i). \tag{4.28}$$

Proof. Using Corollary 4.1, with probability at least $1-\delta$,

$$\sup_{f\in F} |p(f) - p_n(f)| \leq \epsilon_\delta,$$

where ϵ_δ is determined by

$$\epsilon_\delta = \sqrt{\frac{128}{n}\left(\ln N(\epsilon_\delta/8, F, L_\infty) + \ln\frac{8}{\delta}\right)}.$$

Obviously,

$$\epsilon_\delta \geq \sqrt{\frac{128}{n}} = 8\sqrt{\frac{2}{n}}$$

which gives that

$$N(\epsilon_\delta/8, F, L_\infty) \leq N(\sqrt{\frac{2}{n}}, F, L_\infty)$$

Then, we have

$$\epsilon_\delta \leq \sqrt{\frac{128}{n}\left(\ln N(\sqrt{\frac{2}{n}}, F, L_\infty) + \ln\frac{8}{\delta}\right)}$$

which proves the corollary. □

If the entropy bound of F is known, the upper bound of $\sup_{f\in F}|p(f) - p_n(f)|$ follows from Corollary 4.2. Now, the key point left becomes bounding the entropy of some given function class F.

Let us start from the function class $\mathcal{R}_t$. The functions in $\mathcal{R}_t$ are bounded uniformly, and the bound only depends on the kernel function R. To apply above corollary, we need to normalize $\mathcal{R}_t$ to make it lie in $[-1, 1]$. Here we also use $\mathcal{R}_t$ to denote the normalized function class and absorb the bound of $\mathcal{R}_t$ into the generic constant C. We do same normalize procedure for all function classes defined in Section 4.

Since the kernel $R \in C^2(\mathcal{M})$ and $\mathcal{M} \in C^\infty$, we have for any $\mathbf{x}, \mathbf{y} \in \mathcal{M}$

$$|R\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right) - R\left(\frac{\|\mathbf{z}-\mathbf{y}\|^2}{4t}\right)| \leq \frac{C}{\sqrt{t}}\|\mathbf{x}-\mathbf{z}\|.$$

This gives an easy bound of $N(\epsilon, \mathcal{R}_t, L_\infty)$,

$$N(\epsilon, \mathcal{R}_t, L_\infty) \leq \left(\frac{C}{\epsilon\sqrt{t}}\right)^k \tag{4.29}$$

Using Corollary 4.2, with probability at least $1 - 1/(2n)$,

$$\sup_{f\in\mathcal{R}_t\cup\mathcal{R}_{t'}\cup\mathcal{R}_{8t}} |p(f) - p_n(f)| \leq \frac{C}{\sqrt{n}}\left(\ln n - \ln t + 1\right)^{1/2} \tag{4.30}$$

This gives that

Corollary 4.3. *With probability at least* $1 - 1/(2n)$,

$$\sup_{f\in\mathcal{R}_t\cup\mathcal{R}_{t'}\cup\mathcal{R}_{8t}} |p(f) - p_n(f)| \leq \frac{w_{\min}}{2}$$

as long as n *is large enough such that the right hand side of* (4.30) *is less than* $w_{\min}/2$.

To get the covering number of $\mathcal{K}_{t,n}$, we need the assumption that $\sup_{f\in\mathcal{R}_t}|p(f) - p_n(f)| \leq \frac{w_{\min}}{2}$. Notice that

$$\left|\frac{1}{w_{t,n}(\mathbf{y})}\left[R\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right) - R\left(\frac{\|\mathbf{z}-\mathbf{y}\|^2}{4t}\right)\right]\right| \leq \frac{2}{w_{\min}}|R\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right) - R\left(\frac{\|\mathbf{z}-\mathbf{y}\|^2}{4t}\right)| \leq \frac{C}{\sqrt{t}}|\mathbf{x}-\mathbf{y}|$$

The first inequality comes from the fact that $\min_{\mathbf{z}\in\mathcal{M}} w_{t,n}(\mathbf{z}) \geq w_{\min}/2$ which is guaranteed by the assumption that $\sup_{f\in\mathcal{R}_t}|p(f) - p_n(f)| \leq \frac{w_{\min}}{2}$. This gives a bound

$$N(\epsilon, \mathcal{K}_{t,n}, L_\infty) \leq \left(\frac{C}{\epsilon\sqrt{t}}\right)^k. \tag{4.31}$$

Similarly, we can get

$$N(\epsilon, \mathcal{K}_{t,n} \cdot \mathcal{K}_{t,n}, L_\infty) \leq \left(\frac{C}{\epsilon\sqrt{t}}\right)^{2k} \tag{4.32}$$

Using Corollary 4.2, if $\sup_{f\in\mathcal{R}_t} |p(f) - p_n(f)| \leq \frac{w_{\min}}{2}$, then

$$\sup_{f\in\mathcal{K}_{t,n}\cup\mathcal{K}_{t,n}\cdot\mathcal{K}_{t,n}} |p(f) - p_n(f)| \leq \frac{C}{\sqrt{n}} \left(\ln n - \ln t + 1\right)^{1/2} \tag{4.33}$$

with probability at least $1 - 1/(2n)$. Moreover, from Corollary 4.3, the assumption $\sup_{f\in\mathcal{R}_t} |p(f) - p_n(f)| \leq \frac{w_{\min}}{2}$ holds with probability at least $1 - 1/(2n)$. By integrating these results together, we obtain

Corollary 4.4. *With probability at least* $1 - 1/n$,

$$\sup_{f\in\mathcal{K}_{t,n}\cup\mathcal{K}_{t,n}\cdot\mathcal{K}_{t,n}} |p(f) - p_n(f)| \leq \frac{\delta^2}{2\max\{w_{\max} + w_{\min}/2, 2/w_{\min}\}}$$

as long as n *is large enough. Here* $\delta = \frac{w_{\min}}{4w_{\max}+3w_{\min}}$.

Using similar techniques, we can get the estimate of h_0 and $h(f)$ in (4.24) and (4.25). Putting all bounds in Theorem 4.4, we get

Theorem 4.7. *Let* ϕ *be an eigenfunction of* T. *With probability at least* $1 - 1/n$,

$$\begin{aligned}
\|(T_t - T_{t,n})T_{t,n}\|_{C^1} \leq & \frac{C}{t^{3k/4+3/2}\sqrt{n}} \left(\ln n - \ln t + 1\right)^{1/2}, \\
\|(T_t - T_{t,n})\phi\|_{C^1} \leq & \frac{C_\phi}{t^{3k/4+3/2}\sqrt{n}} \left(\ln n - \ln t + 1\right)^{1/2}
\end{aligned}$$

as long as n *is large enough. Here* C_ϕ *is a constant depends on* $\mathcal{M}$, *kernel function* R, *distribution* p *and eigenfunction* ϕ.

The main theorem, Theorem 3.1, is an easy corollary of Theorem 4.3 and 4.7.

5 Proof of Theorem 4.5

To prove Theorem 4.5, first we prove the convergence in L^2 and then lift the convergence from L^2 to C^1 by using the regularity of the kernel function. The calculus is a little tedious. However, the method is rather standard.

In $L^2(\mathcal{M})$ space, we have

Theorem 5.1. *Under the assumptions in Assumption 1. Let* $f \in C(\mathcal{M})$ *in both problems, then there exists constants* $C > 0$, *so that*

$$\begin{aligned}
\|(T_{t,n} - T_t)T_{t,n}f\|_{L^2(\mathcal{M})} \leq & \frac{C}{t^{k/2+1}}\|f\|_\infty \left(\sup_{g\in\mathcal{R}_t\cup\mathcal{R}_t\cdot\mathcal{K}_{t,n}} |p_n(g) - p(g)| + t \sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{R}}_t\cup\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t} |p_n(g) - p(g)| \right) \\
\|(T_{t,n} - T_t)f\|_{L^2(\mathcal{M})} \leq & \frac{C}{t^{k/2+1}}\|f\|_\infty \left(\sup_{g\in\mathcal{R}_t\cup\mathcal{R}_t\cdot\mathcal{K}_{t,n}} |p_n(g) - p(g)| + t \sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t\cup f\cdot\overline{\mathcal{R}}_t} |p_n(g) - p(g)| \right),
\end{aligned}$$

as long as t *small enough and* (4.22), (4.23) *are satisfied.*

To prove Theorem 5.1, we need two theorems regarding the *a prior* estimate of the discrete solution and the stability of the integral operator L_t.

Theorem 5.2. *Suppose* $\mathbf{u} = (u_1, \cdots, u_n)^t$ *with* $\sum_{i=1}^n u_i = 0$ *solves the problem* (2.4) *and* $f \in C(\mathcal{M})$. *Then there exists a constant* $C > 0$ *only depends on* $\mathcal{M}$ *and kernel function* R, *such that*

$$\left(\frac{1}{n}\sum_{i=1}^n u_i^2\right)^{1/2} \leq C\left(\frac{1}{n}\sum_{i=1}^n f(\mathbf{x}_i)^2\right)^{1/2} \leq C\|f\|_\infty,$$

as long as (4.22), (4.23) *are satisfied.*

Theorem 5.3. *([14]) Under the assumptions in Assumption 1, assume* $u(\mathbf{x})$ *solves the following equation*

$$-L_t u = r, \tag{5.1}$$

where

$$L_t u = \frac{C_t}{t}\int_{\mathcal{M}} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right)(u(\mathbf{x}) - u(\mathbf{y}))p(\mathbf{y})\mathrm{d}\mathbf{y}. \tag{5.2}$$

Then, there exist constants $C > 0, T_0 > 0$ *independent on* t, *such that*

$$\|u\|_{L^2(\mathcal{M})} \leq C\|r\|_{L^2(\mathcal{M})}. \tag{5.3}$$

as long as $t \leq T_0$.

Theorem 5.3 has been proved in [14]. Theorem 5.2 is an easy corollary of following theorem which is proved in the appendix.

Theorem 5.4. *Under the assumption in Assumption 1 and assume* (4.22), (4.23) *hold. There exist constants* $C > 0$ *only depends on* $\mathcal{M}$ *and kernel function* R, *so that for any* $\mathbf{u} = (u_1, \cdots, u_n)^t \in \mathbb{R}^d$ *with* $\sum_{i=1}^n u_i = 0$,

$$\frac{1}{n^2 t}\sum_{i,j=1}^n R_t(\mathbf{x}_i, \mathbf{x}_j)(u_i - u_j)^2 \geq \frac{C}{n}\sum_{i=1}^n u_i^2. \tag{5.4}$$

Now we can give the proof of Theorem 5.1.

Proof. of Theorem 5.1
First, denote

$$u_{t,n}(\mathbf{x}) = T_{t,n}f = \frac{1}{n\, w_{t,n}(\mathbf{x})}\left(\sum_{j=1}^n R_t(\mathbf{x}, \mathbf{x}_j)u_j - t\sum_{j=1}^n \bar{R}_t(\mathbf{x}, \mathbf{x}_j)f_j\right) \tag{5.5}$$

where $\mathbf{u} = (u_1, \cdots, u_n)^t$ with $\sum_{i=1}^n u_i = 0$ solves the problem (2.4), $f_j = f(\mathbf{x}_j)$ and $w_{t,n}(\mathbf{x}) = \frac{1}{n}\sum_{j=1}^n R_t(\mathbf{x}, \mathbf{x}_j)$. And denote

$$v_{t,n}(\mathbf{x}) = T_{t,n}u_{t,n} = \frac{1}{n\, w_{t,n}(\mathbf{x})}\left(\sum_{j=1}^n R_t(\mathbf{x}, \mathbf{x}_j)v_j - t\sum_{j=1}^n \bar{R}_t(\mathbf{x}, \mathbf{x}_j)u_j\right) \tag{5.6}$$

where $\mathbf{v} = (v_1, \cdots, v_n)^t$ with $\sum_{i=1}^n v_i = 0$ solves

$$-\frac{1}{nt}\sum_{j=1}^n R_t(\mathbf{x}_i, \mathbf{x}_j)(v_i - v_j) = \frac{1}{n}\sum_{j=1}^n \bar{R}_t(\mathbf{x}_i, \mathbf{x}_j)u_j. \tag{5.7}$$

It follows from Theorem 5.4 that there exists a constant $C > 0$ independent on t and n such that

$$\left(\frac{1}{n}\sum_{i=1}^{n} u_i^2\right)^{1/2} \le C\|f\|_\infty, \quad \left(\frac{1}{n}\sum_{i=1}^{n} v_i^2\right)^{1/2} \le C\left(\frac{1}{n}\sum_{i=1}^{n} u_i^2\right)^{1/2} \le C\|f\|_\infty \tag{5.8}$$

The idea to prove the theorem is using Theorem 5.3. Then we need to estimate $\|L_t(T_{t,n} - T_t)T_{t,n}f\|_2$ and $\|L_t(T_{t,n} - T_t)f\|_2$ for any $f \in C(\mathcal{M})$.

For any $f \in C(\mathcal{M})$,

$$\begin{aligned} L_t(T_{t,n} - T_t)T_{t,n}f &= (L_tT_{t,n}T_{t,n}f - L_{t,n}T_{t,n}T_{t,n}f) + (L_{t,n}T_{t,n}T_{t,n}f - L_tT_tT_{t,n}f) \\ &= (L_tv_{t,n} - L_{t,n}v_{t,n}) + (L_{t,n}T_{t,n}u_{t,n} - L_tT_tu_{t,n})\,. \end{aligned} \tag{5.9}$$

Next, we estimate two terms of right hand side of (5.9) separately. For convenience, we split $v_{t,n} = a_{t,n} + b_{t,n}$ and

$$a_{t,n}(\mathbf{x}) = \frac{1}{n\,w_{t,n}(\mathbf{x})}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j)v_j, \tag{5.10}$$

$$b_{t,n}(\mathbf{x}) = -\frac{t}{n\,w_{t,n}(\mathbf{x})}\sum_{j=1}^{n} \bar{R}_t(\mathbf{x},\mathbf{x}_j)u_j. \tag{5.11}$$

For $\|L_tb_{t,n} - L_{t,n}b_{t,n}\|_2$, we have

$$\begin{aligned} &|(L_tb_{t,n} - L_{t,n}b_{t,n})(\mathbf{x})| \\ =\ & \frac{1}{t}\left|\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})(b_{t,n}(\mathbf{x}) - b_{t,n}(\mathbf{y}))p(\mathbf{y})\mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j)(b_{t,n}(\mathbf{x}) - b_{t,n}(\mathbf{x}_j))\right| \\ \le\ & \frac{1}{t}\,|b_{t,n}(\mathbf{x})|\left|\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j)\right| \\ &+\frac{1}{t}\left|\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})b_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j)b_{t,n}(\mathbf{x}_j)\right| \end{aligned} \tag{5.12}$$

The first term of (5.12) can be bounded as following,

$$\left\|b_{t,n}(\mathbf{x})\left(\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j)\right)\right\|_{L^2} \le C_t\|b_{t,n}\|_{L^2}\sup_{g\in\mathcal{R}_t}|p_n(g) - p(g)| \tag{5.13}$$

and

$$\begin{aligned} \|b_{t,n}\|_{L^2}^2 &= \frac{t^2}{n^2}\int_{\mathcal{M}}\left(\frac{1}{w_{t,n}(\mathbf{x})}\sum_{j=1}^{n}\bar{R}_t(\mathbf{x},\mathbf{x}_j)u_j\right)^2 p(\mathbf{x})\mathrm{d}\mathbf{x} \\ &\le \frac{Ct^2}{n}\int_{\mathcal{M}}\left(\frac{1}{n}\sum_{j=1}^{n}\bar{R}_t(\mathbf{x},\mathbf{x}_j)\right)\left(\sum_{j=1}^{n}\bar{R}_t(\mathbf{x},\mathbf{x}_j)u_j^2\right)p(\mathbf{x})\mathrm{d}\mathbf{x} \\ &\le \frac{Ct^2}{n}\sum_{j=1}^{n}\left(u_j^2\int_{\mathcal{M}}\bar{R}_t(\mathbf{x},\mathbf{x}_j)p(\mathbf{x})\mathrm{d}\mathbf{x}\right) \\ &\le \frac{Ct^2}{n}\sum_{j=1}^{n}u_j^2 \le Ct^2\|f\|_\infty, \end{aligned} \tag{5.14}$$

where last inequality comes from (5.8).

For the second term of (5.12),

$$\begin{aligned}
&\left|\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y}) b_{t,n}(\mathbf{y}) p(\mathbf{y}) \mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j) b_{t,n}(\mathbf{x}_j)\right| \\
=\ & \frac{t}{n}\left|\int_{\mathcal{M}} \frac{R_t(\mathbf{x},\mathbf{y})}{w_{t,n}(\mathbf{y})}\left(\sum_{\mathbf{x}_k\in P}\bar{R}_t(\mathbf{y},\mathbf{x}_k)u_k\right) p(\mathbf{y})\mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n}\frac{R_t(\mathbf{x},\mathbf{x}_j)}{w_{t,n}(\mathbf{x}_j)}\sum_{\mathbf{x}_k\in P}\bar{R}_t(\mathbf{x}_j,\mathbf{x}_k)u_k\right| \\
\le\ & \frac{t}{n}\sum_{k=1}^{n}|u_k|\left|\int_{\mathcal{M}}\frac{R_t(\mathbf{x},\mathbf{y})}{w_{t,n}(\mathbf{y})}\bar{R}_t(\mathbf{y},\mathbf{x}_k)p(\mathbf{y})\mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n}\frac{R_t(\mathbf{x},\mathbf{x}_j)}{w_{t,n}(\mathbf{x}_j)}\bar{R}_t(\mathbf{x}_j,\mathbf{x}_k)\right| \qquad (5.15)
\end{aligned}$$

Let

$$\begin{aligned}
A\ =\ & C_t\int_{\mathcal{M}}\frac{1}{w_{t,n}(\mathbf{y})}R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right)\bar{R}\left(\frac{|\mathbf{x}_i-\mathbf{y}|^2}{4t}\right)p(\mathbf{y})\mathrm{d}\mathbf{y} \\
-\ & \frac{C_t}{n}\sum_{j=1}^{n}\frac{1}{w_{t,n}(\mathbf{x}_j)}R\left(\frac{|\mathbf{x}-\mathbf{x}_j|^2}{4t}\right)\bar{R}\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t}\right). \qquad (5.16)
\end{aligned}$$

We have

$$|A| < C_t \sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| \qquad (5.17)$$

for some constant C independent of t. In addition, notice that only when $|\mathbf{x}-\mathbf{x}_i|^2 \le 16t$ is $A \neq 0$, which implies

$$|A| \le \frac{1}{\delta_0}|A|R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{32t}\right). \qquad (5.18)$$

Using these properties of A, we obtain

$$\begin{aligned}
&\left|\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y}) b_{t,n}(\mathbf{y}) p(\mathbf{y}) \mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j) b_{t,n}(\mathbf{x}_j)\right| \\
\le\ & \frac{Ct}{n}|A|_\infty\sum_{k=1}^{n}|u_k|R\left(\frac{|\mathbf{x}-\mathbf{x}_k|^2}{32t}\right) \\
\le\ & \frac{Ct}{n}\sum_{k=1}^{n}C_t|u_k|R\left(\frac{|\mathbf{x}-\mathbf{x}_k|^2}{32t}\right)C_t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| \qquad (5.19)
\end{aligned}$$

It follows that

$$\begin{aligned}
&\left\|\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y}) b_{t,n}(\mathbf{y}) p(\mathbf{y}) \mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{x}_j) b_{t,n}(\mathbf{x}_j)\right\|_2 \\
\le\ & Ct\left(\int_{\mathcal{M}}\left(\frac{1}{n}\sum_{k=1}^{n}C_t|u_k|R\left(\frac{|\mathbf{x}-\mathbf{x}_k|^2}{32t}\right)\right)^2 p(\mathbf{x})\mathrm{d}\mathbf{x}\right)^{1/2} C_t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| \\
\le\ & Ct\left(\frac{1}{n}\sum_{k=1}^{n}u_k^2\right)^{1/2} C_t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| \\
\le\ & Ct\|f\|_\infty C_t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| \qquad (5.20)
\end{aligned}$$

To get the second inequality, we use the condition that $C_t \sup_{g\in\mathcal{R}_{8t}} \le w_{\min}/2$.

Now we have complete upper bound of $\|L_t b_{t,n} - L_{t,n}b_{t,n}\|_{L_2}$ using (5.12), (5.13) and (5.20) and $C_t = \frac{1}{(4\pi t)^{k/2}}$,

$$\|L_t b_{t,n} - L_{t,n}b_{t,n}\|_{L^2(\mathcal{M})} \le \frac{C}{t^{k/2}}\|f\|_\infty \left(\sup_{g\in\mathcal{R}_t\cup\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t} |p_n(g)-p(g)|\right). \tag{5.21}$$

Mimicking the derivation of (5.21), we have

$$\|L_t a_{t,n} - L_{t,n}a_{t,n}\|_{L^2(\mathcal{M})} \le \frac{C}{t^{k/2+1}}\|f\|_\infty \left(\sup_{g\in\mathcal{R}_t\cup\mathcal{K}_{t,n}\cdot\mathcal{R}_t} |p_n(g)-p(g)|\right) \tag{5.22}$$

And consequently,

$$\begin{aligned}
&\|L_t v_{t,n} - L_{t,n}v_{t,n}\|_{L^2(\mathcal{M})}\\
\le\ &\|L_t a_{t,n} - L_{t,n}a_{t,n}\|_{L^2(\mathcal{M})} + \|L_t b_{t,n} - L_{t,n}b_{t,n}\|_{L^2(\mathcal{M})}\\
\le\ &\frac{C}{t^{k/2+1}}\|f\|_\infty \left(\sup_{g\in\mathcal{R}_t\cup\mathcal{K}_{t,n}\cdot\mathcal{R}_t} |p_n(g)-p(g)| + t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t} |p_n(g)-p(g)|\right).
\end{aligned} \tag{5.23}$$

The second term of (5.9) can be bounded as following,

$$\begin{aligned}
&L_t(T_t u_{t,n}) - L_{t,n}(T_{t,n}u_{t,n})\\
\le\ &\int_{\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y} - \frac{1}{n}\sum_{j=1}^n \bar{R}_t(\mathbf{x},\mathbf{x}_j)u_j\\
\le\ &\frac{1}{n^2}\sum_{j=1}^n \frac{\overline{R}_t(\mathbf{x},\mathbf{x}_j)}{w_{t,n}(\mathbf{x}_j)}\left(\sum_{k=1}^n R_t(\mathbf{x}_j,\mathbf{x}_k)u_k - t\sum_{k=1}^n \overline{R}_t(\mathbf{x}_j,\mathbf{x}_k)f_k\right)\\
&-\frac{1}{n}\int_{\mathcal{M}} \frac{\overline{R}_t(\mathbf{x},\mathbf{y})}{w_{t,n}(\mathbf{y})}\left(\sum_{k=1}^n R_t(\mathbf{y},\mathbf{x}_k)u_k - t\sum_{k=1}^n \overline{R}_t(\mathbf{y},\mathbf{x}_k)f_k\right)p(\mathbf{y})\mathrm{d}\mathbf{y}\\
=\ &\frac{1}{n}\sum_{k=1}^n u_k\left(\frac{1}{n}\sum_{j=1}^n \frac{\overline{R}_t(\mathbf{x},\mathbf{x}_j)}{w_{t,n}(\mathbf{x}_j)}R_t(\mathbf{x}_j,\mathbf{x}_k) - \int_{\mathcal{M}}\frac{\overline{R}_t(\mathbf{x},\mathbf{y})}{w_{t,n}(\mathbf{y})}R_t(\mathbf{y},\mathbf{x}_k)p(\mathbf{y})\mathrm{d}\mathbf{y}\right)\\
&-\frac{t}{n}\sum_{k=1}^n f_k\left(\frac{1}{n}\sum_{j=1}^n \frac{\overline{R}_t(\mathbf{x},\mathbf{x}_j)}{w_{t,n}(\mathbf{x}_j)}\overline{R}_t(\mathbf{x}_j,\mathbf{x}_k) - \int_{\mathcal{M}}\frac{\overline{R}_t(\mathbf{x},\mathbf{y})}{w_{t,n}(\mathbf{y})}\overline{R}_t(\mathbf{y},\mathbf{x}_k)p(\mathbf{y})\mathrm{d}\mathbf{y}\right).
\end{aligned} \tag{5.24}$$

Using the similar derivation from (5.15) to (5.21), we get

$$\begin{aligned}
&\|L_t(T_t u_{t,n}) - L_{t,n}(T_{t,n}u_{t,n})\|_{L^2}\\
\le\ &C\left(\frac{1}{n}\sum_{j=1}^n u_j^2\right)^{1/2} C_t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| + Ct\|f\|_\infty C_t\sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)|\\
\le\ &\frac{C}{t^{k/2}}\|f\|_\infty\left(\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| + t\sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)|\right).
\end{aligned} \tag{5.25}$$

The complete estimate follows from Equation (5.23) and (5.24).

$$\begin{aligned}
\|L_t(T_{t,n}-T_t)T_{t,n}f\|_{L^2(\mathcal{M})} \le\ &\frac{C}{t^{k/2+1}}\|f\|_\infty\left(\sup_{g\in\mathcal{R}_t\cup\mathcal{R}_t\cdot\mathcal{K}_{t,n}}|p_n(g)-p(g)|\right.\\
&\left.+t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| + t^2\sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)|\right).
\end{aligned} \tag{5.26}$$

Similarly, we can also get

$$\|L_t(T_{t,n}-T_t)f)\|_{L^2(\mathcal{M})} \leq \frac{C}{t^{k/2+1}}\|f\|_\infty \left(\sup_{g\in\mathcal{R}_t\cup\mathcal{R}_t\cdot\mathcal{K}_{t,n}} |p_n(g)-p(g)| \right.$$
$$\left. +t \sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t} |p_n(g)-p(g)| + t^2 \sup_{g\in f\cdot\overline{\mathcal{R}}_t} |p_n(g)-p(g)| \right). \tag{5.27}$$

The theorem is proved by using Theorem 5.3 and above two estimates (5.26), (5.27)

□

Now, we can prove Theorem 4.5 after one techinical lemma.

Lemma 5.1. *Under the assumption in Assumption 1 and assume* (4.22), (4.23) *hold. Then, there exist constants $C>0$ only depends on $\mathcal{M}$ and kernel function R, such that for any $f\in C(\mathcal{M})$,*

$$\|T_{t,n}f\|_\infty \leq Ct^{-k/4}\|f\|_\infty, \quad \|T_{t,n}f\|_{L^2} \leq C\|f\|_\infty.$$

Proof. From the definition of $T_{t,n}$, we have for any $f\in C(\mathcal{M})$

$$T_{t,n}f = \frac{C_t}{nw_{t,n}(\mathbf{x})}\sum_{i=1}^n R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{4t}\right)u_i + \frac{tC_t}{nw_{t,n}(\mathbf{x})}\sum_{i=1}^n R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{4t}\right)f(\mathbf{x}_i)$$

where $(u_1,\cdots,u_n)$ satisfies the equation

$$\frac{C_t}{nt}\sum_{j=1}^n R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t}\right)(u_i-u_j) = \frac{C_t}{n}\sum_{j=1}^n R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t}\right)f(\mathbf{x}_j).$$

Using Theorem 5.4, it is easy to get that

$$\left(\frac{1}{n}\sum_{i=1}^n u_i^2\right)^{1/2} \leq C\|f\|_\infty$$

where $C>0$ is a constant only depends on $\mathcal{M}$ and kernel function R.

Then

$$\begin{aligned}
|T_{t,n}f| \leq& \left(\frac{C_t}{nw_{t,n}(\mathbf{x})}\sum_{i=1}^n R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{4t}\right)\right)^{1/2}\left(\frac{C_t}{nw_{t,n}(\mathbf{x})}\sum_{i=1}^n R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{4t}\right)u_i^2\right)^{1/2} \\
&+ \frac{tC_t}{nw_{t,n}(\mathbf{x})}\sum_{i=1}^n R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{4t}\right)\|f\|_\infty \\
\leq& \left(\frac{C_t}{nw_{t,n}(\mathbf{x})}\sum_{i=1}^n R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{4t}\right)u_i^2\right)^{1/2} + t\|f\|_\infty \\
\leq& \left(\frac{2C_t}{w_{\min}}\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^n u_i^2\right)^{1/2} + t\|f\|_\infty \leq C\|f\|_\infty.
\end{aligned}$$

and

$$\begin{aligned}
\|T_{t,n}f\|_{L^2}^2 \leq& 2\int_{\mathcal{M}} \frac{C_t}{nw_{t,n}(\mathbf{x})}\sum_{i=1}^n R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{4t}\right)u_i^2 p(\mathbf{x})\mathrm{d}\mathbf{x} + 2t^2\|f\|_\infty^2 \\
\leq& C\left(\frac{1}{n}\sum_{i=1}^n u_i^2 + t^2\|f\|_\infty^2\right) \leq C\|f\|_\infty^2.
\end{aligned}$$

□

Finally, we get fully prepared to prove Theorem 4.5.

Proof. of Theorem 4.5:

For any $f \in C^1(\mathcal{M})$, let $u_{t,n} = T_{t,n}f$ and $v_i = T_{t,n}u_{t,n}(\mathbf{x}_i),\ i = 1, \cdots, n$. Using the definition of T_t and $T_{t,n}$, $T_t u_{t,n}$ and $T_{t,n}u_{t,n}$ have following representations

$$\begin{aligned} T_t u_{t,n} &= \frac{1}{w_t(\mathbf{x})}\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})T_t u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y} + \frac{t}{w_t(\mathbf{x})}\int_{\mathcal{M}} \bar{R}(\mathbf{x},\mathbf{y})u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}, \\ T_{t,n}u_{t,n} &= \frac{1}{n\,w_{t,n}(\mathbf{x})}\sum_{i=1}^n R_t(\mathbf{x},\mathbf{x}_i)v_i + \frac{t}{n\,w_{t,n}(\mathbf{x})}\sum_{i=1}^n \bar{R}(\mathbf{x},\mathbf{x}_i)u_i. \end{aligned} \tag{5.28}$$

where $u_i = u_{t,n}(\mathbf{x}_i),\ i = 1, \cdots, n$. We know that $(u_1, \cdots, u_n)$ and $(v_1, \cdots, v_n)$ satisfy following equations respectively

$$\begin{aligned} \frac{1}{nt}\sum_{j=1}^n R_t(\mathbf{x}_i,\mathbf{x}_j)(u_i - u_j) &= \frac{1}{n}\sum_{i=1}^n R_t(\mathbf{x}_i,\mathbf{x}_j)f(\mathbf{x}_j), \\ \frac{1}{nt}\sum_{j=1}^n R_t(\mathbf{x}_i,\mathbf{x}_j)(v_i - v_j) &= \frac{1}{n}\sum_{i=1}^n R_t(\mathbf{x}_i,\mathbf{x}_j)u_j. \end{aligned}$$

Using Theorem 5.2, we have

$$\left(\frac{1}{n}\sum_{i=1}^n u_i^2\right)^{1/2} \le C\|f\|_\infty, \quad \left(\frac{1}{n}\sum_{i=1}^n v_i^2\right)^{1/2} \le C\left(\frac{1}{n}\sum_{i=1}^n u_i^2\right)^{1/2} \le C\|f\|_\infty \tag{5.29}$$

Denote

$$\begin{aligned} T_t^1 u_{t,n} &= \frac{1}{w_{t,n}(\mathbf{x})}\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})T_t u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y} + \frac{t}{w_{t,n}(\mathbf{x})}\int_{\mathcal{M}} \bar{R}(\mathbf{x},\mathbf{y})u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}, \\ T_t^2 u_{t,n} &= \frac{1}{w_{t,n}(\mathbf{x})}\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})T_{t,n} u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y} + \frac{t}{w_{t,n}(\mathbf{x})}\int_{\mathcal{M}} \bar{R}(\mathbf{x},\mathbf{y})u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}. \end{aligned}$$

We will prove the theorem by upper bound $T_t u_{t,n} - T_t^1 u_{t,n}$, $T_t^1 u_{t,n} - T_t^2 u_{t,n}$ and $T_t^2 u_{t,n} - T_{t,n}u_{t,n}$ separately.

First, let us see $T_t u_{t,n} - T_t^1 u_{t,n}$.

$$\begin{aligned} &\left|T_t u_{t,n} - T_t^1 u_{t,n}\right| \\ \le&\left|\frac{1}{w_{t,n}(\mathbf{x})} - \frac{1}{w_t(\mathbf{x})}\right|\left(\left|\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})T_t u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}\right| + t\left|\int_{\mathcal{M}} \bar{R}(\mathbf{x},\mathbf{y})u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}\right|\right) \\ \le&\frac{2C_t}{w_{\min}^2}\sup_{g\in\mathcal{R}_t}(|p_n(g) - p(g)|)\left(\left|\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})T_t u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}\right| + t\left|\int_{\mathcal{M}} \bar{R}(\mathbf{x},\mathbf{y})u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}\right|\right) \\ \le&\frac{C}{t^{3k/4}}(\|T_t u_{t,n}\|_{L^2} + t\|u_{t,n}\|_{L^2})\sup_{g\in\mathcal{R}_t}(|p_n(g) - p(g)|) \\ \le&\frac{C}{t^{3k/4}}\|u_{t,n}\|_{L^2}\sup_{g\in\mathcal{R}_t}(|p_n(g) - p(g)|) \\ \le&\frac{C}{t^{3k/4}}\|f\|_\infty\sup_{g\in\mathcal{R}_t}(|p_n(g) - p(g)|), \end{aligned}$$

Similarly, we have

$$\left|\nabla(T_t u_{t,n} - T_t^1 u_{t,n})\right| \le \frac{C}{t^{(3k+2)/4}}\|f\|_\infty\sup_{g\in\mathcal{R}_t\cup\mathcal{D}_t}(|p_n(g) - p(g)|),$$

which proves that

$$\left\|T_t u_{t,n} - T_t^1 u_{t,n}\right\|_{C^1} \leq \frac{C}{t^{(3k+2)/4}} \|f\|_\infty \sup_{g \in \mathcal{R}_t \cup \mathcal{D}_t} (|p_n(g) - p(g)|). \tag{5.30}$$

Secondly, using Theorem 5.1 we have

$$\begin{aligned}
&\left|T_t^1 u_{t,n} - T_t^2 u_{t,n}\right| \\
=\ &\left|\frac{1}{w_{t,n}(\mathbf{x})} \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y}) \left(T_t u_{t,n}(\mathbf{y}) - T_{t,n} u_{t,n}(\mathbf{y})\right) p(\mathbf{y}) \mathrm{d}\mathbf{y}\right| \\
\leq\ & C t^{-k/4} \left\|T_t u_{t,n} - T_{t,n} u_{t,n}\right\|_{L^2} \\
=\ & C t^{-k/4} \left\|(T_t - T_{t,n}) T_{t,n} f\right\|_{L^2} \\
\leq\ & \frac{C}{t^{3k/4+1}} \|f\|_\infty \left(\sup_{g \in \mathcal{R}_t \cup \mathcal{R}_t \cdot \mathcal{K}_{t,n}} |p_n(g) - p(g)| + t \sup_{g \in \mathcal{K}_{t,n} \cdot \overline{\mathcal{R}}_t} |p_n(g) - p(g)| + t^2 \sup_{g \in \overline{\mathcal{K}}_{t,n} \cdot \overline{\mathcal{R}}_t} |p_n(g) - p(g)| \right).
\end{aligned}$$

and

$$\begin{aligned}
&\left|\nabla \left(T_t^1 u_{t,n} - T_t^2 u_{t,n}\right)\right| \\
=\ &\left|\nabla_{\mathbf{x}} \left(\frac{1}{w_{t,n}(\mathbf{x})} \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y}) \left(T_t u_{t,n}(\mathbf{y}) - T_{t,n} u_{t,n}(\mathbf{y})\right) p(\mathbf{y}) \mathrm{d}\mathbf{y}\right)\right| \\
\leq\ & C t^{-k/4+1/2} \left\|T_t u_{t,n} - T_{t,n} u_{t,n}\right\|_{L^2} \\
=\ & C t^{-k/4+1/2} \left\|(T_t - T_{t,n}) T_{t,n} f\right\|_{L^2} \\
\leq\ & \frac{C}{t^{k/4+3/2}} \|f\|_\infty \left(\sup_{g \in \mathcal{R}_t \cup \mathcal{R}_t \cdot \mathcal{K}_{t,n}} |p_n(g) - p(g)| + t \sup_{g \in \mathcal{K}_{t,n} \cdot \overline{\mathcal{R}}_t} |p_n(g) - p(g)| + t^2 \sup_{g \in \overline{\mathcal{K}}_{t,n} \cdot \overline{\mathcal{R}}_t} |p_n(g) - p(g)| \right).
\end{aligned}$$

This implies that

$$\begin{aligned}
&\left\|T_t^1 u_{t,n} - T_t^2 u_{t,n}\right\|_{C^1} \\
\leq\ & \frac{C}{t^{k/4+3/2}} \|f\|_\infty \left(\sup_{g \in \mathcal{R}_t \cup \mathcal{R}_t \cdot \mathcal{K}_{t,n}} |p_n(g) - p(g)| + t \sup_{g \in \mathcal{K}_{t,n} \cdot \overline{\mathcal{R}}_t} |p_n(g) - p(g)| + t^2 \sup_{g \in \overline{\mathcal{K}}_{t,n} \cdot \overline{\mathcal{R}}_t} |p_n(g) - p(g)| \right).
\end{aligned} \tag{5.31}$$

Now, we turn to estimate $T_{t,n} u_{t,n} - T_t^2 u_{t,n}$. Using (5.28), we have

$$\begin{aligned}
T_{t,n} u_{t,n} - T_t^2 u_{t,n} \ =\ & \frac{1}{w_{t,n}(\mathbf{x})} \left(\frac{1}{n} \sum_{i=1}^n R_t(\mathbf{x}, \mathbf{x}_i) v_i - \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y}) T_{t,n} u_{t,n}(\mathbf{y}) p(\mathbf{y}) \mathrm{d}\mathbf{y} \right) \\
& + \frac{t}{w_{t,n}(\mathbf{x})} \left(\frac{1}{n} \sum_{i=1}^n \bar{R}(\mathbf{x}, \mathbf{x}_i) u_i - \int_{\mathcal{M}} \bar{R}(\mathbf{x}, \mathbf{y}) u_{t,n}(\mathbf{y}) p(\mathbf{y}) \mathrm{d}\mathbf{y} \right).
\end{aligned}$$

Using (5.28) again, the first term becomes

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n} R_t(\mathbf{x},\mathbf{x}_i)v_i - \int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})T_{t,n}u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}\right| \\
\leq\ &\left|\frac{1}{n}\sum_{i=1}^{n} R_t(\mathbf{x},\mathbf{x}_i)\left(\frac{1}{nw_{t,n}(\mathbf{x}_i)}\sum_{j=1}^{n} R_t\left(\mathbf{x}_i,\mathbf{x}_j\right)v_j + \frac{t}{nw_{t,n}(\mathbf{x}_i)}\sum_{j=1}^{n}\bar{R}_t\left(\mathbf{x}_i-\mathbf{x}_j\right)u_j\right)\right. \\
&\left. - \int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})\left(\frac{1}{nw_{t,n}(\mathbf{y})}\sum_{j=1}^{n} R_t\left(\mathbf{y},\mathbf{x}_j\right)v_j + \frac{t}{nw_{t,n}(\mathbf{y})}\sum_{j=1}^{n}\bar{R}_t\left(\mathbf{y}-\mathbf{x}_j\right)u_j\right)p(\mathbf{y})\mathrm{d}\mathbf{y}\right| \\
\leq\ &\left|\frac{1}{n}\sum_{j=1}^{n} v_j\left(\frac{1}{n}\sum_{i=1}^{n}\frac{R_t(\mathbf{x},\mathbf{x}_i)}{w_{t,n}(\mathbf{x}_i)}R_t\left(\mathbf{x}_i,\mathbf{x}_j\right) - \int_{\mathcal{M}}\frac{R_t(\mathbf{x},\mathbf{y})}{w_{t,n}(\mathbf{y})}R_t\left(\mathbf{y},\mathbf{x}_j\right)p(\mathbf{y})\mathrm{d}\mathbf{y}\right)\right| \\
&+\left|\frac{t}{n}\sum_{j=1}^{n} u_j\left(\frac{1}{n}\sum_{i=1}^{n}\frac{R_t(\mathbf{x},\mathbf{x}_i)}{w_{t,n}(\mathbf{x}_i)}\bar{R}_t\left(\mathbf{x}_i,\mathbf{x}_j\right) - \int_{\mathcal{M}}\frac{R_t(\mathbf{x},\mathbf{y})}{w_{t,n}(\mathbf{y})}\bar{R}_t\left(\mathbf{y},\mathbf{x}_j\right)p(\mathbf{y})\mathrm{d}\mathbf{y}\right)\right|
\end{aligned}$$

Using the similar techniques from (5.15) to (5.21), we get

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n} R_t(\mathbf{x},\mathbf{x}_i)v_i - \int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})T_{t,n}u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}\right| \\
\leq\ &\frac{C}{t^{k/4}}\left(\frac{1}{n}\sum_{j=1}^{n} v_j^2\right)^{1/2} C_t \sup_{g\in\mathcal{K}_{t,n}\cdot\mathcal{R}_t}|p_n(g)-p(g)| + \frac{C}{t^{k/4-1}}\left(\frac{1}{n}\sum_{j=1}^{n} u_j^2\right)^{1/2} C_t \sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| \\
\leq\ &\frac{C}{t^{3k/4}}\|f\|_\infty\left(\sup_{g\in\mathcal{K}_{t,n}\cdot\mathcal{R}_t}|p_n(g)-p(g)| + t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)|\right)
\end{aligned}$$

The second term can be bounded similarly,

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n} \bar{R}(\mathbf{x},\mathbf{x}_i)u_i - \int_{\mathcal{M}} \bar{R}(\mathbf{x},\mathbf{y})u_{t,n}(\mathbf{y})p(\mathbf{y})\mathrm{d}\mathbf{y}\right| \\
\leq\ &\frac{C}{t^{k/4}}\left(\frac{1}{n}\sum_{j=1}^{n} u_j^2\right)^{1/2} C_t \sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| + \frac{C}{t^{k/4-1}}\left(\frac{1}{n}\sum_{j=1}^{n} f_j^2\right)^{1/2} C_t \sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| \\
\leq\ &\frac{C}{t^{3k/4}}\|f\|_\infty\left(\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| + t\sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)|\right) \qquad (5.32)
\end{aligned}$$

Now, we have

$$|T_{t,n}u_{t,n} - T_t^2 u_{t,n}| \leq \frac{C}{t^{3k/4}}\|f\|_\infty\left(\sup_{g\in\mathcal{K}_{t,n}\cdot\mathcal{R}_t}|p_n(g)-p(g)| + t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)| + t^2\sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{R}}_t}|p_n(g)-p(g)|\right)$$

Using the similar method, we can get

$$|\nabla(T_{t,n}u_{t,n} - T_t^2 u_{t,n})| \leq \frac{C}{t^{3k/4+1/2}}\|f\|_\infty\left(\sup_{g\in\mathcal{K}_{t,n}\cdot\mathcal{D}_t}|p_n(g)-p(g)| + t\sup_{g\in\mathcal{K}_{t,n}\cdot\overline{\mathcal{D}}_t}|p_n(g)-p(g)| + t^2\sup_{g\in\overline{\mathcal{K}}_{t,n}\cdot\overline{\mathcal{D}}_t}|p_n(g)-p(g)|\right)$$

The estimate of $\|(T_t - T_{t,n})T_{t,n}\|_{C^1}$ in Theorem 4.5 is proved.

Similarly, we can obtain the estimate of $\|(T_t - T_{t,n})f\|_{C^1}$ for any $f \in C(\mathcal{M})$ which complete the proof.

□

6 Conclusions

In this paper, we proved that the spectra of the normalized graph laplacian (1.1) will converge to the spectral of Laplace-Beltrami operator with Neumann boundary condition (1.2) as $t \to 0$ and the number of sample points goes to infinity. The samples points are assumed to be drawn on a smooth manifold according to some probability distribution p. Moreover, we also give an estimate of the convergence rate. However, the estimate of the convergence rate in this paper is far from optimal. In the analysis, we believe that *a prior* estimate of the integral equation (1.3) can be improved. Now, we only get L^2 estimate. In the spectra convergence analysis, we need C^1 estimate. In this paper, the regularity is lifted by using the regularity of the kernel function. The trade off is that a large factor $t^{-k/4}$ emerges which reduces the rate of convergence.

Appendix A: Proof of Theorem 5.4

Proposition A.1. *([20]) Assume both $\mathcal{M}$ and $\partial\mathcal{M}$ are C^2 smooth. There are constants $w_{\min} > 0, w_{\max} < +\infty$ and $T_0 > 0$ depending only on the geometry of $\mathcal{M}$, so that*

$$w_{\min} \leq w_t(\mathbf{x}) = \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y}) \mathrm{d}\mathbf{y} \leq w_{\max}$$

as long as $t < T_0$.

We have the following lemma about the function $w_{t,n}$.

Lemma A.1. *Under the assumptions in Assumption 1, if $C_t \sup_{f \in \mathcal{R}_t} |p(f) - p_n(f)| \leq w_{\min}/2$,*

$$w_{\min}/2 \leq w_{t,n}(\mathbf{x}) \leq w_{\max} + w_{\min}/2.$$

This lemma is a direct consequence of Proposition A.1 and the fact that

$$\left| w_{t,n}(\mathbf{x}) - C_t \int_{\mathcal{M}} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) p(\mathbf{y}) \mathrm{d}\mathbf{y} \right| \leq C_t \sup_{f \in \mathcal{R}_t} |p(f) - p_n(f)|.$$

Lemma A.2. *([20, 14]) For any function $u \in L^2(\mathcal{M})$, there exists a constant $C > 0$ only depends on $\mathcal{M}$, such that*

$$\int_{\mathcal{M}} \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y})(u(\mathbf{x}) - u(\mathbf{y}))^2 p(\mathbf{x}) p(\mathbf{y}) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y} \geq C \int_{\mathcal{M}} |u(\mathbf{x}) - \bar{u}|^2 p(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{A.1}$$

where

$$\bar{u} = \int_{\mathcal{M}} u(\mathbf{x}) p(\mathbf{x}) \mathrm{d}\mathbf{x}.$$

Now, we can prove Theorem 5.4.

Proof. of Theorem 5.4

First, we introduce a smooth function u that approximates $\mathbf{u}$ at the samples X_n.

$$u(\mathbf{x}) = \frac{C_t}{n w_{t',n}(\mathbf{x})} \sum_{i=1}^{n} R\left(\frac{|\mathbf{x} - \mathbf{x}_i|^2}{4t'}\right) u_i, \tag{A.2}$$

where $w_{t',n}(\mathbf{x}) = \frac{C_t}{n} \sum_{i=1}^{n} R\left(\frac{|\mathbf{x}-\mathbf{x}_i|^2}{4t'}\right)$ and $t' = t/18$.

Then, we have

$$
\begin{aligned}
&\int_{\mathcal{M}}\int_{\mathcal{M}} R_{t'}(\mathbf{x},\mathbf{y})\left(u(\mathbf{x})-u(\mathbf{y})\right)^2 p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\\
=&\int_{\mathcal{M}}\int_{\mathcal{M}} R_{t'}(\mathbf{x},\mathbf{y})\left(\frac{1}{nw_{t',n}(\mathbf{x})}\sum_{i=1}^{n}R_{t'}(\mathbf{x},\mathbf{x}_i)u_i-\frac{1}{nw_{t',n}(\mathbf{y})}\sum_{j=1}^{n}R_{t'}(\mathbf{x}_j,\mathbf{y})u_j\right)^2 p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\\
=&\int_{\mathcal{M}}\int_{\mathcal{M}} R_{t'}(\mathbf{x},\mathbf{y})\left(\frac{1}{n^2w_{t',n}(\mathbf{x})w_{t',n}(\mathbf{y})}\sum_{i,j=1}^{n}R_{t'}(\mathbf{x},\mathbf{x}_i)R_{t'}(\mathbf{x}_j,\mathbf{y})(u_i-u_j)\right)^2 p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\\
\leq&\int_{\mathcal{M}}\int_{\mathcal{M}} R_{t'}(\mathbf{x},\mathbf{y})\frac{1}{n^2w_{t',n}(\mathbf{x})w_{t',n}(\mathbf{y})}\sum_{i,j=1}^{n}R_{t'}(\mathbf{x},\mathbf{x}_i)R_{t'}(\mathbf{x}_j,\mathbf{y})(u_i-u_j)^2 p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\\
=&\frac{1}{n^2}\sum_{i,j=1}^{n}\left(\int_{\mathcal{M}}\int_{\mathcal{M}}\frac{1}{w_{t',n}(\mathbf{x})w_{t',n}(\mathbf{y})}R_{t'}(\mathbf{x},\mathbf{x}_i)R_{t'}(\mathbf{x}_j,\mathbf{y})R_{t'}(\mathbf{x},\mathbf{y})p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\right)(u_i-u_j)^2.
\end{aligned}
\tag{A.3}
$$

Denote

$$
A=\int_{\mathcal{M}}\int_{\mathcal{M}}\frac{1}{w_{t',n}(\mathbf{x})w_{t',n}(\mathbf{y})}R_{t'}(\mathbf{x},\mathbf{x}_i)R_{t'}(\mathbf{x}_j,\mathbf{y})R_{t'}(\mathbf{x},\mathbf{y})p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}
$$

and then notice only when $|\mathbf{x}_i-\mathbf{x}_j|^2\leq 36t'$ is $A\neq 0$. For $|\mathbf{x}_i-\mathbf{x}_j|^2\leq 36t'$, we have

$$
\begin{aligned}
A\ \leq&\int_{\mathcal{M}}\int_{\mathcal{M}} R_{t'}(\mathbf{x},\mathbf{x}_i)R_{t'}(\mathbf{x}_j,\mathbf{y})R_{t'}(\mathbf{x},\mathbf{y})R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{72t'}\right)^{-1}R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{72t'}\right)p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\\
\leq&\frac{CC_t}{\delta_0}\int_{\mathcal{M}}\int_{\mathcal{M}} R_{t'}(\mathbf{x},\mathbf{x}_i)R_{t'}(\mathbf{x}_j,\mathbf{y})R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{72t'}\right)p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\\
\leq&CC_t\int_{\mathcal{M}}\int_{\mathcal{M}} R_{t'}(\mathbf{x},\mathbf{x}_i)R_{t'}(\mathbf{x}_j,\mathbf{y})R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{72t'}\right)p(\mathbf{x})p(\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\\
\leq&CC_tR\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t}\right).
\end{aligned}
\tag{A.4}
$$

Combining Equation (A.3), (A.4) and Lemma A.2, we obtain

$$
\frac{CC_t}{n^2t}\sum_{i,j=1}^{n}R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t}\right)(u_i-u_j)^2\geq\int_{\mathcal{M}}(u(\mathbf{x})-\bar{u})^2p(\mathbf{x})\mathrm{d}\mathbf{x}
\tag{A.5}
$$

We now lower bound the RHS of the above equation using $\frac{1}{n}\sum_{j=1}^{n}u_i^2$.

$$
|\bar{u}|\ =\ \left|\int_{\mathcal{M}}u(\mathbf{x})p(\mathbf{x})\mathrm{d}\mathbf{x}\right|=\left|\frac{1}{n}\sum_{j=1}^{n}\left(u_j\int_{\mathcal{M}}\frac{C_t}{w_{t',n}(\mathbf{x})}R\left(\frac{|\mathbf{x}-\mathbf{x}_j|^2}{4t'}\right)p(\mathbf{x})\mathrm{d}\mathbf{x}\right)\right|.
\tag{A.6}
$$

Notice that

$$
\left|\int_{\mathcal{M}}\frac{C_t}{w_{t',n}(\mathbf{x})}R\left(\frac{|\mathbf{x}-\mathbf{x}_j|^2}{4t'}\right)p(\mathbf{x})\mathrm{d}\mathbf{x}-\frac{1}{n}\sum_{i=1}^{n}\frac{C_t}{w_{t',n}(\mathbf{x}_i)}R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t'}\right)\right|\leq C_t\sup_{f\in\mathcal{K}_{t',n}}|p(f)-p_n(f)|.
$$

Thus we have

$$
\begin{aligned}
|\bar{u}| &\leq \left| \frac{1}{n^2} \sum_{i,j=1}^{n} \frac{C_t}{w_{t',n}(\mathbf{x}_i)} R\left(\frac{|\mathbf{x}_i - \mathbf{x}_j|^2}{4t'} \right) u_j \right| + \left(\frac{1}{n} \sum_{j=1}^{n} |u_j| \right) \sup_{f \in \mathcal{K}_{t',n}} |p(f) - p_n(f)| \\
&\leq \left| \frac{1}{n} \sum_{i=1}^{n} u(\mathbf{x}_i) \right| + \left(\frac{1}{n} \sum_{j=1}^{n} |u_j| \right) \sup_{f \in \mathcal{K}_{t',n}} |p(f) - p_n(f)| \\
&\leq \left| \frac{1}{n^2} \sum_{i,j=1}^{n} \frac{C_t}{w_{t',n}(\mathbf{x}_i)} R\left(\frac{|\mathbf{x}_i - \mathbf{x}_j|^2}{4t'} \right) (u_j - u_i) \right| + \left(\frac{1}{n} \sum_{j=1}^{n} u_j^2 \right)^{1/2} \sup_{f \in \mathcal{K}_{t',n}} |p(f) - p_n(f)| \\
&\leq \frac{2}{w_{\min}} \left(\frac{C_t}{n^2} \sum_{i,j=1}^{n} R\left(\frac{|\mathbf{x}_i - \mathbf{x}_j|^2}{4t'} \right) (u_i - u_j)^2 \right)^{1/2} + \left(\frac{1}{n} \sum_{j=1}^{n} u_j^2 \right)^{1/2} \sup_{f \in \mathcal{K}_{t',n}} |p(f) - p_n(f)|,
\end{aligned}
\tag{A.7}
$$

Denote

$$
\begin{aligned}
A = \int_{\mathcal{M}} \frac{C_t}{w_{t',n}^2(\mathbf{x})} R\left(\frac{|\mathbf{x} - \mathbf{x}_i|^2}{4t'} \right) R\left(\frac{|\mathbf{x} - \mathbf{x}_l|^2}{4t'} \right) p(\mathbf{x}) \mathrm{d}\mathbf{x} - \\
\frac{1}{n} \sum_{j=1}^{n} \frac{C_t}{w_{t',n}^2(\mathbf{x}_j)} R\left(\frac{|\mathbf{x}_j - \mathbf{x}_i|^2}{4t'} \right) R\left(\frac{|\mathbf{x}_j - \mathbf{x}_l|^2}{4t'} \right)
\end{aligned}
$$

and then $|A| \leq C_t \sup_{f \in \mathcal{K}_{t',n} \cdot \mathcal{K}_{t',n}} |p(f) - p_n(f)|$. At the same time, notice that only when $|\mathbf{x}_i - \mathbf{x}_l|^2 < 16t'$ is $A \neq 0$. Thus we have

$$
|A| \leq \frac{1}{\delta_0} |A| R(\frac{|\mathbf{x}_i - \mathbf{x}_l|^2}{72t'}).
$$

Then

$$
\begin{aligned}
&\left| \int_{\mathcal{M}} u^2(\mathbf{x}) \mathrm{d}\mathbf{x} - \frac{1}{n} \sum_{j=1}^{n} u^2(\mathbf{x}_j) \right| \\
\leq\ & \frac{1}{n^2} \sum_{i,l=1}^{n} |C_t u_i u_l| |A| \\
\leq\ & \frac{C_t}{n^2} \sup_{f \in \mathcal{K}_{t',n} \cdot \mathcal{K}_{t',n}} |p(f) - p_n(f)| \sum_{i,l=1}^{n} \left| C_t R\left(\frac{|\mathbf{x}_i - \mathbf{x}_l|^2}{72t'} \right) u_i u_l \right| \\
\leq\ & \frac{C_t}{n^2} \sup_{f \in \mathcal{K}_{t',n} \cdot \mathcal{K}_{t',n}} |p(f) - p_n(f)| \sum_{i,l=1}^{n} C_t R\left(\frac{|\mathbf{x}_i - \mathbf{x}_l|^2}{72t'} \right) u_i^2 \\
\leq\ & (w_{\max} + w_{\min}/2) C_t \sup_{f \in \mathcal{K}_{t',n} \cdot \mathcal{K}_{t',n}} |p(f) - p_n(f)| \left(\frac{1}{n} \sum_{i=1}^{n} u_i^2 \right).
\end{aligned}
\tag{A.8}
$$

In the last inequality, we use the condition that $C_t \sup_{f \in \mathcal{R}_t} |p(f) - p_n(f)| \leq w_{\min}/2$.

Now combining Equation (A.5), (A.7) and (A.8), we have for small t

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n} u^2(\mathbf{x}_i) &= \int_{\mathcal{M}} u^2(\mathbf{x})p(\mathbf{x})\mathrm{d}\mathbf{x} + (w_{\max}+w_{\min}/2)C_t \sup_{f\in\mathcal{K}_{t',n}\cdot\mathcal{K}_{t',n}} |p(f)-p_n(f)| \left(\frac{1}{n}\sum_{i=1}^{n} u_i^2\right) \\
&\le 2\int_{\mathcal{M}} (u(\mathbf{x})-\bar{u})^2 p(\mathbf{x})\mathrm{d}\mathbf{x} + 2\bar{u}^2 + (w_{\max}+w_{\min}/2)C_t \sup_{f\in\mathcal{K}_{t',n}\cdot\mathcal{K}_{t',n}} |p(f)-p_n(f)| \left(\frac{1}{n}\sum_{i=1}^{n} u_i^2\right) \\
&\le \frac{CC_t}{n^2 t}\sum_{i,j=1}^{n} R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t}\right)(u_i-u_j)^2 \\
&\quad + \max\{w_{\max}+w_{\min}/2, 2/w_{\min}\}C_t \sup_{f\in\mathcal{K}_{t',n}\cdot\mathcal{K}_{t',n}\cup\mathcal{K}_{t',n}} |p(f)-p_n(f)| \left(\frac{1}{n}\sum_{i=1}^{n} u_i^2\right).
\end{aligned}$$

Let $\delta = \frac{w_{\min}}{4w_{\max}+3w_{\min}}$. If $\frac{1}{n}\sum_{i=1}^{n} u^2(\mathbf{x}_i) \ge \frac{\delta^2}{n}\sum_{i=1}^{n} u_i^2$, and

$$\max\{w_{\max}+w_{\min}/2, 2/w_{\min}\}C_t \sup_{f\in\mathcal{K}_{t',n}\cdot\mathcal{K}_{t',n}\cup\mathcal{K}_{t',n}} |p(f)-p_n(f)| \le \delta^2/2$$

then we have completed the proof. Otherwise, we have

$$\frac{1}{n}\sum_{i=1}^{n}(u_i-u(\mathbf{x}_i))^2 = \frac{1}{n}\sum_{i=1}^{n} u_i^2 + \frac{1}{n}\sum_{i=1}^{n} u(\mathbf{x}_i)^2 - \frac{2}{n}\sum_{i=1}^{n} u_i u(\mathbf{x}_i) \ge \frac{(1-\delta)^2}{n}\sum_{i=1}^{n} u_i^2. \tag{A.9}$$

This enables us to prove the theorem in the case of $\frac{1}{n}\sum_{i=1}^{n} u^2(\mathbf{x}_i) < \frac{\delta^2}{n}\sum_{i=1}^{n} u_i^2$ as follows.

$$\begin{aligned}
&\frac{C_t}{n^2}\sum_{i,j=1}^{n} R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t'}\right)(u_i-u_j)^2 \\
=\ &\frac{2C_t}{n^2}\sum_{i,j=1}^{n} R\left(\frac{|\mathbf{x}_i-\mathbf{x}_j|^2}{4t'}\right)u_i(u_i-u_j) \\
=\ &\frac{2}{n}\sum_{i=1}^{n} u_i(u_i-u(\mathbf{x}_i))w_{t',n}(\mathbf{x}_i) \\
=\ &\frac{2}{n}\sum_{i=1}^{n}(u_i-u(\mathbf{x}_i))^2 w_{t',n}(\mathbf{x}_i) + \frac{2}{n}\sum_{i=1}^{n} u(\mathbf{x}_i)(u_i-u(\mathbf{x}_i))w_{t',n}(\mathbf{x}_i) \\
\ge\ &\frac{2}{n}\sum_{i=1}^{n}(u_i-u(\mathbf{x}_i))^2 w_{t',n}(\mathbf{x}_i) - 2\left(\frac{1}{n}\sum_{i=1}^{n} u^2(\mathbf{x}_i)w_{t',n}(\mathbf{x}_i)\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}(u_i-u(\mathbf{x}_i))^2 w_{t,n}(\mathbf{x}_i)\right)^{1/2} \\
\ge\ &\frac{w_{\min}}{n}\sum_{i=1}^{n}(u_i-u(\mathbf{x}_i))^2 - 2(w_{\max}+w_{\min}/2)\left(\frac{1}{n}\sum_{i=1}^{n} u^2(\mathbf{x}_i)\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}(u_i-u(\mathbf{x}_i))^2\right)^{1/2} \\
\ge\ &(w_{\min}(1-\delta) - 2(w_{\max}+w_{\min}/2)\delta)\left(\frac{1}{n}\sum_{i=1}^{n} u_i^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}(u_i-u(\mathbf{x}_i))^2\right)^{1/2} \\
\ge\ &w_{\min}(1-\delta)^2\left(\frac{1}{n}\sum_{i=1}^{n} u_i^2\right).
\end{aligned} \tag{A.10}$$

□

References

[1] K. Atkinson. The numerical solution of the eigenvalue problem for compact integral operators. *Transactions of the American Mathematical Society*, 129(3):458–465, 1967.

[2] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396, 2003.

[3] M. Belkin and P. Niyogi. Towards a theoretical foundation for laplacian-based manifold methods. In *COLT*, pages 486–500, 2005.

[4] M. Belkin and P. Niyogi. Convergence of laplacian eigenmaps. In *Adv. Neur. In.: Proceedings of the 2006 Conference*, volume 19, page 129, 2007.

[5] M. Belkin, Q. Que, Y. Wang, and X. Zhou. Toward understanding complex spaces: Graph laplacians on manifolds with singularities and boundaries. In *25th Annual Conference on Learning Theory*, volume 23, pages 36.1–36.24, 2012.

[6] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, F. Warner, and S. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. In *Proceedings of the National Academy of Sciences*, pages 7426–7431, 2005.

[7] E. Giné and V. Koltchinskii. Empirical graph laplacian approximation of laplace-beltrami operaetors: Large sample results. *High Dimensional Probability*, 51:238–259, 2006.

[8] M. Hein. Uniform convergence of adaptive graph-based regularization. In *Proceedings of the 19th Annual Conference on Learning Theory*, COLT'06, pages 50–64, Berlin, Heidelberg, 2006. Springer-Verlag.

[9] M. Hein, J.-Y. Audibert, and U. von Luxburg. From graphs to manifolds – weak and strong pointwise consistency of graph laplacians. In *Proceedings of the 18th Annual Conference on Learning Theory*, COLT'05, pages 470–485, Berlin, Heidelberg, 2005. Springer-Verlag.

[10] M. Hein, J.-Y. Audibert, and U. von Luxburg. On the convergence of eigenspaces in kernel principal component analysis. In *Advances in Neural Information Processing Systems*, pages 1649–1656, Cambridge, MA, 2006. MIT Press.

[11] V. Koltchinskii. Asymptotics of spectral projections of some random matrices approximating integral operators. *Progress in Probabilty*, 43:191–227, 1998.

[12] V. Koltchinskii and E. Giné. Random matrix approximation of spectra of integral operators. *Bernoulli*, 6:113–167, 2000.

[13] S. Lafon. *Diffusion Maps and Geodesic Harmonics*. PhD thesis, 2004.

[14] Z. Li and Z. Shi. A convergent point integral method for isotropic elliptic equations on point cloud. *SIAM: Multiscale Modeling & Simulation*, 14:874–905, 2016.

[15] Z. Li, Z. Shi, and J. Sun. Point integral method for solving poisson-type equations on manifolds from point clouds with convergence guarantees. *Communications in Computational Physics*, 22:228–258, 2017.

[16] F. Mémoli. A spectral notion of Gromov-Wasserstein distances and related methods. *Applied and Computational Mathematics.*, 30:363–401, 2011.

[17] S. Mendelson. A few notes on statistical learning theory. In *Lecture Notes in Computer Science*, volume 2600, pages 1–40, 2003.

[18] M. Ovsjanikov, J. Sun, and L. J. Guibas. Global intrinsic symmetries of shapes. *Comput. Graph. Forum*, 27(5):1341–1348, 2008.

[19] J. Shawe-Taylor, C. Williams, N. Cristianini, and J. Kandola. On the eigenspectrum of the gram matrix and the generalization error of kernel-pca. *IEEE Trans. Inform. Theory*, 51:2510–2522, 2005.

[20] Z. Shi and J. Sun. Convergence of the point integral method for poisson equation on point cloud. *to appear in Research in the Mathematical Sciences*.

[21] A. Singer. From graph to manifold Laplacian: The convergence rate. *Applied and Computational Harmonic Analysis*, 21(1):128–134, July 2006.

[22] A. Singer and H. Wu. Spectral convergence of the connection laplacian from random samples. *arXiv:1306.1587*.

[23] N. G. Trillos and D. Slepčev. A variational approach to the consistency of spectral clustering. *arXiv:1508.01928*.

[24] U. von Luxburg, M. Belkin, and O. Bousquet. Consistency of spectral clustering. *Ann. Statist.*, 36(2):555–586, 2008.

[25] X. Wang. Spectral convergence rate of graph laplacian. *arXiv:1510.08110*.