HARMONIC EXTENSION ON POINT CLOUD *

ZUOQIANG SHI†, JIAN SUN ‡, AND MINGHAO TIAN§

Abstract. In this paper, we consider the harmonic extension problem, which is widely used in many applications of machine learning. We formulate the harmonic extension as solving a Laplace-Beltrami equation with Dirichlet boundary condition. We use the point integral method (PIM) proposed in [14, 18, 13] to solve the Laplace-Beltrami equation. The basic idea of the PIM method is to approximate the Laplace equation using an integral equation, which is easy to be discretized from points. Based on the integral equation, we found that traditional graph Laplacian method (GLM) may fail to approximate the harmonic functions in the classical sense. For the Laplace-Beltrami equation with Dirichlet boundary, we can prove the convergence of the point integral method. The point integral method is also very easy to implement, which only requires a minor modification of the graph Laplacian. One important application of the harmonic extension in machine learning is semi-supervised learning. We run a popular semi-supervised learning algorithm by Zhu et al. [23] over a couple of well-known datasets and compare the performance of the aforementioned approaches. Our experiments show the PIM performs the best. We also apply PIM to an image recovery problem and show it outperforms GLM. Finally, on a model problem of Laplace-Beltrami equation with Dirichlet boundary, we prove the convergence of the point integral method.

Keywords: harmonic extension; point cloud; point integral method; Laplace-Beltrami operator; Dirichlet boundary.

1. Introduction.

In this paper, we consider interpolation on a point cloud in high dimensional space. The problem is described as follows. Let $P = \{\mathbf{p}_1, \cdots, \mathbf{p}_n\}$ be a set of points in $\mathbb{R}^d$ and $S = \{\mathbf{s}_1, \cdots, \mathbf{s}_m\}$ be a subset of P. Let u be a function on the point set P and the value of u on $S \subset P$ is given as a function g over S, i.e. $u(\mathbf{s}) = g(\mathbf{s})$, $\forall \mathbf{s} \in S$. In this paper, S is called the labelled set. From the given value on S, we want to refer the value of u on the whole data set P. This is a fundamental mathematical model in many data analysis and machine learning problem.

This is an ill-posed problem. The function of u can be any value on $P \backslash S$, if we do not have any assumption on u. To make this problem well-posed, usually, we assume that the point cloud P sample a smooth manifold $\mathcal{M}$ embedded in $\mathbb{R}^d$ and u is a smooth function on $\mathcal{M}$. Based on this assumption, one idea is to find the smoothest u such that $u(\mathbf{s}) = g(\mathbf{s})$, $\forall \mathbf{s} \in S$. One of the simplest measurement of the smoothness of a function u is the L_2 norm of the gradient of u, which gives following objective function to minimize:

$$\mathcal{J}_{\mathcal{M}}(u) = \frac{1}{2}\int_{\mathcal{M}} \|\nabla_{\mathcal{M}} u(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}. \tag{1.1}$$

where $\mathcal{M}$ is the underlying manifold, P is a sample of $\mathcal{M}$, $\nabla_{\mathcal{M}}$ is the gradient on $\mathcal{M}$.

The classical harmonic extension problem, also known as the Dirichlet problem for Laplace equation, has been studied by mathematicians for more than a century and has many applications in mathematics. The discrete harmonicity has also been extensively studied in the graph theory [5]. For instance, it is closely related to random

*Research supported by NSFC Grant 11371220 and 11671005.

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

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

§Department of Mathematical Sciences, Tsinghua University, Beijing, China, 100084. *Email: tianmh0918@hotmail.com.*

walk and electric networks on graphs [6]. In machine learning, the discrete harmonic extension and its variants have been used for semi-supervised learning [23, 21].

One method which is widely used in many applications in image processing and data analysis is using the nonlocal gradient to discretize the gradient in (1.1), which gives following discrete objective function:

$$\mathcal{J}_P(u) = \frac{1}{2}\sum_{\mathbf{x},\mathbf{y}\in P} w(\mathbf{x},\mathbf{y})(u(\mathbf{x}) - u(\mathbf{y}))^2, \tag{1.2}$$

Here $w(\mathbf{x},\mathbf{y})$ is a given weight function. One often used weight is the Gaussian weight, $w(\mathbf{x},\mathbf{y}) = \exp(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{\sigma^2})$, σ is a parameter, $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^d$.

Based on the nonlocal gradient, the interpolation on point cloud is formulated as an optimization problem:

$$\min_u \quad \frac{1}{2}\sum_{\mathbf{x},\mathbf{y}\in P} w(\mathbf{x},\mathbf{y})(u(\mathbf{x}) - u(\mathbf{y}))^2, \tag{1.3}$$

with the constraint

$$u(\mathbf{x}) = g(\mathbf{x}), \quad \mathbf{x}\in S. \tag{1.4}$$

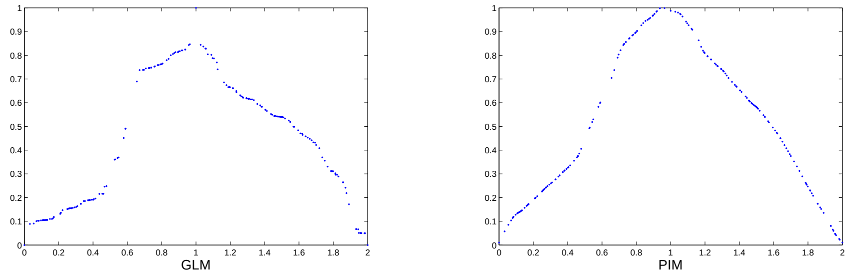
The optimal solution of above optimization problem is given by solving a linear system:

$$\begin{cases} \displaystyle\sum_{\mathbf{y}\in P}(w(\mathbf{x},\mathbf{y}) + w(\mathbf{y},\mathbf{x}))(u(\mathbf{x}) - u(\mathbf{y})) = 0, & \mathbf{x}\in P\backslash S, \\ u(\mathbf{x}) = g(\mathbf{x}), & \mathbf{x}\in S. \end{cases} \tag{1.5}$$

This is the well known graph Laplacian [5, 23] which has been used widely in many problems.

Much of research has been done on the theoretical analysis of the graph Laplacian. When the manifold has no boundary, the pointwise convergence of the graph Laplacian to the manifold Laplacian was shown in [1, 12, 11, 19], and the spectral convergence of the graph Laplacian was shown in [2]. When there are boundaries, Singer and Wu [20] and independently Shi and Sun [17] have shown that the spectra of the graph Laplacian converge to that of manifold Laplacian with Neumann boundary. However, for the graph Laplacian with Dirichlet type boundary, such as (1.5), the graph Laplacian approach has inconsisitent problem. In other word, the solution given by the graph Laplacian is not continuous on the labeled set. This inconsistency can be seen clearly in a simple example.

Let P be the union of 200 randomly sampled points over the interval $[0, 2]$ and $S = \{0, 1, 2\}$. Set $g = 0$ at $0, 2$ and $g = 1$ at 1. We run the above graph Laplacian method over this example. Figure 1 (a) shows the resulting minimizer. It is well-known that the harmonic function over the interval $(0, 2)$ with the Dirichlet boundary $\mathbf{g}$, in the classical sense, is a piece linear function, i.e., $u(x) = x$ for $x \in (0, 1)$ and $u(x) = 2 - x$ for $x \in (1, 2)$; Clearly, the function computed by GLM does not approximate the harmonic function in the classical sense. In particular, the Dirichlet boundary has not been enforced properly, and in fact the obtained function is not even continuous near the boundary.



(a) (b)

FIG. 1. *1D examples. (a): interpolation given by the graph Laplacian; (b): interpolation given by the point integral method. Note in (a), the recovered function is not continuous at the labelled set* $\{0,\ 1,\ 2\}$.

1.1. Harmonic Extension. In this paper, to derive a consistent method, we consider the harmonic extension in the continuous form, as shown in Figure 2. Assume $\mathcal{M}$ is a submanifold embedded in $\mathbb{R}^d$. Consider a function $u(\mathbf{x})$ defined on $\mathcal{M}$ and $u(\mathbf{x})$ is known in some regions $\Omega_1 \cup \cdots \cup \Omega_k \subset \mathcal{M}$. Now, we want to extend the function $u(\mathbf{x})$ from $\Omega_1 \cup \cdots \cup \Omega_k$ to the entire manifold $\mathcal{M}$.

$$\min_{u \in H^1(\mathcal{M})} \frac{1}{2} \int_{\mathcal{M}} \|\nabla_{\mathcal{M}} u(\mathbf{x})\|^2 \mathrm{d}\mathbf{x}. \tag{1.6}$$

with the constraint

$$u(\mathbf{x}) = g(\mathbf{x}), \quad \mathbf{x} \in \Omega_1 \cup \cdots \cup \Omega_k.$$

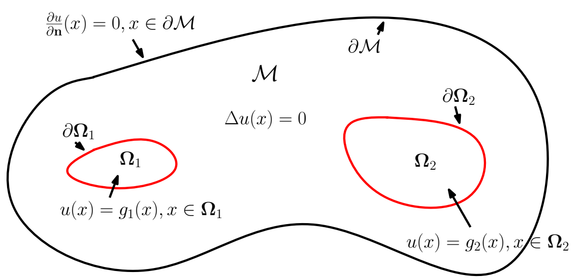


FIG. 2. *Sketch of the manifold.*

It is well known that above optimization problem (1.6) is solved by a harmonic extension problem:

$$\begin{cases} -\Delta_{\mathcal{M}} u(\mathbf{x}) = 0, & \mathbf{x} \in \mathcal{M}, \\ u(\mathbf{x}) = g(\mathbf{x}), & \mathbf{x} \in \partial\mathcal{M}_D, \\ \dfrac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) = 0, & \mathbf{x} \in \partial\mathcal{M}_N. \end{cases} \tag{1.7}$$

In the aforementioned harmonic extension problem, one can think of $\partial\mathcal{M}_D$ as the boundary of $\Omega_1 \cup \cdots \cup \Omega_k$, and $\partial\mathcal{M}_N$ as the actual boundary of $\mathcal{M}$.

In this paper, we use the point integral method (PIM) [14, 18, 13] to solve the harmonic extension problem (1.7). The key step in the point integral method is to use an integral equation to approximate the original Laplace-Beltrami equation (1.7) with small parameter $0 < \beta \ll 1$:

$$\frac{1}{t}\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x})-u(\mathbf{y}))\mathrm{d}\mathbf{y} - \frac{2}{\beta}\int_{\partial\mathcal{M}_D} \bar{R}_t(\mathbf{x},\mathbf{y})(g(\mathbf{y})-u(\mathbf{y}))\mathrm{d}\tau_{\mathbf{y}} = 0, \tag{1.8}$$

The kernel function R_t and $\bar{R}_t$ is defined as

$$R_t(\mathbf{x},\mathbf{y}) = R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right), \quad \bar{R}_t(\mathbf{x},\mathbf{y}) = \bar{R}\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) \tag{1.9}$$

where $R : \mathbb{R}^+ \to \mathbb{R}^+$ is a C^2 function which is integrable over $[0,+\infty)$, $t > 0$ is a parameter and

$$\bar{R}(r) = \int_r^{+\infty} R(s)\mathrm{d}s.$$

When $R(r) = e^{-r}$, $\bar{R}_t(\mathbf{x},\mathbf{y}) = R_t(\mathbf{x},\mathbf{y}) = \exp\left(-\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right)$ are the well-known Gaussian.

The integral equation (1.8) is discretized over the point cloud P and the labelled set S. We get a linear system as follows:

$$\sum_{\mathbf{y}\in P} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x})-u(\mathbf{y})) - \mu\sum_{\mathbf{y}\in S} \bar{R}_t(\mathbf{x},\mathbf{y})(g(\mathbf{y})-u(\mathbf{y})) = 0, \quad \mathbf{x}\in P \tag{1.10}$$

Here μ is a parameter associated with β whose choice will be described in Section 2.

Figure 1 (b) shows the interpolation computed by the point integral method in the simple 1D example. It is shown clearly that the solution given by the point integral method is continuous in the labeled set S while the graph Laplacian gives an discontinuous solution. Nevertheless, for the Laplace-Beltrami equation with Dirichlet boundary, we prove that the point integral method converges to the true solution as the number of sample points goes to infinity, the whole point cloud P sample the whole manifolds and the labeled set S sample the boundary. The result is given in Section 5.

One important application of the harmonic extension is semi-supervised learning [22]. We will perform the semi-supervised learning using the PIM over a couple of well-known data sets, and compare its performance to GLM as well as the closely related method by Zhou et al. [21]. The experimental results show that the PIM have the best performance. We also consider the image recovery problem and harmonic extension is used to recover the subsampled image based on the patch manifold.

The rest of the paper is organized as follows. The point integral method for the harmonic extension is given in Section 2. The examples of the semi-supervised learning and the image recovery are shown in Section 3 and Section 4 respectively. In Section 5, the convergence of the point integral method is proved for the Laplace-Beltrami equation with Dirichlet boundary.

2. Point Integral Method. The key observation in the point integral method is an integral approximation of the Laplace-Beltrami operator.

$$-\int_{\mathcal{M}} \nabla_{\mathcal{M}} u(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y} \approx \frac{1}{t}\int_{\mathcal{M}}(u(\mathbf{x})-u(\mathbf{y}))R_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y} - 2\int_{\partial\mathcal{M}} \frac{\partial u(\mathbf{y})}{\partial \mathbf{n}}\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}, \tag{2.1}$$

Next, we give a brief derivation of the integral approximation (2.1) in the Euclidean space $\mathbb{R}^d$. For a general submanifold, the rigorous analysis of the error in (2.1) can be found in Theorem 5.5. First, by integration by parts,

$$
\begin{aligned}
&\int_{\mathcal{M}} \Delta u(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y} \\
=&-\int_{\mathcal{M}} \nabla u(\mathbf{y})\cdot\nabla\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y}+\int_{\partial\mathcal{M}}\frac{\partial u(\mathbf{y})}{\partial \mathbf{n}}\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} \\
=&\frac{1}{2t}\int_{\mathcal{M}}(\mathbf{y}-\mathbf{x})\cdot\nabla u(\mathbf{y})R_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y}+\int_{\partial\mathcal{M}}\frac{\partial u(\mathbf{y})}{\partial \mathbf{n}}\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}.
\end{aligned}
\tag{2.2}
$$

The Taylor expansion of the function u gives that

$$
u(\mathbf{y})-u(\mathbf{x})=(\mathbf{y}-\mathbf{x})\cdot\nabla u(\mathbf{y})-\frac{1}{2}(\mathbf{y}-\mathbf{x})^T\mathbf{H}_u(\mathbf{y})(\mathbf{y}-\mathbf{x})+O(\|\mathbf{y}-\mathbf{x}\|^3),
\tag{2.3}
$$

where $\mathbf{H}_u(\mathbf{y})$ is the Hessian matrix of u at $\mathbf{y}$. The second order term is derived as follows under the assumption that u is smooth enough,

$$
\begin{aligned}
&\frac{1}{4t}\int_{\mathcal{M}}(\mathbf{y}-\mathbf{x})^T\mathbf{H}_u(\mathbf{y})(\mathbf{y}-\mathbf{x})R_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y} \\
=&\frac{1}{4t}\sum_{i,j=1}^{d}\int_{\mathcal{M}}(\mathbf{y}_i-\mathbf{x}_i)(\mathbf{y}_j-\mathbf{x}_j)\partial_{ij}u(\mathbf{y})R_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y} \\
=&-\frac{1}{2}\sum_{i,j=1}^{d}\int_{\mathcal{M}}(\mathbf{y}_i-\mathbf{x}_i)\partial_{ij}u(\mathbf{y})\partial_j\left(\bar{R}_t(\mathbf{x},\mathbf{y})\right)\mathrm{d}\mathbf{y} \\
=&\frac{1}{2}\sum_{i,j=1}^{d}\int_{\mathcal{M}}\partial_j(\mathbf{y}_i-\mathbf{x}_i)\partial_{ij}u(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y}+\frac{1}{2}\sum_{i,j=1}^{d}\int_{\mathcal{M}}(\mathbf{y}_i-\mathbf{x}_i)\partial_{ijj}u(\mathbf{y})\bar{R}_t(\mathbf{y},\mathbf{x})\mathrm{d}\mathbf{y} \\
&-\frac{1}{2}\sum_{i,j=1}^{d}\int_{\partial\mathcal{M}}(\mathbf{y}_i-\mathbf{x}_i)\mathbf{n}_j\partial_{ij}u(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} \\
=&\frac{1}{2}\int_{\mathcal{M}}\Delta u(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y}-\frac{1}{2}\sum_{i,j=1}^{d}\int_{\partial\mathcal{M}}(\mathbf{y}_i-\mathbf{x}_i)\mathbf{n}_j\partial_{ij}u(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}+O(t^{1/2}).
\end{aligned}
\tag{2.4}
$$

Then the integral approximation (2.1) is obtained following from the equations (2.2), (2.3) and (2.4).

Using the integral equation (2.1) and the boundary condition of (1.7), we know the Laplace equation with the mixed boundary condition can be approximated by the following integral equation:

$$
\frac{1}{t}\int_{\mathcal{M}}(u(\mathbf{x})-u(\mathbf{y}))R_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y}-2\int_{\partial\mathcal{M}_D}\frac{\partial u(\mathbf{y})}{\partial\mathbf{n}}\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}=0,
\tag{2.5}
$$

However, on $\partial\mathcal{M}_D$, $\frac{\partial u}{\partial \mathbf{n}}$ is not known. To address this issue, we use the Robin boundary condition to approximate the original Dirichlet boundary condition on $\partial\mathcal{M}_D$. Then, we consider the following Robin/Neumann mixed boundary problem.

$$
\begin{cases}
-\Delta_{\mathcal{M}} u(\mathbf{x}) &= 0, \quad \mathbf{x} \in \mathcal{M}, \\
u(\mathbf{x}) + \beta \dfrac{\partial u(\mathbf{x})}{\partial \mathbf{n}} &= g(\mathbf{x}), \quad \mathbf{x} \in \partial \mathcal{M}_D, \\
\dfrac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) &= 0, \quad \mathbf{x} \in \partial \mathcal{M}_N.
\end{cases} \tag{2.6}
$$

where $\beta > 0$ is a parameter. It is easy to prove that the solution of the above Robin/Neumann problem (2.6) converges to the solution of the Dirichlet/Neumann problem (1.7) as β goes to 0 (see Theorem 5.3).

By substituting the Robin boundary $\frac{\partial u(\mathbf{x})}{\partial \mathbf{n}} = \frac{1}{\beta}(g(\mathbf{x}) - u(\mathbf{x}))$ in the integral equation (2.5), we get an integral equation to solve the Robin/Neumann problem.

$$
\frac{1}{t}\int_{\mathcal{M}} (u(\mathbf{x}) - u(\mathbf{y}))w_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y} - \frac{2}{\beta}\int_{\partial\mathcal{M}_D} (g(\mathbf{y}) - u(\mathbf{y}))w_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} = 0, \tag{2.7}
$$

When $\beta > 0$ is small enough, this integral equation also gives a good approximation to the original harmonic extension problem (1.7).

Assume that the point set $P = \{\mathbf{p}_1, \cdots, \mathbf{p}_n\}$ samples the submanifold $\mathcal{M}$ and it is uniformly distributed, then the integral $\int_{\mathcal{M}} (u(\mathbf{x}) - u(\mathbf{y}))R_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y}$ is discretized and well approximated by

$$
\frac{|\mathcal{M}|}{n}\sum_{\mathbf{y}\in P} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x}) - u(\mathbf{y})),
$$

$\frac{|\mathcal{M}|}{n}$ is the volume weight of the point cloud P.

The boundary term $I_t u(\mathbf{x})$ is actually corresponding to the labeled set $S = \{\mathbf{s}_1, \cdots, \mathbf{s}_m\} \subset P$ where the values of function u are given. From the continuous point of view, for each point $\mathbf{s}_i \in S$, in a small area around it, the value of u is given. In this sense, each $\mathbf{s}_i$ actually stands for one part of the boundary $\partial\mathcal{M}_D$. Based on the above discussion, the boundary term $\int_{\partial\mathcal{M}_D} (g(\mathbf{y}) - u(\mathbf{y}))\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}$ can be discretized as

$$
\sum_{\mathbf{y}\in S} \bar{R}_t(\mathbf{x},\mathbf{y})(g(\mathbf{y}) - u(\mathbf{y}))
$$

up to the surface area weight $\frac{|\partial\mathcal{M}_D|}{m}$.

Therefore, the complete discretization is given by

$$
\sum_{\mathbf{y}\in P} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x}) - u(\mathbf{y})) - \mu\sum_{\mathbf{y}\in S} \bar{R}_t(\mathbf{x},\mathbf{y})(g(\mathbf{y}) - u(\mathbf{y})) = 0, \quad \mathbf{x} \in P \tag{2.8}
$$

where $\mu = \frac{2}{\beta}\frac{n|\partial\mathcal{M}_D|}{m|\mathcal{M}|}$. The parameter μ seems to be very complicated. However, in the computation of the harmonic extension, we give μ directly instead of using above formula of μ. One typical choice of μ is $|P|/|S|$, which is the inverse of the sample rate.

Remark 2.1. *Based on the discussion in this section, we can see clearly the reason that the traditional graph Laplacian may fail to approximate the classic harmonic functions. The reason is that in the graph Laplacian approach, the boundary term is dropped. However, this boundary term is not small. Without this term, the boundary condition may not be enforced correctly. This effect has been shown in Figure 1 and more evidence will be given in the example section.*

Algorithm 1 Semi-Supervised Learning

Require: A point set $P = \{\mathbf{p}_1, \cdots, \mathbf{p}_m, \mathbf{p}_{m+1}, \cdots, \mathbf{p}_n\} \subset \mathbb{R}^d$ and a partial label assignment $L : S = \{\mathbf{p}_1, \cdots, \mathbf{p}_m\} \to \{1, 2, \cdots, l\}$

Ensure: A complete label assignment $L : P \to \{1, 2, \cdots, l\}$

for $i = 1 : l$ **do**

  for $j = 1 : m$ **do**

    For any $\mathbf{p}_j \in S$, set $g_i(\mathbf{p}_j) = 1$ if $L(\mathbf{p}_j) = i$, and otherwise set $g_i(\mathbf{p}_j) = 0$.

  end for

  Compute the harmonic extension u_i of g_i by solving

$$\sum_{\mathbf{y}\in P} R_t(\mathbf{x}, \mathbf{y})(u_i(\mathbf{x}) - u_i(\mathbf{y})) - \mu \sum_{\mathbf{y}\in S} \bar{R}_t(\mathbf{x}, \mathbf{y})(g_i(\mathbf{y}) - u_i(\mathbf{y})) = 0, \quad \mathbf{x} \in P. \tag{3.1}$$

end for

for $j = m + 1 : n$ **do**

  $L(\mathbf{p}_j) = k$ where $k = \arg\max_{i\le l} u_i(\mathbf{p}_j)$.

end for

3. Semi-supervised Learning. In this section, we briefly describe the algorithm of semi-supervised learning based on the harmonic extension proposed by Zhu et al. [23]. We plug into the algorithm the aforementioned approach for harmonic extension, and apply them to several well-known data sets, and compare their performance.

Assume we are given a point set $P = \{\mathbf{p}_1, \cdots, \mathbf{p}_m, \mathbf{p}_{m+1}, \cdots, \mathbf{p}_n\} \subset \mathbb{R}^d$, and a label set $\{1, 2, \cdots, l\}$, and the label assignment on the first m points $L : \{\mathbf{p}_1, \cdots, \mathbf{p}_m\} \to \{1, 2, \cdots, l\}$. In a typical setting, m is much smaller than n. The purpose of the semi-supervised learning is to extend the label assignment L to the entire P, namely, infer the labels for the unlabeled points.

Think of the label points as the boundary $S = \{\mathbf{p}_1, \cdots, \mathbf{p}_m\}$. For the label $i \in \{1, 2, \cdots, l\}$, we set up the Dirichlet boundary g_i as follows. If a point $\mathbf{p}_j \in S$ is labeled as i, set $g_i(\mathbf{p}_j) = 1$, and otherwise set $g_i(\mathbf{p}_j) = 0$. Then we compute the harmonic extension u_i of g_i using the aforementioned approaches. In this way, we obtain a set of l harmonic functions $u_1, u_2, \cdots, u_l$. We label $\mathbf{p}_j$ using k where $k = \arg\max_{i\le l} u_i(\mathbf{p}_j)$. The algorithm is summarized in Algorithm 1. Note that this algorithm is slightly different from the original algorithm by Zhu et al. [23] where only one harmonic extension was computed by setting $g_i(\mathbf{p}_j) = k$ if $\mathbf{p}_j$ has a label k.

3.1. Experiments. We now apply the above semi-supervised learning algorithm to a couple of well-known data sets: MNIST and 20 Newsgroups. We do not claim the state of the art performance on these datasets. The purpose of these experiments is to compare the performance of different approaches of harmonic extension. We also compare to the closely related method of local and global consistency by Zhou et al. [21].

In the computations, the kernel function is chosen to be Gaussian, such that $\bar{R}_t(\mathbf{x}, \mathbf{y}) = R_t(\mathbf{x}, \mathbf{y})$ and

$$R_t(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{4t}\right).$$

The parameter t will be given later. In the computaitons, we set $\mu = 10^4$ in (3.1).

MNIST : In this experiment, we use the MNIST of dataset of handwritten digits [4], which contains $60k$ 28×28 gray scale digit images with labels. We view digits $0 \sim 9$ as ten classes. Each digit can be seen as a point in a common 784-dimensional Euclidean space. We randomly choose 16k images. Specifically, there are 1606, 1808, 1555, 1663, 1552, 1416, 1590, 1692, 1521 and 1597 digits in $0 \sim 9$ class respectively.

To set the parameter t, we build a graph by connecting a point x_i to its 10 nearest neighbors under the standard Euclidean distance. We compute the average of the distances for x_i to its neighbors on the graph, denoted h_i. Let h be the average of h_i's over all points and set $t = h^2$. The distance $|\mathbf{x}_i - \mathbf{x}_j|$ is computed as the graph distance between x_i and x_j. In the method of local and global consistency, we follow the paper [21] and set the width of the RBF kernel to be 0.3 and the parameter α in the iteration process to be 0.3.

For a particular trial, we choose k ($k = 1, 2, \cdots, 10$) images randomly from each class to assemble the labeled set B and assume all the other images are unlabeled. For each fixed k, we do 100 trials. The error bar of the tests is presented in Figure 3 (a). It is quite clear that the PIM has the best performance when there are more than 5 labeled points in each class, and the GLM has the worst performance.

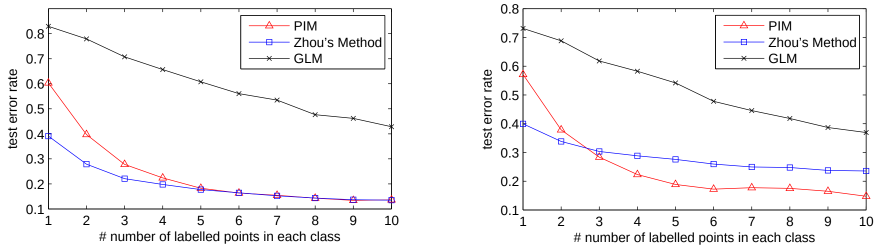
Newsgroup: In this experiment, we use the 20-newsgroups dataset, which is a classic dataset in text classification. We only choose the articles from topic *rec* containing four classes from the version 20-news-18828. We use Rainbow (version:20020213) to pre-process the dataset and finally vectorize them. The following command-line options are required[1]: (1)*--skip-header*: to avoid lexing headers; (2)*--use-stemming*: to modify lexed words with the 'Porter' stemmer; (3)*--use-stoplist*: to toss lexed words that appear in the SMART stoplist; (4)*--prune-vocab-by-doc-count=5*: to remove words that occur in 5 or fewer documents; Then, we use TF-IDF algorithm to normalize the word count matrix. Finally, we obtain 3970 documents (990 from rec.autos, 994 from rec.motorcycles, 994 from rec.sport.baseball and 999 from rec.sport.hockey) and a list of 8014 words. Each document will be treated as a point in a 8014-dimensional space.

To deal with text-kind data, we define a new distance introduced by Zhu et al. [23]: the distance between x_i and x_j is $d(x_i, x_j) = 1 - \cos\alpha$, where α is the angle between x_i and x_j in Euclidean space. Under this new distance, we ran the same experiment with the same parameter as we process the above MNIST dataset. The error bar of the tests for 20-newsgroups is presented in Figure 3 (b). A similar pattern result is observed, namely the PIM has the best performance when there are more than 2 labeled points in each class, and the GLM has the worst performance.

4. Image Recovery. In this example, we consider an image recovery problem. The original image is the well known image of Barbara (256×256) which is shown in Figure 4(a). Then, we subsample the image and only retain 1% of the pixels. The positions of the retained pixels are selected at random. The subsampled image is shown in Figure 4(b). Now, we want to recover the original image from the subsampled image. This is a classical problem in image processing which has been studied extensively. Here, we only use this example to demonstrate the difference between PIM method and the Graph Laplacian approach, rather than presenting an image recovery method.

First, we construct a point cloud from the original image, denoted by f, by using

[1]all the following options are offered by Rainbow



(a) (b)

FIG. 3. *(a) the error rates of digit recognition with a 16000-size subset of MNIST dataset; (b) the error rates of text classification with 20-newsgroups.rec(a 8014-dimensional space with 3970 data points).*

so called patch approach which is widely used in image processing [3, 10]. For each pixel x_i in the image f, we extract a patch around it of size 9×9 which is denoted as $p_{x_i}(f)$. Here $i=1,\cdots,256^2$. For the pixels on the boundary, the patch is obtained by extending the image symmetrically. Then, we can get 256^2 patches and each patch is 9×9. These patches consist of a point cloud in $\mathbb{R}^{81}$. Denote this point cloud as $P=\{p_{x_i}(f): i=1,\cdots,256^2\}$. And function u on P is defined as $u(p_{x_i}(f))=f(x_i)$, $f(x_i)$ is the value of image f at pixel x_i. Using this definition, at some patches which around the retained pixels, the value of u is known. The collection of these patches is denoted as S which is a subset of P.

Here, we recover the whole function u by harmonic extension, i.e. solving following linear system

$$\sum_{\mathbf{y}\in P} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x})-u(\mathbf{y}))-\mu\sum_{\mathbf{y}\in S}\bar{R}_t(\mathbf{x},\mathbf{y})(g(\mathbf{y})-u(\mathbf{y}))=0,\quad \mathbf{x}\in P. \tag{4.1}$$

We also compute the solution given by the graph Laplacian.

$$\left\{\begin{array}{rcll}\displaystyle\sum_{\mathbf{y}\in P} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x})-u(\mathbf{y})) &=& 0, & \mathbf{x}\in P\backslash S,\\ u(\mathbf{x}) &=& g(\mathbf{x}), & \mathbf{x}\in S.\end{array}\right. \tag{4.2}$$

We remark that in this example the point cloud P is constructed using the original image shown in Figure 4(a). So this is not an full image recovery method since the original image is used. However, by update the image iteratively, we can get a real image recovery method [15].

In the computations, we take the weight $R_t(\mathbf{x},\mathbf{y})$ as the Gaussian kernel. In this case, $\bar{R}_t(\mathbf{x},\mathbf{y})=R_t(\mathbf{x},\mathbf{y})$ and

$$R_t(\mathbf{x},\mathbf{y})=\exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{t(\mathbf{x})}\right).$$

Here, we choose t adaptive to the distribution of the point cloud. More specifically, $t(\mathbf{x})=\sigma(\mathbf{x})^2$ and $\sigma(\mathbf{x})$ is chosen to be the distance between $\mathbf{x}$ and its 20th nearest neighbor, To make the weight matrix sparse, the weight is truncated to the 50 nearest neighbors. The parameter μ in (4.1) is set to be $|P|/|S|$.

The solution of PIM (4.1) is given in Figure 4(c) and the solution of GLM (4.2) is given in Figure 4(d). Obviously, the result given by PIM is much better. To get

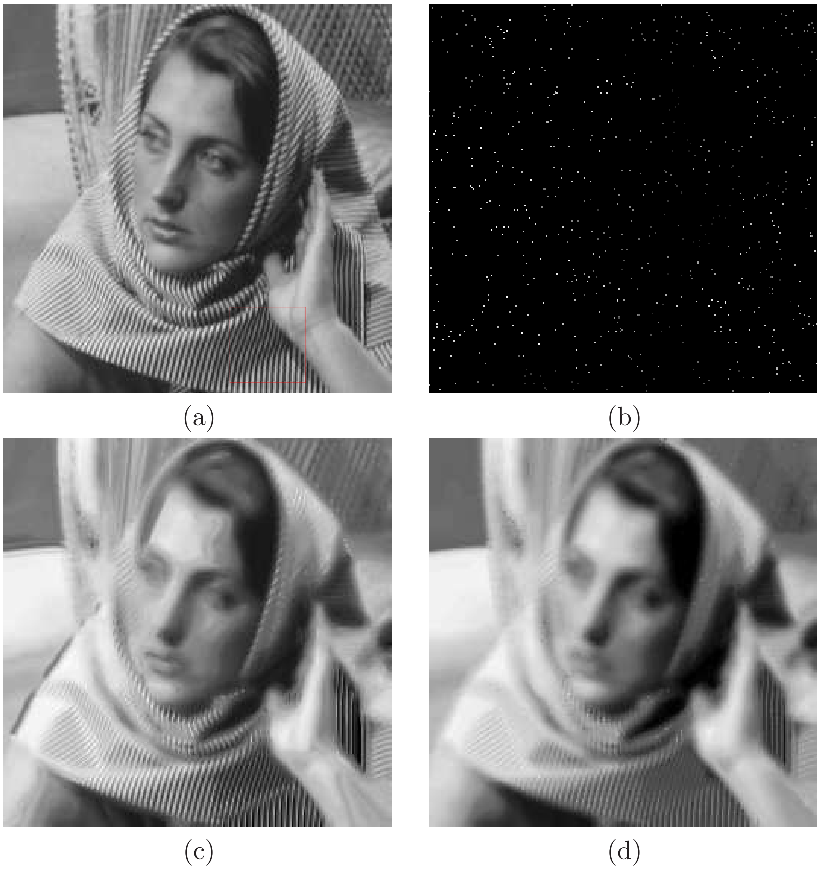
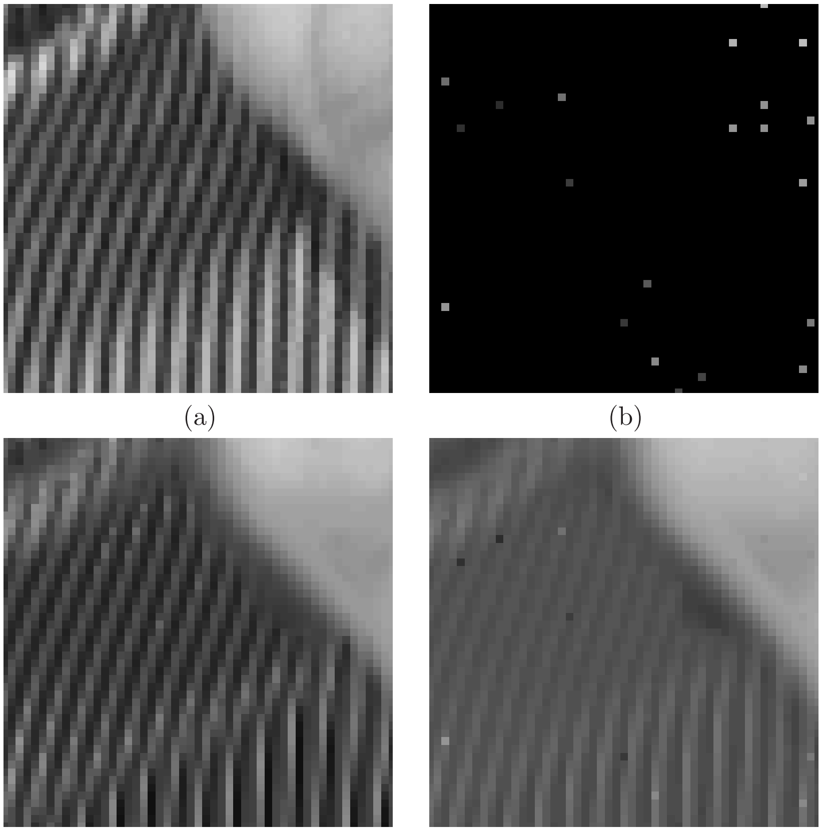


FIG. 4. *(a) original image; (b) subsampled image (1% of the pixels are retained); (c) recovered image by PIM; (d) recovered image by GL.*

a closer look at of the recovery, Figure 5 shows the zoom in image enclosed by the box in Figure 4(a). In Figure 5(d), there are many pixels which are not consistent with their neighbors. Comparing with the subsampled image 5(b), it is easy to check that these pixels are actually the retained pixels. This phenomenon suggests that in GLM, (4.2), the values at the retained pixels are not spread to their neighbours properly. The reason is that in GLM a non-negligible boundary term is dropped as we pointed in this paper. On the contrary, in PIM, the boundary term is retained and the resultant recovery is much better and smoother as shown in Figure 4(c) and 5(c).

As the sample rate grows, the inconsistency in the graph Laplacian may be alleviated. Figure 6(a) gives the recovery obtained by graph Laplacian from 30% subsamples. Visually, the result is much better and the inconsistent pixels disappear. This can be explained qualitatively by the theory of volume constraint [7, 16]. When the number of sample points increase, the sample points may accumulate together. Then the value of function is given on some volumes rather than the discrete points. Based on the theory of volume constraint, in this case, the Dirichlet boundary condition may be correctly enforced as long as the volume is larger than the support of the weight function. But the inconsistency can not completely removed in the graph Laplacian.



(a) (b) (c) (d)

FIG. 5. *Zoom in images. (a) original image; (b) subsampled image (1% of the pixels are retained); (c) recovered image by PIM; (d) recovered image by GL;*

Figure 6(c) shows the zoom in image enclosed by the box in Figure 6(a). Comparing with the result given by PIM, Figure 6(b)(d), the reconstruction given by PIM is much smoother and better.

5. Convergence of the Point Integral Method. In this section, we will establish the convergence results for the point integral method for solving the Laplace-Beltrami equation with Dirichlet boundary (5.1). To simplify the notation and make the proof concise, we consider the homogeneous Dirichlet boundary conditions, i.e.

$$\begin{cases} -\Delta_{\mathcal{M}} u(\mathbf{x}) = f(\mathbf{x}), & \mathbf{x} \in \mathcal{M} \\ u(\mathbf{x}) = 0, & \mathbf{x} \in \partial\mathcal{M} \end{cases} \tag{5.1}$$

The analysis can be easily generalized to the non-homogeneous boundary conditions.

In this section, we assume the point cloud P samples the submanifold $\mathcal{M}$ and a subset $S \subset P$ samples the boundary, $\partial\mathcal{M}$. List the points in P in a fixed order $P = (\mathbf{p}_1, \cdots, \mathbf{p}_n)$ where $\mathbf{p}_i \in \mathbb{R}^d, 1 \le i \le n$ and $S = (\mathbf{p}_1, \cdots, \mathbf{p}_m) \subset P$ with $m < n$. In addition, assume we are given two vectors $\mathbf{V} = (V_1, \cdots, V_n)$ where V_i is an volume weight of $\mathbf{p}_i$ in $\mathcal{M}$, and $\mathbf{A} = (A_1, \cdots, A_m)$ where A_i is an area weight of $\mathbf{p}_i$ in $\partial\mathcal{M}$.

The discretization of (5.1) in the point integral method over the point cloud

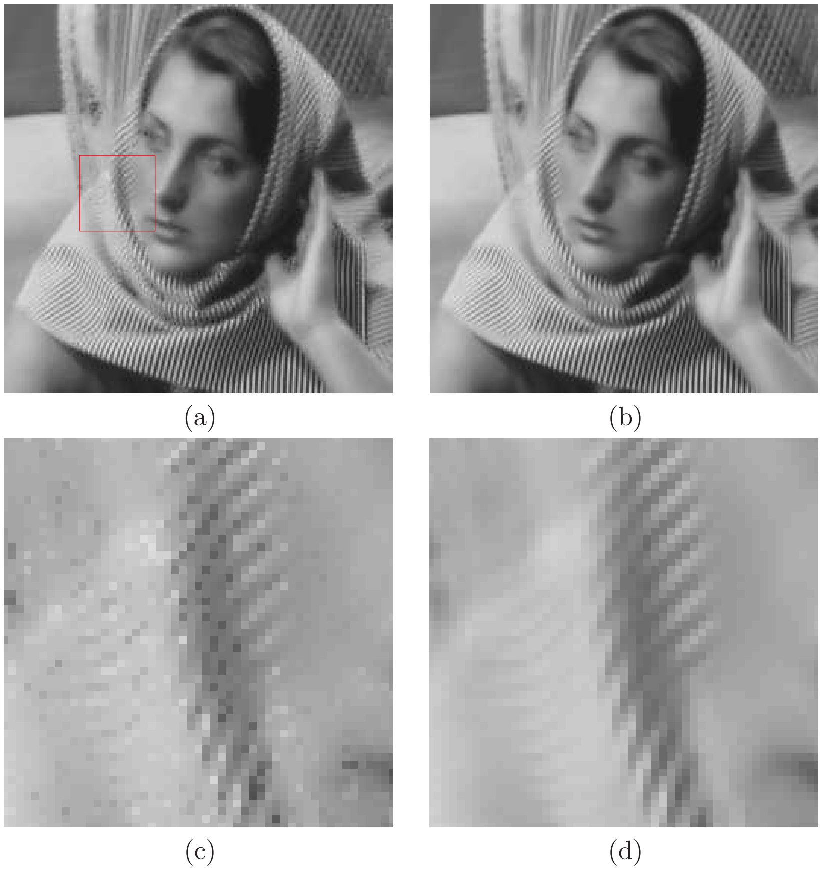


FIG. 6. *Recovered image by graph Laplacian from 30% subsamples. (a) recovered image by GL; (b) recovered image by PIM; (c) zoom in of the image reconstructed by GL; (c) zoom in of the image reconstructed by PIM.*

$(P, S, \mathbf{V}, \mathbf{A})$ is

$$\frac{1}{t}\sum_{j=1}^{n} R_t(\mathbf{p}_i,\mathbf{p}_j)(u_i-u_j)V_j + \frac{2}{\beta}\sum_{j=1}^{m}\bar{R}_t(\mathbf{p}_i,\mathbf{p}_j)u_jA_j = \sum_{j=1}^{n}\bar{R}_t(\mathbf{p}_i,\mathbf{p}_j)f_jV_j. \tag{5.2}$$

where $f_j = f(\mathbf{p}_j)$. In this section, we add a normalization factor C_t in the kernel function,

$$R_t(\mathbf{x},\mathbf{y}) = C_t R\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right), \quad \bar{R}_t(\mathbf{x},\mathbf{y}) = C_t \bar{R}\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right) \tag{5.3}$$

with $C_t = \frac{1}{(4\pi t)^{k/2}}$ and k is the dimension of the manifold $\mathcal{M}$. This factor does not change the discretization. It is introduced to normalize the kernel function which would be more convenient in theoretical analysis.

5.1. Assumptions and Results. Before proving the convergence of the point integral method, we need to clarify the meaning of the convergence between the point cloud $(P, S, \mathbf{V}, \mathbf{A})$ and the manifold $\mathcal{M}$. In this paper, we consider the convergence in the sense that

$$h(P, S, \mathbf{V}, \mathbf{A}, \mathcal{M}, \partial\mathcal{M}) \to 0$$

where $h(P, S, \mathbf{V}, \mathbf{A}, \mathcal{M}, \partial\mathcal{M})$ is the *integral accuracy index* defined as following,

Definition 5.1 (Integral Accuracy Index). *For the point cloud $(P, S, \mathbf{V}, \mathbf{A})$ which samples the manifold $\mathcal{M}$ and $\partial\mathcal{M}$, the integral accuracy index $h(P, S, \mathbf{V}, \mathbf{A}, \mathcal{M}, \partial\mathcal{M})$ is defined as*

$$h(P, S, \mathbf{V}, \mathbf{A}, \mathcal{M}, \partial\mathcal{M}) = \max\{h(P, \mathbf{V}, \mathcal{M}), h(S, \mathbf{A}, \partial\mathcal{M})\}$$

and

$$\begin{aligned} h(P, \mathbf{V}, \mathcal{M}) &= \sup_{f \in C^1(\mathcal{M})} \frac{\left|\int_{\mathcal{M}} f(\mathbf{y})\mathrm{d}\mathbf{y} - \sum_{i=1}^n f(\mathbf{p}_i)V_i\right|}{|\mathrm{supp}(f)|\|f\|_{C^1(\mathcal{M})}}, \\ h(S, \mathbf{A}, \partial\mathcal{M}) &= \sup_{g \in C^1(\partial\mathcal{M})} \frac{\left|\int_{\partial\mathcal{M}} g(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} - \sum_{i=1}^m g(\mathbf{p}_i)A_i\right|}{|\mathrm{supp}(g)|\|g\|_{C^1(\partial\mathcal{M})}} \end{aligned}$$

To simplify the notation, we denote $h = h(P, S, \mathbf{V}, \mathbf{A}, \mathcal{M}, \partial\mathcal{M})$ in the rest of the paper.

Using the definition of integrable index, we say that the point cloud $(P, S, \mathbf{V}, \mathbf{A})$ converges to the manifold $\mathcal{M}$ if $h \to 0$. The convergence analysis in this paper is based on the assumption that h is small enough.

To get the convergence, we also need some assumptions on the regularity of the submanifold $\mathcal{M}$ and the integral kernel function R.

Assumption 5.1.

- Smoothness of the manifold: $\mathcal{M}, \partial\mathcal{M}$ are both compact and C^∞ smooth k-dimensional submanifolds isometrically embedded in a Euclidean space $\mathbb{R}^d$.
- Assumptions on the kernel function $R(r)$:
  - (a) Smoothness: $R \in C^2(\mathbb{R}^+)$;
  - (b) Nonnegativity: $R(r) \geq 0$ for any $r \geq 0$.
  - (c) Compact support: $R(r) = 0$ for $\forall r > 1$;
  - (d) Nondegeneracy: $\exists \delta_0 > 0$ so that $R(r) \geq \delta_0$ for $0 \leq r \leq \frac{1}{2}$.

Remark 5.1. *The assumption on the kernel function is very mild. The compact support assumption can be relaxed to exponentially decay, like Gaussian kernel. In the nondegeneracy assumption, $1/2$ may be replaced by a positive number θ_0 with $0 < \theta_0 < 1$. Similar assumptions on the kernel function is also used in analysis the nonlocal diffusion problem [9].*

Remark 5.2. *The assumption that $(P, S, \mathbf{V}, \mathbf{A})$ is an h-integral approximation of $(\mathcal{M}, \partial\mathcal{M})$ is pretty mild. If the points in P and S are independent samples from uniform distribution on $\mathcal{M}$ and $\partial\mathcal{M}$ respectively, then $\mathbf{V}$ and $\mathbf{A}$ can be taken as the constant vector. From the central limit theorem, $(P, S, \mathbf{V}, \mathbf{A})$ is an h-integral approximation with h is of the order of $1/\sqrt{n}$.*

All the analysis in this paper is under the assumptions in Assumption 5.1 and h, t are small enough. In the theorems and the proof, without introducing any confusions, we omit the statement of the assumptions.

To compare the discrete numerical solution with the continuous exact solution, we interpolate the discrete solution $\mathbf{u} = (u_1, \cdots, u_n)$ of the problem (5.2) onto the smooth manifold using following interpolation operator:

$$I_{\mathbf{f}}(\mathbf{u})(\mathbf{x}) = \frac{\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{p}_j)u_j V_j - \frac{2t}{\beta}\sum_{j=1}^{m} \bar{R}_t(\mathbf{x},\mathbf{p}_j)u_j A_j + t\sum_{j=1}^{n} \bar{R}_t(\mathbf{x},\mathbf{p}_j)f_j V_j}{\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{p}_j)V_j}. \tag{5.4}$$

where $\mathbf{f} = [f_1, \cdots, f_n] = [f(\mathbf{p}_1), \cdots, f(\mathbf{p}_n)]$. It is easy to verify that $I_{\mathbf{f}}(\mathbf{u})$ interpolates $\mathbf{u}$ at the sample points P, i.e., $I_{\mathbf{f}}(\mathbf{u})(\mathbf{p}_j) = u_j$ for any j. In the analysis, $I_{\mathbf{f}}(\mathbf{u})$ is used as the numerical solution of (5.1) instead of the discrete solution $\mathbf{u}$.

Now, we can state the main result.

THEOREM 5.2. *Let u is the solution to Problem* (5.1) *with $f \in C^1(\mathcal{M})$. Set $\mathbf{f} = (f(\mathbf{p}_1), \cdots, f(\mathbf{p}_n))$. If the vector $\mathbf{u}$ is the solution to the problem* (5.2). *There exists constants C, T_0 and r_0 only depend on $\mathcal{M}$ and $\partial\mathcal{M}$, so that for any $t \le T_0$,*

$$\|u - I_{\mathbf{f}}(\mathbf{u})\|_{H^1(\mathcal{M})} \le C\left(\frac{h}{t^{3/2}} + t^{1/2} + \beta^{1/2}\right)\|f\|_{C^1(\mathcal{M})}. \tag{5.5}$$

as long as $\frac{h}{t^{3/2}} \le r_0$ and $\frac{\sqrt{t}}{\beta} \le r_0$.

5.2. Structure of the Proof. In the point integral method, we use Robin boundary problem (5.6) to approximate the Dirichlet boundary problem (5.1). First, we show that the solution of the Robin problem converges to the solution of the Dirichlet problem as the parameter $\beta \to 0$.

THEOREM 5.3. *Suppose u is the solution of the Dirichlet problem* (5.1) *and u_β is the solution of the Robin problem*

$$\begin{cases} -\Delta_{\mathcal{M}} u(\mathbf{x}) = f(\mathbf{x}), & \mathbf{x} \in \mathcal{M} \\ u(\mathbf{x}) + \beta\frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) = 0, & \mathbf{x} \in \partial\mathcal{M} \end{cases} \tag{5.6}$$

then

$$\|u - u_\beta\|_{H^1(\mathcal{M})} \le C\beta^{1/2}\|f\|_{L^2(\mathcal{M})}.$$

Proof. Let $w = u - u_\beta$, then w satisfies

$$\begin{cases} \Delta_{\mathcal{M}} w = 0, & \text{on } \mathcal{M}, \\ w + \beta\frac{\partial w}{\partial \mathbf{n}} = \beta\frac{\partial u}{\partial \mathbf{n}}, & \text{on } \partial\mathcal{M}. \end{cases}$$

By multiplying w on both sides of the equation and integrating by parts, we can get

$$\begin{aligned} 0 &= \int_{\mathcal{M}} w\Delta_{\mathcal{M}} w \mathrm{d}\mathbf{x} \\ &= -\int_{\mathcal{M}} |\nabla w|^2 \mathrm{d}\mathbf{x} + \int_{\partial\mathcal{M}} w\frac{\partial w}{\partial \mathbf{n}}\mathrm{d}\tau_{\mathbf{x}} \\ &= -\int_{\mathcal{M}} |\nabla w|^2 \mathrm{d}\mathbf{x} - \frac{1}{\beta}\int_{\partial\mathcal{M}} w^2 \mathrm{d}\tau_{\mathbf{x}} + \int_{\partial\mathcal{M}} w\frac{\partial u}{\partial \mathbf{n}}\mathrm{d}\tau_{\mathbf{x}} \\ &\le -\int_{\mathcal{M}} |\nabla w|^2 \mathrm{d}\mathbf{x} - \frac{1}{2\beta}\int_{\partial\mathcal{M}} w^2 \mathrm{d}\tau_{\mathbf{x}} + 2\beta\int_{\partial\mathcal{M}} \left|\frac{\partial u}{\partial \mathbf{n}}\right|^2 \mathrm{d}\tau_{\mathbf{x}}, \end{aligned}$$

which implies that

$$\int_{\mathcal{M}} |\nabla w|^2 \mathrm{d}\mathbf{x} + \frac{1}{2\beta}\int_{\partial\mathcal{M}} w^2 \mathrm{d}\tau_{\mathbf{x}} \le 2\beta \int_{\partial\mathcal{M}} \left|\frac{\partial u}{\partial \mathbf{n}}\right|^2 \mathrm{d}\tau_{\mathbf{x}}.$$

Moreover, we have

$$\|w\|_{L^2(\mathcal{M})}^2 \le C\left(\int_{\mathcal{M}} |\nabla w|^2 \mathrm{d}\mathbf{x} + \frac{1}{2\beta}\int_{\partial\mathcal{M}} w^2 \mathrm{d}\tau_{\mathbf{x}}\right) \le C\beta \int_{\partial\mathcal{M}} \left|\frac{\partial u}{\partial \mathbf{n}}\right|^2 \mathrm{d}\tau_{\mathbf{x}}.$$

Combining above two inequalities and using the trace theorem, we get

$$\|u-u_\beta\|_{H^1(\mathcal{M})} \le C\beta^{1/2}\left\|\frac{\partial u}{\partial \mathbf{n}}\right\|_{L^2(\partial\mathcal{M})} \le C\beta^{1/2}\|u\|_{H^2(\mathcal{M})}.$$

The proof is complete using that

$$\|u\|_{H^2(\mathcal{M})} \le C\|f\|_{L^2(\mathcal{M})}.$$

□

Next, we prove the solution of (5.2) converges to the solution of the Robin prolem (5.6) as h, t go to 0. Comparing to the Neumann boundary problem considered in [18], in (5.2), the unknown variables u_i not only appear in the discrete Laplace operator L_t, but also appear in an integral over the boundary. Therefore, instead of showing the stability for the integral Laplace operator L_t as in [18], we need to consider the stability for the following integral operator

$$K_t u(\mathbf{x}) = \frac{1}{t}\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x})-u(\mathbf{y}))\mathrm{d}\mathbf{y} + \frac{2}{\beta}\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}. \tag{5.7}$$

This is the most difficult part in this paper.

THEOREM 5.4. *Let $u(\mathbf{x})$ solves following equation with $r \in H^1(\mathcal{M})$*

$$K_t u = r.$$

Then, there exist constants $C, T_0, r_0 > 0$ independent on t, such that

$$\|u\|_{H^1(\mathcal{M})} \le C\left(\|r\|_{L^2(\mathcal{M})} + \frac{t}{\sqrt{\beta}}\|r\|_{H^1(\mathcal{M})}\right),$$

as long as $t \le T_0$ and $\frac{\sqrt{t}}{\beta} \le r_0$.

To apply the stability result, we need L_2 estimate of $K_t(u_\beta - I_{\mathbf{f}}(\mathbf{u}))$ and $\nabla K_t(u_\beta - I_{\mathbf{f}}(\mathbf{u}))$. In the analysis, the truncation error $K_t(u_\beta - I_{\mathbf{f}}(\mathbf{u}))$ is further splitted to two terms

$$K_t(u_\beta - I_{\mathbf{f}}(\mathbf{u})) = K_t(u_\beta - u_{\beta,t})) + K_t(u_{\beta,t} - I_{\mathbf{f}}(\mathbf{u}))$$

where $u_{\beta,t}$ is the solution of the integral equation

$$\frac{1}{t}\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x})-u(\mathbf{y}))\mathrm{d}\mathbf{y} + \frac{2}{\beta}\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} = \int_{\mathcal{M}} f(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y}. \tag{5.8}$$

The first term $K_t(u_\beta - u_{\beta,t})$ is same as that in the Neumann boundary problem [18]. It also has boundary layer structure.

THEOREM 5.5. *Let $u(\mathbf{x})$ be the solution of the problem* (5.1) *and $u_t(\mathbf{x})$ be the solution of the corresponding integral equation* (5.8). *Let*

$$I_{bd} = \sum_{j=1}^{d} \int_{\partial\mathcal{M}} n^j(\mathbf{y})(\mathbf{x}-\mathbf{y}) \cdot \nabla(\nabla^j u(\mathbf{y}))\bar{R}_t(\mathbf{x},\mathbf{y})p(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}, \tag{5.9}$$

and

$$K_t(u-u_t) = I_{in} + I_{bd}.$$

where $\mathbf{n}(\mathbf{y}) = (n^1(\mathbf{y}), \cdots, n^d(\mathbf{y}))$ is the out normal vector of $\partial\mathcal{M}$ at $\mathbf{y}$, ∇^j is the jth component of gradient ∇.

If $u \in H^3(\mathcal{M})$, then there exists constants C, T_0 depending only on $\mathcal{M}$ and $p(\mathbf{x})$, so that,

$$\|I_{in}\|_{L^2(\mathcal{M})} \leq Ct^{1/2}\|u\|_{H^3(\mathcal{M})}, \quad \|\nabla I_{in}\|_{L^2(\mathcal{M})} \leq C\|u\|_{H^3(\mathcal{M})}, \tag{5.10}$$

as long as $t \leq T_0$.

The estimate of the second term, $K_t(u_{\beta,t} - I_{\mathbf{f}}(\mathbf{u}))$, is given in following theorem.

THEOREM 5.6. *Let $u_t(\mathbf{x})$ be the solution of the problem* (5.8) *and $\mathbf{u}$ be the solution of the problem* (5.2). *If $f \in C^1(\mathcal{M})$, then there exists constants C, T_0 depending only on $\mathcal{M}$, so that*

$$\|K_t\left(I_{\mathbf{f}}\mathbf{u} - u_t\right)\|_{L^2(\mathcal{M})} \leq \frac{Ch}{t^{3/2}}\|f\|_{C^1(\mathcal{M})}, \tag{5.11}$$

$$\|\nabla K_t\left(I_{\mathbf{f}}\mathbf{u} - u_t\right)\|_{L^2(\mathcal{M})} \leq \frac{Ch}{t^2}\|f\|_{C^1(\mathcal{M})}. \tag{5.12}$$

as long as $t \leq T_0$ and $\frac{h}{\sqrt{t}} \leq T_0$.

Corresponding to the boundary layer structure in Theorem 5.5, we need stability of K_t for the boundary term.

THEOREM 5.7. *Let $u(\mathbf{x})$ solves the integral equation*

$$K_t u(\mathbf{x}) = \int_{\partial\mathcal{M}} \mathbf{b}(\mathbf{y}) \cdot (\mathbf{x}-\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}.$$

There exist constant $C > 0, T_0 > 0$ independent on t, such that

$$\|u\|_{H^1(\mathcal{M})} \leq C\sqrt{t}\;\|\mathbf{b}\|_{H^1(\mathcal{M})}.$$

as long as $t \leq T_0$.

Theorem 5.2 is an easy corollary from Theorems 5.3, Theorems 5.4, 5.6, 5.5 and 5.7. The detailed proof is omitted here.

Proof of Theorem 5.5 is essentially a special case with constant coefficients of Theorem 3.5 in [13]. In the rest of the paper, we prove Theorem 5.4, 5.6 and 5.7 respectively.

6. Discussion and Future Work. In this paper, we applied the point integral method to solve the harmonic extension problem. We found that the graph Laplacian has inconsistent problem since one important boundary term is dropped. The point integral method gives a consistent discretization for the harmonic extension. We compared the performance of the point integral method with that of graph Laplacian in the application of semi-supervised learning and image recovery. In the future, we will test this method on more datasets and find different applications of harmonic extension.

We also prove the convergence of the point integral method for Laplace-Beltrami equation on manifolds with the Dirichlet boundary. In point integral method, the Dirichlet boundary can not be enforced directly. In this paper, we use Robin boundary to approximate the Dirichlet boundary and use point integral method to solve the Poisson equation with Robin boundary condition.

Another way to deal with the Dirichlet boundary condition in point integral method is using the volume constraint proposed by Du et.al. [8]. The volume constraint has been integrated into the point integral method to enforce the Dirichlet boundary condition and the convergence has been proved [16].

Appendix A. Stability of K_t (Theorem 5.4 and 5.7). In this section, we will prove Theorem 5.4 and 5.7. Both these two theorems are concerned with the stability of K_t, which are essential in the convergence analysis.

To simplify the notation, we introduce an integral operator, L_t,

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

In the proof, we need following theorem which has been proved in [18].

THEOREM A.1. *For any function $u \in L^2(\mathcal{M})$, there exists a constant $C > 0$ independent on t and u, such that*

$$\langle u, L_t u\rangle_{\mathcal{M}} \geq C\int_{\mathcal{M}} |\nabla v|^2 \mathrm{d}\mathbf{x}$$

where $\langle f, g\rangle_{\mathcal{M}} = \int_{\mathcal{M}} f(\mathbf{x})g(\mathbf{x})\mathrm{d}\mathbf{x}$ for any $f, g \in L_2(\mathcal{M})$, and

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

and $w_t(\mathbf{x}) = C_t\int_{\mathcal{M}} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right)\mathrm{d}\mathbf{y}$.

A.1. Stability of K_t for interior term (Theorem 5.4). Using Theorem A.1, we have

$$\|\nabla v\|^2_{L^2(\mathcal{M})} \leq C\langle u, L_t u\rangle = \int_{\mathcal{M}} u(\mathbf{x})r(\mathbf{x})\mathrm{d}\mathbf{x} - \frac{2}{\beta}\int_{\mathcal{M}} u(\mathbf{x})\left(\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}. \tag{A.3}$$

where v is the same as defined in Theorem A.1. We control the second term on the right hand side of (A.3) as follows.

$$\begin{aligned}
&\left|\int_{\mathcal{M}} u(\mathbf{x})\left(\int_{\partial\mathcal{M}}\left(\bar{R}_t(\mathbf{x},\mathbf{y}) - \frac{\bar{w}_t(\mathbf{y})}{w_t(\mathbf{y})}R_t(\mathbf{x},\mathbf{y})\right)u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}\right| \\
=& \left|\int_{\partial\mathcal{M}} u(\mathbf{y})\left(\int_{\mathcal{M}}\left(\bar{R}_t(\mathbf{x},\mathbf{y}) - \frac{\bar{w}_t(\mathbf{y})}{w_t(\mathbf{y})}R_t(\mathbf{x},\mathbf{y})\right)u(\mathbf{x})\mathrm{d}\mathbf{x}\right)\mathrm{d}\tau_{\mathbf{y}}\right| \\
=& \left|\int_{\partial\mathcal{M}} \frac{1}{w_t(\mathbf{y})}u(\mathbf{y})\left(\int_{\mathcal{M}}\left(w_t(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y}) - \bar{w}_t(\mathbf{y})R_t(\mathbf{x},\mathbf{y})\right)u(\mathbf{x})\mathrm{d}\mathbf{x}\right)\mathrm{d}\tau_{\mathbf{y}}\right| \\
\leq& C\|u\|_{L^2(\partial\mathcal{M})}\left(\int_{\partial\mathcal{M}}\left(\int_{\mathcal{M}}\left(w_t(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y}) - \bar{w}_t(\mathbf{y})R_t(\mathbf{x},\mathbf{y})\right)u(\mathbf{x})\mathrm{d}\mathbf{x}\right)^2\mathrm{d}\tau_{\mathbf{y}}\right)^{1/2},
\end{aligned}$$

where $\bar{w}_t(\mathbf{x}) = \int_{\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{y}$. Noticing that

$$\begin{aligned}
&\int_{\mathcal{M}} \left(w_t(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y}) - \bar{w}_t(\mathbf{y})R_t(\mathbf{x},\mathbf{y})\right) u(\mathbf{x})\mathrm{d}\mathbf{x} \\
&= \int_{\mathcal{M}}\int_{\mathcal{M}} R_t(\mathbf{y},\mathbf{z})\bar{R}_t(\mathbf{x},\mathbf{y})\left(u(\mathbf{x}) - u(\mathbf{z})\right)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z},
\end{aligned}$$

we have

$$\begin{aligned}
&\int_{\partial\mathcal{M}} \left(\int_{\mathcal{M}} \left(w_t(\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y}) - \bar{w}_t(\mathbf{y})R_t(\mathbf{x},\mathbf{y})\right) u(\mathbf{x})\mathrm{d}\mathbf{x}\right)^2 \mathrm{d}\tau_{\mathbf{y}} \\
&\leq \int_{\partial\mathcal{M}} \left(\int_{\mathcal{M}}\int_{\mathcal{M}} R_t(\mathbf{y},\mathbf{z})\bar{R}_t(\mathbf{x},\mathbf{y})\left(u(\mathbf{x}) - u(\mathbf{z})\right)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}\right)^2 \mathrm{d}\tau_{\mathbf{y}} \\
&\leq \int_{\partial\mathcal{M}} \left(\int_{\mathcal{M}}\int_{\mathcal{M}} R_t(\mathbf{y},\mathbf{z})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}\right)\left(\int_{\mathcal{M}}\int_{\mathcal{M}} R_t(\mathbf{y},\mathbf{z})\bar{R}_t(\mathbf{x},\mathbf{y})\left(u(\mathbf{x}) - u(\mathbf{z})\right)^2 \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}\right) \mathrm{d}\tau_{\mathbf{y}} \\
&\leq C\left(\int_{\mathcal{M}}\int_{\mathcal{M}} \left(\int_{\partial\mathcal{M}} R_t(\mathbf{y},\mathbf{z})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\left(u(\mathbf{x}) - u(\mathbf{z})\right)^2 \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}\right) \\
&= C\left(\int_{\mathcal{M}}\int_{\mathcal{M}} Q(\mathbf{x},\mathbf{z})\left(u(\mathbf{x}) - u(\mathbf{z})\right)^2 \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}\right),
\end{aligned}$$

where

$$Q(\mathbf{x},\mathbf{z}) = \int_{\partial\mathcal{M}} R_t(\mathbf{y},\mathbf{z})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}.$$

Notice that $Q(\mathbf{x},\mathbf{z}) = 0$ if $\|\mathbf{y}-\mathbf{z}\|^2 \geq 16t$, and $|Q(\mathbf{x},\mathbf{z})| \leq CC_t/\sqrt{t}$. We have

$$|Q(\mathbf{x},\mathbf{z})| \leq \frac{CC_t}{\sqrt{t}} R\left(\frac{\|\mathbf{x}-\mathbf{z}\|^2}{32t}\right).$$

Then, we obtain the following estimate,

$$\begin{aligned}
&\left|\left(\int_{\mathcal{M}}\int_{\mathcal{M}} Q(\mathbf{x},\mathbf{z})\left(u(\mathbf{x}) - u(\mathbf{z})\right)^2 \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}\right)\right| \\
&\leq \left|\frac{C}{\sqrt{t}}\left(\int_{\mathcal{M}}\int_{\mathcal{M}} C_t R\left(\frac{\|\mathbf{x}-\mathbf{z}\|^2}{32t}\right)\left(u(\mathbf{x}) - u(\mathbf{z})\right)^2 \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}\right)\right| \\
&\leq \left|\frac{C}{\sqrt{t}}\left(\int_{\mathcal{M}}\int_{\mathcal{M}} C_t R\left(\frac{\|\mathbf{x}-\mathbf{z}\|^2}{4t}\right)\left(u(\mathbf{x}) - u(\mathbf{z})\right)^2 \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{z}\right)\right| \\
&\leq C\sqrt{t}\left(\left|\int_{\mathcal{M}} u(\mathbf{x})r(\mathbf{x})\mathrm{d}\mathbf{x}\right| + \frac{1}{\beta}\left|\int_{\mathcal{M}} u(\mathbf{x})\left(\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}\right|\right) \\
&\leq C\sqrt{t}\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} + \frac{C\sqrt{t}}{\beta}\left|\int_{\mathcal{M}} u(\mathbf{x})\left(\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}\right|.
\end{aligned} \tag{A.4}$$

On the other hand,

$$\begin{aligned}
&\int_{\mathcal{M}} u(\mathbf{x})\left(\int_{\partial\mathcal{M}} \frac{\bar{w}_t(\mathbf{y})}{w_t(\mathbf{y})} R_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x} \\
&= \int_{\partial\mathcal{M}} \frac{\bar{w}_t(\mathbf{y})}{w_t(\mathbf{y})} u(\mathbf{y})\left(\int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y})(u(\mathbf{x}) - u(\mathbf{y}))\mathrm{d}\mathbf{x}\right)\mathrm{d}\tau_{\mathbf{y}} + \int_{\partial\mathcal{M}} \bar{w}_t(\mathbf{y})u^2(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} \\
&= \int_{\partial\mathcal{M}} \bar{w}_t(\mathbf{y})u(\mathbf{y})\left(v(\mathbf{y}) - u(\mathbf{y})\right)\mathrm{d}\tau_{\mathbf{y}} + \int_{\partial\mathcal{M}} \bar{w}_t(\mathbf{y})u^2(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}},
\end{aligned}$$

where v is the same as defined in (A.2). Since u solves $K_t u = r(\mathbf{x})$, we have

$$w_t(\mathbf{x})u(\mathbf{x}) = w_t(\mathbf{x})v(\mathbf{x}) - \frac{2t}{\beta}\int_{\partial\mathcal{M}} R_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} - t\, r(\mathbf{x}). \tag{A.5}$$

Then, we obtain

$$\begin{aligned}
&\int_{\partial\mathcal{M}} \bar{w}_t(\mathbf{y})u(\mathbf{y})\left(v(\mathbf{y}) - u(\mathbf{y})\right)\mathrm{d}\tau_{\mathbf{y}} \\
&= \int_{\partial\mathcal{M}} \frac{\bar{w}_t(\mathbf{y})}{w_t(\mathbf{y})}u(\mathbf{y})\left(\frac{2t}{\beta}\int_{\partial\mathcal{M}} R_t(\mathbf{x},\mathbf{y})u(\mathbf{x})\mathrm{d}\tau_{\mathbf{x}} - t\, r(\mathbf{y})\right)\mathrm{d}\tau_{\mathbf{y}} \\
&\le \frac{C\sqrt{t}}{\beta}\|u\|^2_{L^2(\partial\mathcal{M})} + Ct\|u\|_{L^2(\partial\mathcal{M})}\|r\|_{L^2(\partial\mathcal{M})} \\
&\le \frac{C\sqrt{t}}{\beta}\|u\|^2_{L^2(\partial\mathcal{M})} + Ct\|u\|_{L^2(\partial\mathcal{M})}\|r\|_{H^1(\mathcal{M})}.
\end{aligned}$$

Combining above estimates together, we have

$$\begin{aligned}
&\int_{\mathcal{M}} u(\mathbf{x})\left(\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x} \\
\ge& \int_{\partial\mathcal{M}} \bar{w}_t(\mathbf{y})u^2(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} - \frac{C\sqrt{t}}{\beta}\|u\|^2_{L^2(\partial\mathcal{M})} - Ct\|u\|_{L^2(\partial\mathcal{M})}\|r\|_{H^1(\mathcal{M})} \\
&-C\sqrt{t}\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} - \frac{C\sqrt{t}}{\beta}\left|\int_{\mathcal{M}} u(\mathbf{x})\left(\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}\right|.
\end{aligned}$$

We can choose $\frac{\sqrt{t}}{\beta}$ small enough such that $\frac{C\sqrt{t}}{\beta} \le \min\{\frac{1}{2}, \frac{w_{\min}}{6}\}$ with $w_{\min} = \min_{\mathbf{x}\in\mathcal{M}} w_t(\mathbf{x})$, which gives us

$$\begin{aligned}
&\int_{\mathcal{M}} u(\mathbf{x})\left(\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x} \\
\ge& \frac{2}{3}\int_{\partial\mathcal{M}} \bar{w}_t(\mathbf{y})u^2(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} - \frac{C\sqrt{t}}{\beta}\|u\|^2_{L^2(\partial\mathcal{M})} - Ct\|u\|_{L^2(\partial\mathcal{M})}\|r\|_{H^1(\mathcal{M})} - C\sqrt{t}\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} \\
\ge& \frac{w_{\min}}{2}\|u\|^2_{L^2(\partial\mathcal{M})} - Ct\|u\|_{L^2(\partial\mathcal{M})}\|r\|_{H^1(\mathcal{M})} - C\sqrt{t}\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} \\
\ge& \frac{w_{\min}}{4}\|u\|^2_{L^2(\partial\mathcal{M})} - Ct^2\|r\|^2_{H^1(\mathcal{M})} - C\sqrt{t}\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})}
\end{aligned}$$

Substituting the above estimate to the first inequality (A.3), we obtain

$$\begin{aligned}
&\|\nabla v\|_{L^2(\mathcal{M})} + \frac{w_{\min}}{4\beta}\|u\|^2_{L^2(\partial\mathcal{M})} \\
&\le -C\int_{\mathcal{M}} u(\mathbf{x})r(\mathbf{x})\mathrm{d}\mathbf{x} + \frac{Ct^2}{\beta}\|r\|^2_{H^1(\mathcal{M})} + \frac{C\sqrt{t}}{\beta}\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} \\
&\le C\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} + \frac{Ct^2}{\beta}\|r\|^2_{H^1(\mathcal{M})}.
\end{aligned} \tag{A.6}$$

Here we require that $\frac{\sqrt{t}}{\beta}$ is bounded by a constant independent on β and t. Now, using the representation of u given in (A.5), we obtain

$$\begin{aligned}
&\|\nabla u\|_{L^2(\mathcal{M})}^2 + \frac{w_{\min}}{8\beta}\|u\|_{L^2(\partial\mathcal{M})}^2 \\
\le\ & C\|\nabla v\|_{L^2(\mathcal{M})}^2 + \frac{Ct^2}{\beta^2}\left\|\nabla\left(\frac{1}{w_t(\mathbf{x})}\int_{\partial\mathcal{M}} R_t(\mathbf{x},\mathbf{y})u(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\right\|_{L^2(\mathcal{M})}^2 \\
&+Ct^2\left\|\nabla\left(\frac{r(\mathbf{x})}{w_t(\mathbf{x})}\right)\right\|_{L^2(\mathcal{M})}^2 + \frac{w_{\min}}{8\beta}\|u\|_{L^2(\partial\mathcal{M})}^2 \\
\le\ & C\|\nabla v\|_{L^2(\mathcal{M})}^2 + \left(\frac{C\sqrt{t}}{\beta^2} + \frac{w_{\min}}{8\beta}\right)\|u\|_{L^2(\partial\mathcal{M})}^2 + Ct\|r\|_{L^2(\mathcal{M})}^2 + Ct^2\|r\|_{H^1(\mathcal{M})}^2 \\
\le\ & C\|\nabla v\|_{L^2(\mathcal{M})}^2 + \frac{w_{\min}}{4\beta}\|u\|_{L^2(\partial\mathcal{M})}^2 + Ct\|r\|_{L^2(\mathcal{M})}^2 + Ct^2\|r\|_{H^1(\mathcal{M})}^2 \\
\le\ & C\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} + Ct\|r\|_{L^2(\mathcal{M})}^2 + \frac{Ct^2}{\beta}\|r\|_{H^1(\mathcal{M})}^2.
\end{aligned}$$

Here we require that $\frac{C\sqrt{t}}{\beta} \le \frac{w_{\min}}{8}$ in the third inequality. Furthermore, we have

$$\begin{aligned}
&\|u\|_{L^2(\mathcal{M})}^2 \le C\left(\|\nabla u\|_{L^2(\mathcal{M})}^2 + \frac{w_{\min}}{8\beta}\|u\|_{L^2(\partial\mathcal{M})}^2\right) \\
\le\ & C\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} + Ct\|r\|_{L^2(\mathcal{M})}^2 + \frac{Ct^2}{\beta}\|r\|_{H^1(\mathcal{M})}^2 \\
\le\ & \frac{1}{2}\|u\|_{L^2(\mathcal{M})}^2 + C\|r\|_{L^2(\mathcal{M})}^2 + \frac{Ct^2}{\beta}\|r\|_{H^1(\mathcal{M})}^2,
\end{aligned}$$

which implies that

$$\|u\|_{L^2(\mathcal{M})} \le C\left(\|r\|_{L^2(\mathcal{M})} + \frac{t}{\sqrt{\beta}}\|r\|_{H^1(\mathcal{M})}\right).$$

Finally, we obtain

$$\begin{aligned}
\|\nabla u\|_{L^2(\mathcal{M})}^2 &\le C\|u\|_{L^2(\mathcal{M})}\|r\|_{L^2(\mathcal{M})} + \frac{Ct^2}{\beta}\|r\|_{H^1(\mathcal{M})}^2 \\
&\le C\left(\|r\|_{L^2(\mathcal{M})} + \frac{t}{\sqrt{\beta}}\|r\|_{H^1(\mathcal{M})}\right)^2,
\end{aligned}$$

which completes the proof.

A.2. Stability of K_t for boundary term (Theorem 5.7). First, we denote

$$r(\mathbf{x}) = \int_{\partial\mathcal{M}} \mathbf{b}(\mathbf{y})\cdot(\mathbf{x}-\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}.$$

The key point of the proof is to show that

$$\left|\int_{\mathcal{M}} u(\mathbf{x})r(\mathbf{x})\mathrm{d}\mathbf{x}\right| \le C\sqrt{t}\,\|\mathbf{b}\|_{H^1(\mathcal{M})}\|u\|_{H^1(\mathcal{M})}. \tag{A.7}$$

Direct calculation gives that

$$|2t\nabla\bar{\bar{R}}_t(\mathbf{x},\mathbf{y}) - (\mathbf{x}-\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y})| \le C|\mathbf{x}-\mathbf{y}|^2\bar{R}_t(\mathbf{x},\mathbf{y}),$$

where $\bar{\bar{R}}_t(\mathbf{x},\mathbf{y}) = C_t\bar{\bar{R}}\left(\frac{\|\mathbf{x}-\mathbf{y}\|^2}{4t}\right)$ and $\bar{\bar{R}}(r) = \int_r^\infty \bar{R}(s)\mathrm{d}s$. This implies that

$$\begin{aligned}
\text{(A.8)}\quad & \left|\int_{\mathcal{M}} u(\mathbf{x})\int_{\partial\mathcal{M}} \mathbf{b}(\mathbf{y})\left((\mathbf{x}-\mathbf{y})\bar{R}_t(\mathbf{x},\mathbf{y}) + 2t\nabla\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\right)\mathrm{d}\tau_{\mathbf{y}}\mathrm{d}\mathbf{x}\right| \\
\leq & C\int_{\mathcal{M}} |u(\mathbf{x})|\int_{\partial\mathcal{M}} |\mathbf{b}(\mathbf{y})||\mathbf{x}-\mathbf{y}|^2\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\mathrm{d}\mathbf{x} \\
\leq & Ct\|\mathbf{b}\|_{L^2(\partial\mathcal{M})}\left(\int_{\partial\mathcal{M}}\left(\int_{\mathcal{M}}\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\right)\left(\int_{\mathcal{M}}|u(\mathbf{x})|^2\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\right)\mathrm{d}\tau_{\mathbf{y}}\right)^{1/2} \\
\leq & Ct\|\mathbf{b}\|_{H^1(\mathcal{M})}\left(\int_{\mathcal{M}}|u(\mathbf{x})|^2\left(\int_{\partial\mathcal{M}}\bar{R}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}\right)^{1/2} \\
\leq & Ct^{3/4}\|\mathbf{b}\|_{H^1(\mathcal{M})}\|u\|_{L^2(\mathcal{M})}.
\end{aligned}$$

On the other hand, using the Gauss integral formula, we have

$$\begin{aligned}
\text{(A.9)}\quad & \int_{\mathcal{M}} u(\mathbf{x})\int_{\partial\mathcal{M}}\mathbf{b}(\mathbf{y})\cdot\nabla\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\mathrm{d}\mathbf{x} \\
= & \int_{\partial\mathcal{M}}\int_{\mathcal{M}} u(\mathbf{x})T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))\cdot\nabla\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\tau_{\mathbf{y}} \\
= & \int_{\partial\mathcal{M}}\int_{\partial\mathcal{M}}\mathbf{n}(\mathbf{x})\cdot T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))u(\mathbf{x})\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{x}}\mathrm{d}\tau_{\mathbf{y}} \\
& -\int_{\partial\mathcal{M}}\int_{\mathcal{M}}\mathrm{div}_{\mathbf{x}}[u(\mathbf{x})T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))]\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\tau_{\mathbf{y}}.
\end{aligned}$$

Here $T_{\mathbf{x}}$ is the projection operator to the tangent space on $\mathbf{x}$. To get the first equality, we use the fact that $\nabla\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})$ belongs to the tangent space on $\mathbf{x}$, such that $\mathbf{b}(\mathbf{y})\cdot\nabla\bar{\bar{R}}_t(\mathbf{x},\mathbf{y}) = T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))\cdot\nabla\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})$ and $\mathbf{n}(\mathbf{x})\cdot T_{\mathbf{x}}(\mathbf{b}(\mathbf{y})) = \mathbf{n}(\mathbf{x})\cdot\mathbf{b}(\mathbf{y})$ where $\mathbf{n}(\mathbf{x})$ is the out normal of $\partial\mathcal{M}$ at $\mathbf{x}\in\partial\mathcal{M}$.

For the first term, we have

$$\begin{aligned}
\text{(A.10)}\quad & \left|\int_{\partial\mathcal{M}}\int_{\partial\mathcal{M}}\mathbf{n}(\mathbf{x})\cdot T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))u(\mathbf{x})\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{x}}\mathrm{d}\tau_{\mathbf{y}}\right| \\
= & \left|\int_{\partial\mathcal{M}}\int_{\partial\mathcal{M}}\mathbf{n}(\mathbf{x})\cdot\mathbf{b}(\mathbf{y})u(\mathbf{x})\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{x}}\mathrm{d}\tau_{\mathbf{y}}\right| \\
\leq & C\|\mathbf{b}\|_{L^2(\partial\mathcal{M})}\left(\int_{\partial\mathcal{M}}\left(\int_{\partial\mathcal{M}}|u(\mathbf{x})|\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{x}}\right)^2\mathrm{d}\tau_{\mathbf{y}}\right)^{1/2} \\
\leq & C\|\mathbf{b}\|_{H^1(\mathcal{M})}\left(\int_{\partial\mathcal{M}}\left(\int_{\partial\mathcal{M}}\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{x}}\right)\left(\int_{\partial\mathcal{M}}|u(\mathbf{x})|^2\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{x}}\right)\mathrm{d}\tau_{\mathbf{y}}\right)^{1/2} \\
\leq & Ct^{-1/2}\ \|\mathbf{b}\|_{H^1(\mathcal{M})}\|u\|_{L^2(\partial\mathcal{M})} \leq Ct^{-1/2}\ \|\mathbf{b}\|_{H^1(\mathcal{M})}\|u\|_{H^1(\mathcal{M})}.
\end{aligned}$$

We can also bound the second term on the right hand side of (A.9). By using the assumption that $\mathcal{M}\in C^\infty$, we have

$$\begin{aligned}
& |\mathrm{div}_{\mathbf{x}}[u(\mathbf{x})T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))]| \\
\leq & |\nabla u(\mathbf{x})||T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))||| + |u(\mathbf{x})||\mathrm{div}_{\mathbf{x}}[T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))]||| + |\nabla||u(\mathbf{x})T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))| \\
\leq & C(|\nabla u(\mathbf{x})| + |u(\mathbf{x})|)|\mathbf{b}(\mathbf{y})|
\end{aligned}$$

where the constant C depends on the curvature of the manifold $\mathcal{M}$.

Then, we have

$$
\begin{aligned}
\text{(A.11)} \qquad & \left|\int_{\partial\mathcal{M}}\int_{\mathcal{M}} \mathrm{div}_{\mathbf{x}}[u(\mathbf{x})T_{\mathbf{x}}(\mathbf{b}(\mathbf{y}))]\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\tau_{\mathbf{y}}\right| \\
&\le C\int_{\partial\mathcal{M}}\mathbf{b}(\mathbf{y})\int_{\mathcal{M}}(|\nabla u(\mathbf{x})|+|u(\mathbf{x})|)\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\tau_{\mathbf{y}} \\
&\le C\|\mathbf{b}\|_{L^2(\partial\mathcal{M})}\left(\int_{\mathcal{M}}(|\nabla u(\mathbf{x})|^2+|u(\mathbf{x})|^2)\left(\int_{\partial\mathcal{M}}\bar{\bar{R}}_t(\mathbf{x},\mathbf{y})\mathrm{d}\tau_{\mathbf{y}}\right)\mathrm{d}\mathbf{x}\right)^{1/2} \\
&\le Ct^{-1/4}\,\|\mathbf{b}\|_{H^1(\mathcal{M})}\|u\|_{H^1(\mathcal{M})}.
\end{aligned}
$$

Then, the inequality (A.7) is obtained from (A.8), (A.9), (A.10) and (A.11).

Following the proof of Theorem 5.4, in (A.4) and (A.6), we bound $\left|\int_{\mathcal{M}} u(\mathbf{x})r(\mathbf{x})\mathrm{d}\mathbf{x}\right|$ by $C\sqrt{t}\,\|\mathbf{b}\|_{H^1(\mathcal{M})}\|u\|_{H^1(\mathcal{M})}$, which implies that

$$
\begin{aligned}
&\|\nabla u\|^2_{L^2(\mathcal{M})}+\frac{w_{\min}}{8\beta}\|u\|^2_{L^2(\partial\mathcal{M})} \\
&\le C\sqrt{t}\,\|\mathbf{b}\|_{H^1(\mathcal{M})}\|u\|_{H^1(\mathcal{M})}+Ct\|r\|^2_{L^2(\mathcal{M})}+\frac{Ct^2}{\beta}\|r\|^2_{H^1(\mathcal{M})} \\
&\le C\|\mathbf{b}\|_{H^1(\mathcal{M})}\left(\sqrt{t}\|u\|_{H^1(\mathcal{M})}+t\right)
\end{aligned}
$$

where we use the estimates that

$$
\begin{aligned}
\|r(\mathbf{x})\|_{L^2(\mathcal{M})} &\le Ct^{1/4}\|\mathbf{b}\|_{H^1(\mathcal{M})}, \\
\|r(\mathbf{x})\|_{H^1(\mathcal{M})} &\le Ct^{-1/4}\|\mathbf{b}\|_{H^1(\mathcal{M})}.
\end{aligned}
$$

Then, using the fact that

$$
\|u\|^2_{L^2(\mathcal{M})}\le C\left(\|\nabla u\|^2_{L^2(\mathcal{M})}+\frac{w_{\min}}{8\beta}\|u\|^2_{L^2(\partial\mathcal{M})}\right),
$$

we have

$$
\|u\|^2_{H^1(\mathcal{M})}\le C\|\mathbf{b}\|_{H^1(\mathcal{M})}\left(\sqrt{t}\|u\|_{H^1(\mathcal{M})}+t\right),
$$

which completes the proof.

Appendix B. Error analysis of the discretization (Theorem 5.6). In this section, we estimate the discretization error introduced by approximating the integrals in (5.8), that is to prove Theorem 5.6. To simplify the notation, we introduce two intermediate operators defined as follows,

$$
\text{(B.1)} \qquad L_{t,h}u(\mathbf{x}) = \frac{1}{t}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{p}_j)(u(\mathbf{x})-u(\mathbf{p}_j))V_j,
$$

$$
\text{(B.2)} \qquad K_{t,h}u(\mathbf{x}) = \frac{1}{t}\sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{p}_j)(u(\mathbf{x})-u(\mathbf{p}_j))V_j+\frac{2}{\beta}\sum_{j=1}^{m}\bar{R}_t(\mathbf{x},\mathbf{p}_j)u(\mathbf{p}_j)A_j.
$$

If $u_{t,h}=I_{\mathbf{f}}(\mathbf{u})$ with $\mathbf{u}$ satisfying Equation (5.2). One can verify that following equation is satisfied,

$$
\text{(B.3)} \qquad K_{t,h}u_{t,h}(\mathbf{x}) = \sum_{j=1}^{n}\bar{R}_t(\mathbf{x},\mathbf{p}_j)f(\mathbf{p}_j)V_j.
$$

The following lemma is needed for proving Theorem 5.6. Its proof is deferred to appendix.

LEMMA B.1. *Suppose* $\mathbf{u} = (u_1, \cdots, u_n)^t$ *satisfies equation* (5.2), *there exist constants* C, T_0, r_0 *only depend on* $\mathcal{M}$ *and* $\partial\mathcal{M}$, *such that*

$$\left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4}\left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} \leq C\|I_{\mathbf{f}}(\mathbf{u})\|_{H^1(\mathcal{M})} + C\sqrt{h}\, t^{3/4}\|f\|_\infty,$$

as long as $t \leq T_0$, $\frac{\sqrt{t}}{\beta} \leq r_0$, $\frac{h}{t^{3/2}} \leq r_0$.

Proof. of Theorem 5.6

Denote

$$u_{t,h}(\mathbf{x}) = I_{\mathbf{f}}(\mathbf{u}) = \frac{1}{w_{t,h}(\mathbf{x})}\left(\sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j)u_jV_j - \frac{2t}{\beta}\sum_{j=1}^m \bar{R}_t(\mathbf{x}, \mathbf{p}_j)u_jA_j + t\sum_{j=1}^n \bar{R}_t(\mathbf{x}, \mathbf{p}_j)f_jV_j\right),$$

where $\mathbf{u} = (u_1, \cdots, u_N)^t$ solves Equation (5.2), $f_j = f(\mathbf{p}_j)$ and $w_{t,h}(\mathbf{x}) = \sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j)V_j$. For convenience, we set

$$\begin{aligned}
a_{t,h}(\mathbf{x}) =& \frac{1}{w_{t,h}(\mathbf{x})}\sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j)u_jV_j,\\
c_{t,h}(\mathbf{x}) =& \frac{t}{w_{t,h}(\mathbf{x})}\sum_{j=1}^n \bar{R}_t(\mathbf{x}, \mathbf{p}_j)f(\mathbf{p}_j)V_j,\\
d_{t,h}(\mathbf{x}) =& -\frac{2t}{\beta w_{t,h}(\mathbf{x})}\sum_{j=1}^m \bar{R}_t(\mathbf{x}, \mathbf{p}_j)u_jA_j.
\end{aligned}$$

Next we upper bound the approximation error $K_t(u_{t,h}) - K_{t,h}(u_{t,h})$. Since $u_{t,h} = a_{t,h} + c_{t,h} + d_{t,h}$, we only need to upper bound the approximation error for $a_{t,h}, c_{t,h}$ and $d_{t,h}$ separately. For $c_{t,h}$,

$$\begin{aligned}
&\left|\left(K_t c_{t,h} - K_{t,h}c_{t,h}\right)(\mathbf{x})\right|\\
\leq& \frac{1}{t}\left|c_{t,h}(\mathbf{x})\right|\left|\int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y})\mathrm{d}\mathbf{y} - \sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j)V_j\right|\\
&+\frac{1}{t}\left|\int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y})c_{t,h}(\mathbf{y})\mathrm{d}\mathbf{y} - \sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j)c_{t,h}(\mathbf{p}_j)V_j\right|\\
&+\frac{2}{\beta}\left|\int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y})c_{t,h}(\mathbf{y})\mathrm{d}\tau_{\mathbf{y}} - \sum_{j=1}^m \bar{R}_t(\mathbf{x}, \mathbf{p}_j)c_{t,h}(\mathbf{p}_j)A_j\right|\\
\leq& \frac{Ch}{t^{3/2}}\left|c_{t,h}(\mathbf{x})\right| + \frac{Ch}{t^{3/2}}\|c_{t,h}\|_\infty + \frac{Ch}{t}\|\nabla c_{t,h}\|_\infty + \frac{Ch}{\beta}\left(t^{-1}\|c_{t,h}\|_\infty + t^{-1/2}\|\nabla c_{t,h}\|_\infty\right)\\
\leq& \frac{Ch}{\sqrt{t}}\left(1 + \frac{\sqrt{t}}{\beta}\right)\|f\|_\infty.
\end{aligned}$$

Now we upper bound $\|K_t a_{t,h} - K_{t,h} a_{t,h}\|_{L_2(\mathcal{M})}$. First, we have

$$\int_{\mathcal{M}} (a_{t,h}(\mathbf{x}))^2 \left| \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y}) \mathrm{d}\mathbf{y} - \sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j) V_j \right|^2 \mathrm{d}\mathbf{x} \tag{B.4}$$

$$\le \frac{Ch^2}{t} \int_{\mathcal{M}} \left(\frac{1}{w_{t,h}(\mathbf{x})} \sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j) u_j V_j \right)^2 \mathrm{d}\mathbf{x}$$

$$\le \frac{Ch^2}{t} \int_{\mathcal{M}} \left(\sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j) u_j^2 V_j \right) \left(\sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j) V_j \right) \mathrm{d}\mathbf{x}$$

$$\le \frac{Ch^2}{t} \left(\sum_{j=1}^n u_j^2 V_j \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{p}_j) \mathrm{d}\mathbf{x} \right) \le \frac{Ch^2}{t} \sum_{j=1}^n u_j^2 V_j.$$

Let

$$K_1 = C_t \int_{\mathcal{M}} \frac{1}{w_{t,h}(\mathbf{y})} R\left(\frac{|\mathbf{x} - \mathbf{y}|^2}{4t} \right) R\left(\frac{|\mathbf{p}_i - \mathbf{y}|^2}{4t} \right) \mathrm{d}\mathbf{y}$$
$$- C_t \sum_{j=1}^n \frac{1}{w_{t,h}(\mathbf{p}_j)} R\left(\frac{|\mathbf{x} - \mathbf{p}_j|^2}{4t} \right) R\left(\frac{|\mathbf{p}_i - \mathbf{p}_j|^2}{4t} \right) V_j.$$

We have $|K_1| < \frac{Ch}{t^{1/2}}$ for some constant C independent of t. In addition, notice that only when $|\mathbf{x} - \mathbf{p}_i|^2 \le 16t$ is $K_1 \ne 0$, which implies

$$|K_1| \le \frac{1}{\delta_0} |K_1| R\left(\frac{|\mathbf{x} - \mathbf{p}_i|^2}{32t} \right).$$

Then we have

$$\int_{\mathcal{M}} \left| \int_{\mathcal{M}} R_t(\mathbf{x}, \mathbf{y}) a_{t,h}(\mathbf{y}) \mathrm{d}\mathbf{y} - \sum_{j=1}^n R_t(\mathbf{x}, \mathbf{p}_j) a_{t,h}(\mathbf{p}_j) V_j \right|^2 \mathrm{d}\mathbf{x} \tag{B.5}$$

$$= \int_{\mathcal{M}} \left(\sum_{i=1}^n C_t u_i V_i K_1 \right)^2 \mathrm{d}\mathbf{x}$$

$$\le \frac{Ch^2}{t} \int_{\mathcal{M}} \left(\sum_{i=1}^n C_t |u_i| V_i R\left(\frac{|\mathbf{x} - \mathbf{p}_i|^2}{32t} \right) \right)^2 \mathrm{d}\mathbf{x}$$

$$\le \frac{Ch^2}{t} \int_{\mathcal{M}} \left(\sum_{i=1}^n C_t R\left(\frac{|\mathbf{x} - \mathbf{p}_i|^2}{32t} \right) u_i^2 V_i \right) \left(\sum_{i=1}^n C_t R\left(\frac{|\mathbf{x} - \mathbf{p}_i|^2}{32t} \right) V_i \right) \mathrm{d}\mathbf{x}$$

$$\le \frac{Ch^2}{t} \sum_{i=1}^n \left(\int_{\mathcal{M}} C_t R\left(\frac{|\mathbf{x} - \mathbf{p}_i|^2}{32t} \right) \mathrm{d}\mathbf{x} \left(u_i^2 V_i \right) \right) \le \frac{Ch^2}{t} \left(\sum_{i=1}^n u_i^2 V_i \right).$$

Let

$$K_2 = C_t \int_{\partial\mathcal{M}} \frac{1}{w_{t,h}(\mathbf{y})} \bar{R}\left(\frac{|\mathbf{x} - \mathbf{y}|^2}{4t} \right) R\left(\frac{|\mathbf{p}_i - \mathbf{y}|^2}{4t} \right) \mathrm{d}\tau_{\mathbf{y}}$$
$$- C_t \sum_{j=1}^m \frac{1}{w_{t,h}(\mathbf{p}_j)} \bar{R}\left(\frac{|\mathbf{x} - \mathbf{p}_j|^2}{4t} \right) R\left(\frac{|\mathbf{p}_i - \mathbf{p}_j|^2}{4t} \right) A_j.$$

We have $|K_2| < \frac{Ch}{t}$ for some constant C independent of t. In addition, notice that only when $|\mathbf{x}-\mathbf{p}_i|^2 \le 16t$ is $K_2 \neq 0$, which implies

$$|K_2| \le \frac{1}{\delta_0}|K_2| R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right).$$

Then

$$\begin{aligned}
&\int_{\mathcal{M}} \left| \int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y}) a_{t,h}(\mathbf{y}) \mathrm{d}\tau_{\mathbf{y}} - \sum_{j=1}^{m} \bar{R}_t(\mathbf{x},\mathbf{p}_j) a_{t,h}(\mathbf{p}_j) A_j \right|^2 \mathrm{d}\mathbf{x} \quad \text{(B.6)}\\
&= \int_{\mathcal{M}} \left(\sum_{i=1}^{n} C_t u_i V_i K_2 \right)^2 \mathrm{d}\mathbf{x}\\
&\le \frac{Ch^2}{t^2} \int_{\mathcal{M}} \left(\sum_{i=1}^{n} C_t |u_i| V_i R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) \right)^2 \mathrm{d}\mathbf{x}\\
&\le \frac{Ch^2}{t^2} \int_{\mathcal{M}} \left(\sum_{i=1}^{n} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) u_i^2 V_i \right) \left(\sum_{i=1}^{n} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) V_i \right) \mathrm{d}\mathbf{x}\\
&\le \frac{Ch^2}{t^2} \sum_{i=1}^{n} \left(\int_{\mathcal{M}} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) \mathrm{d}\mathbf{x} \left(u_i^2 V_i\right) \right) \le \frac{Ch^2}{t^2} \left(\sum_{i=1}^{n} u_i^2 V_i \right).
\end{aligned}$$

Combining Equation (B.4), (B.5) and (B.6),

$$\|K_t a_{t,h} - K_{t,h} a_{t,h}\|_{L^2(\mathcal{M})} \le \frac{Ch}{t^{3/2}} \left(1 + \frac{\sqrt{t}}{\beta}\right) \left(\sum_{i=1}^{n} u_i^2 V_i \right)^{1/2}$$

Now we upper bound $\|K_t d_{t,h} - K_{t,h} d_{t,h}\|_{L_2}$. We have

$$\begin{aligned}
&\int_{\mathcal{M}} (d_{t,h}(\mathbf{x}))^2 \left| \int_{\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y}) \mathrm{d}\tau_{\mathbf{y}} - \sum_{j=1}^{n} \bar{R}_t(\mathbf{x},\mathbf{p}_j) V_j \right|^2 \mathrm{d}\mathbf{x} \quad \text{(B.7)}\\
&\le \frac{Ch^2}{t^2} \int_{\mathcal{M}} (d_{t,h}(\mathbf{x}))^2 \mathrm{d}\mathbf{x}\\
&\le \frac{Ch^2 t}{\beta^2} \int_{\mathcal{M}} \left(\frac{1}{w_{t,h}(\mathbf{x})} \sum_{j=1}^{m} \bar{R}_t(\mathbf{x},\mathbf{p}_j) u_j A_j \right)^2 \mathrm{d}\mathbf{x}\\
&\le \frac{Ch^2 t}{\beta^2} \int_{\mathcal{M}} \left(\sum_{j=1}^{m} \bar{R}_t(\mathbf{x},\mathbf{p}_j) u_j^2 A_j \right) \left(\sum_{j=1}^{m} \bar{R}_t(\mathbf{x},\mathbf{p}_j) A_j \right) \mathrm{d}\mathbf{x}\\
&\le \frac{Ch^2\sqrt{t}}{\beta^2} \left(\sum_{j=1}^{m} u_j^2 A_j \int_{\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{p}_j) \mathrm{d}\mathbf{x} \right) \le \frac{Ch^2\sqrt{t}}{\beta^2} \sum_{j=1}^{m} u_j^2 A_j.
\end{aligned}$$

Let

$$\begin{aligned}
K_3 = {} & C_t \int_{\mathcal{M}} \frac{1}{w_{t,h}(\mathbf{y})} R\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) \bar{R}\left(\frac{|\mathbf{p}_i-\mathbf{y}|^2}{4t}\right) \mathrm{d}\mathbf{y}\\
& - C_t \sum_{j=1}^{n} \frac{1}{w_{t,h}(\mathbf{p}_j)} R\left(\frac{|\mathbf{x}-\mathbf{p}_j|^2}{4t}\right) \bar{R}\left(\frac{|\mathbf{p}_i-\mathbf{p}_j|^2}{4t}\right) V_j.
\end{aligned}$$

We have $|K_3| < \frac{Ch}{t^{1/2}}$ for some constant K_3 independent of t. In addition, notice that only when $|\mathbf{x}-\mathbf{p}_i|^2 \leq 16t$ is $K_3 \neq 0$, which implies

$$|K_3| \leq \frac{1}{\delta_0}|C|R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{4t}\right).$$

Then we have

$$\begin{aligned}
\text{(B.8)} \qquad & \int_{\mathcal{M}} \left| \int_{\mathcal{M}} R_t(\mathbf{x},\mathbf{y}) d_{t,h}(\mathbf{y}) \mathrm{d}\mathbf{y} - \sum_{j=1}^{n} R_t(\mathbf{x},\mathbf{p}_j) d_{t,h}(\mathbf{p}_j) V_j \right|^2 \mathrm{d}\mathbf{x} \\
&= \frac{4t^2}{\beta^2} \int_{\mathcal{M}} \left(\sum_{i=1}^{m} C_t u_i A_i K_3 \right)^2 \mathrm{d}\mathbf{x} \\
&\leq \frac{Ch^2 t}{\beta^2} \int_{\mathcal{M}} \left(\sum_{i=1}^{m} C_t |u_i| A_i R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) \right)^2 \mathrm{d}\mathbf{x} \\
&\leq \frac{Ch^2 t}{\beta^2} \int_{\mathcal{M}} \left(\sum_{i=1}^{m} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) u_i^2 A_i \right) \left(\sum_{i=1}^{m} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) A_i \right) \mathrm{d}\mathbf{x} \\
&\leq \frac{Ch^2\sqrt{t}}{\beta^2} \sum_{i=1}^{m} \left(\int_{\mathcal{M}} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) \mathrm{d}\mathbf{x} \left(u_i^2 A_i\right) \right) \leq \frac{Ch^2\sqrt{t}}{\beta^2} \left(\sum_{i=1}^{m} u_i^2 A_i \right).
\end{aligned}$$

Let

$$\begin{aligned}
K_4 = {} & C_t \int_{\partial\mathcal{M}} \frac{1}{w_{t,h}(\mathbf{y})} \bar{R}\left(\frac{|\mathbf{x}-\mathbf{y}|^2}{4t}\right) \bar{R}\left(\frac{|\mathbf{p}_i-\mathbf{y}|^2}{4t}\right) \mathrm{d}\tau_{\mathbf{y}} \\
& - C_t \sum_{j=1}^{m} \frac{1}{w_{t,h}(\mathbf{p}_j)} \bar{R}\left(\frac{|\mathbf{x}-\mathbf{p}_j|^2}{4t}\right) \bar{R}\left(\frac{|\mathbf{p}_i-\mathbf{p}_j|^2}{4t}\right) A_j.
\end{aligned}$$

We have $|K_4| < \frac{Ch}{t}$ for some constant C independent of t. In addition, notice that only when $|\mathbf{x}-\mathbf{p}_i|^2 \leq 16t$ is $K_4 \neq 0$, which implies

$$|K_4| \leq \frac{1}{\delta_0}|K_4|R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right).$$

and

$$\begin{aligned}
\text{(B.9)} \qquad & \int_{\mathcal{M}} \left| \int_{\partial\mathcal{M}} \bar{R}_t(\mathbf{x},\mathbf{y}) d_{t,h}(\mathbf{y}) \mathrm{d}\tau_{\mathbf{y}} - \sum_{j} \bar{R}_t(\mathbf{x},\mathbf{p}_j) d_{t,h}(\mathbf{p}_j) A_j \right|^2 \mathrm{d}\mathbf{x} \\
&= \frac{4t^2}{\beta^2} \int_{\mathcal{M}} \left(\sum_{i=1}^{m} C_t u_i A_i K_4 \right)^2 \mathrm{d}\mathbf{x} \\
&\leq \frac{Ch^2}{\beta^2} \int_{\mathcal{M}} \left(\sum_{i=1}^{m} C_t |u_i| A_i R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) \right)^2 \mathrm{d}\mathbf{x} \\
&\leq \frac{Ch^2}{\beta^2} \int_{\mathcal{M}} \left(\sum_{i=1}^{m} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) u_i^2 A_i \right) \left(\sum_{i=1}^{m} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) A_i \right) \mathrm{d}\mathbf{x} \\
&\leq \frac{Ch^2}{\beta^2\sqrt{t}} \sum_{i=1}^{m} \left(\int_{\mathcal{M}} C_t R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{32t}\right) \mathrm{d}\mathbf{x} \left(u_i^2 A_i\right) \right) \leq \frac{Ch^2}{\beta^2\sqrt{t}} \left(\sum_{i=1}^{m} u_i^2 A_i \right).
\end{aligned}$$

Combining Equation (B.7), (B.8) and (B.9),

$$\|K_t d_{t,h} - K_{t,h} d_{t,h}\|_{L^2(\mathcal{M})} \le \frac{Ch}{\beta t^{3/4}} \left(1 + \frac{\sqrt{t}}{\beta}\right) \left(\sum_{i=1}^m u_i^2 A_i\right)^{1/2}$$

Now assembling the parts together, we have the following upper bound.

$$\begin{aligned} &\|K_t u_{t,h} - K_{t,h} u_{t,h}\|_{L^2(\mathcal{M})} \\ \le& \frac{Ch}{t^{3/2}} \left(\|g\|_\infty + t\|f\|_\infty + \left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4} \left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} \right). \end{aligned} \tag{B.10}$$

At the same time, since u_t solves $K_t u_t = \int_{\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) \mathrm{d}\mathbf{y}$ and $u_{t,h}$ solves (B.3) respectively, we have

$$\begin{aligned} &\|K_t(u_t) - K_{t,h}(u_{t,h})\|_{L^2(\mathcal{M})} \\ =& \left(\int_{\mathcal{M}} \left((K_t u_t - K_{t,h} u_{t,h}) (\mathbf{x}) \right)^2 \mathrm{d}\mathbf{x} \right)^{1/2} \\ \le& \left(\int_{\mathcal{M}} \left(\int_{\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) - \sum_{j=1}^n \bar{R}_t(\mathbf{x}, \mathbf{p}_j) f(\mathbf{p}_j) V_j \right)^2 \mathrm{d}\mathbf{x} \right)^{1/2} \\ \le& \frac{Ch}{t^{1/2}} \|f\|_\infty. \end{aligned} \tag{B.11}$$

From Equation (B.10) and (B.11), we get

$$\|K_t u_t - L_t u_{t,h}\|_{L^2(\mathcal{M})} \le \frac{Ch}{t^{3/2}} \left(\left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4} \left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} + t\|f\|_\infty \right). \tag{B.12}$$

Using the similar techniques, we can get the upper bound of $\|\nabla(K_t u_t - L_t u_{t,h})\|_{L_2(\mathcal{M})}$ as following.

$$\|\nabla \left(K_t u_t - L_t u_{t,h}\right) \|_{L^2(\mathcal{M})} \le \frac{Ch}{t^2} \left(t\|f\|_{C^1(\mathcal{M})} + \left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4} \left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} \right). \tag{B.13}$$

In the remaining of the proof, we only need to get a prior estimate of $\left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4} \left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2}$. First, using the estimate (B.12) and (B.13) and the Theorem 5.4, we have

$$\begin{aligned} \|u_{t,h}\|_{H^1(\mathcal{M})} \le& \frac{Ch}{t^{3/2}} \left(\left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4} \left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} + t\|f\|_\infty \right) \\ &+ C\|K_t u_t\|_{L^2(\mathcal{M})} + Ct^{3/4} \|K_t u_t\|_{H^1(\mathcal{M})}. \end{aligned} \tag{B.14}$$

Using the relation that $K_t u_t = -\int_{\mathcal{M}} \bar{R}_t(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) \mathbf{y}$, it is easy to get that

$$\|K_t u_t\|_{L^2(\mathcal{M})} \le C\|f\|_\infty, \tag{B.15}$$

$$\|\nabla(K_t u_t)\|_{L^2(\mathcal{M})} \le \frac{C}{t^{1/2}} \|f\|_\infty. \tag{B.16}$$

Substituting above estimates in (B.14), we have

$$\|u_{t,h}\|_{H^1(\mathcal{M})} \le \frac{Ch}{t^{3/2}}\left(\left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4}\left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} + t\|f\|_\infty\right) + C\|f\|_\infty.$$

Using Lemma B.1, we have

$$\begin{aligned}
&\left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4}\left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} \\
&\le C\|u_{t,h}\|_{H^1(\mathcal{M})} + C\sqrt{h}\left(t^{3/4}\|f\|_\infty + \|g\|_\infty\right) \\
&\le \frac{Ch}{t^{3/2}}\left(t\|f\|_\infty + \left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4}\left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2}\right) \\
&\quad +C\|f\|_\infty + C\sqrt{h}\, t^{3/4}\|f\|_\infty
\end{aligned} \tag{B.17}$$

Using the assumption that $\frac{h}{t^{3/2}}$ is small enough such that $\frac{Ch}{t^{3/2}} \le \frac{1}{2}$, we have

$$\left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4}\left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} \le C\|f\|_\infty \tag{B.18}$$

Then the proof is complete by substituting above estimate (B.18) in (B.12) and (B.13). □

Appendix C. Proof of Lemma B.1.

Proof. First, denote

$$u_{t,h}(\mathbf{x}) = I_{\mathbf{f}}(\mathbf{u}) = \frac{1}{w_{t,h}(\mathbf{x})}\left(\sum_{j=1}^n R_t(\mathbf{x},\mathbf{p}_j)u_j V_j - \frac{2t}{\beta}\sum_{j=1}^m \bar{R}_t(\mathbf{x},\mathbf{p}_j)u_j A_j + t\sum_{j=1}^n \bar{R}_t(\mathbf{x},\mathbf{p}_j)f_j V_j\right),$$

where $f_j = f(\mathbf{p}_j)$ and $w_{t,h}(\mathbf{x}) = \sum_{j=1}^n R_t(\mathbf{x},\mathbf{p}_j)V_j$ and $\mathbf{u} = (u_1, \cdots, u_n)$ solves (5.2). Let

$$\begin{aligned}
v_1(\mathbf{x}) &= \frac{1}{w_{t,h}(\mathbf{x})}\sum_{j=1}^n R_t(\mathbf{x},\mathbf{p}_j)u_j V_j, \text{ and} \\
v_2(\mathbf{x}) &= -\frac{2t}{\beta w_{t,h}(\mathbf{x})}\sum_{j=1}^m \bar{R}_t(\mathbf{x},\mathbf{p}_j)u_j A_j, \text{ and} \\
v_3(\mathbf{x}) &= \frac{t}{w_{t,h}(\mathbf{x})}\sum_{j=1}^n \bar{R}_t(\mathbf{x},\mathbf{p}_j)f_j V_j,
\end{aligned}$$

and then $u_{t,h} = v_1 + v_2 + v_3$ and

$$\begin{aligned}
\left|\|u_{t,h}\|^2_{L^2(\mathcal{M})} - \sum_{j=1}^n u_j^2 V_j\right| &= \left|\sum_{m,m'=1}^3\left(\int_{\mathcal{M}} v_m(\mathbf{x})v_{m'}(\mathbf{x})\mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^n v_m(\mathbf{x}_j)v_{m'}(\mathbf{x}_j)V_j\right)\right| \\
&\le \sum_{m,m'=1}^3\left|\int_{\mathcal{M}} v_m(\mathbf{x})v_{m'}(\mathbf{x})\mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^n v_m(\mathbf{x}_j)v_{m'}(\mathbf{x}_j)V_j\right|.
\end{aligned}$$

We now estimate these six terms in the above summation one by one. First, we consider the term with $m = m' = 1$. Denote

$$A = \int_{\mathcal{M}} \frac{C_t}{w_{t,h}^2(\mathbf{x})} R\left(\frac{|\mathbf{x}-\mathbf{p}_i|^2}{4t}\right) R\left(\frac{|\mathbf{x}-\mathbf{p}_l|^2}{4t}\right) \mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^{n} \frac{C_t}{w_{t,h}^2(\mathbf{p}_j)} R\left(\frac{|\mathbf{p}_j-\mathbf{p}_i|^2}{4t}\right) R\left(\frac{|\mathbf{p}_j-\mathbf{p}_l|^2}{4t}\right) V_j,$$

and then $|A| \leq \frac{Ch}{t^{1/2}}$. At the same time, notice that only when $|\mathbf{p}_i - \mathbf{p}_l|^2 < 16t$ is $A \neq 0$. Thus we have

$$|A| \leq \frac{1}{\delta_0}|A|R(\frac{|\mathbf{p}_i-\mathbf{p}_l|^2}{32t}),$$

and

$$\begin{aligned}
&\left|\int_{\mathcal{M}} v_1^2(\mathbf{x})\mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^{n} v_1^2(\mathbf{p}_j)V_j\right| \\
\leq& \sum_{i,l=1}^{n} |C_t u_i u_l V_i V_l||A| \\
\leq& \frac{Ch}{t^{1/2}} \sum_{i,l=1}^{n} \left|C_t R\left(\frac{|\mathbf{p}_i-\mathbf{p}_l|^2}{32t}\right) u_i u_l V_i V_l\right| \\
\leq& \frac{Ch}{t^{1/2}} \sum_{i=1}^{n} \left(\sum_{l=1}^{n} C_t R\left(\frac{|\mathbf{p}_i-\mathbf{p}_l|^2}{32t}\right) V_l\right)^{1/2} \left(\sum_{l=1}^{n} C_t R\left(\frac{|\mathbf{p}_i-\mathbf{p}_l|^2}{32t}\right) u_l^2 V_l\right)^{1/2} u_i V_i \\
\leq& \frac{Ch}{t^{1/2}} \left(\sum_{i=1}^{n}\sum_{l=1}^{n} C_t R\left(\frac{|\mathbf{p}_i-\mathbf{p}_l|^2}{32t}\right) u_l^2 V_l V_i\right)^{1/2} \left(\sum_{i=1}^{n} u_i^2 V_i\right)^{1/2}. \\
=& \frac{Ch}{t^{1/2}} \left(\sum_{l=1}^{n} u_l^2 V_l \sum_{i=1}^{n} C_t R\left(\frac{|\mathbf{p}_i-\mathbf{p}_l|^2}{32t}\right) V_i\right)^{1/2} \left(\sum_{i=1}^{n} u_i^2 V_i\right)^{1/2} \\
\leq& \frac{Ch}{t^{1/2}} \sum_{i=1}^{n} u_i^2 V_i.
\end{aligned}$$

Using a similar argument, we can obtain the following estimates for the remaining terms,

$$\left|\int_{\mathcal{M}} v_1(\mathbf{x})v_2(\mathbf{x})\mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^{n} v_1(\mathbf{p}_j)v_2(\mathbf{p}_j)V_j\right| \leq \frac{Cht^{1/4}}{\beta} \left(\sum_{i=1}^{n} u_i^2 V_i\right)^{1/2} \left(\sum_{l=1}^{m} u_l^2 A_l\right)^{1/2},$$

$$\left|\int_{\mathcal{M}} v_1(\mathbf{x})v_3(\mathbf{x})\mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^{n} v_1(\mathbf{p}_j)v_3(\mathbf{p}_j)V_j\right| \leq Cht^{1/2} \left(\sum_{i=1}^{n} u_i^2 V_i\right)^{1/2} \left(\sum_{j=1}^{n} f_j^2 V_j\right)^{1/2},$$

$$\left|\int_{\mathcal{M}} v_2^2(\mathbf{x})\mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^n v_2^2(\mathbf{p}_j)V_j\right| \le \frac{Cht}{\beta^2}\sum_{l=1}^m u_l^2 A_l, \text{ and}$$

$$\left|\int_{\mathcal{M}} v_2(\mathbf{x})v_3(\mathbf{x})\mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^n v_2(\mathbf{p}_j)v_3(\mathbf{p}_j)V_j\right| \le \frac{Cht^{5/4}}{\beta}\left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2}\left(\sum_{j=1}^n f_j^2 V_j\right)^{1/2}, \text{ and}$$

$$\left|\int_{\mathcal{M}} v_3^2(\mathbf{x})\mathrm{d}\mu_{\mathbf{x}} - \sum_{j=1}^n v_3^2(\mathbf{p}_j)V_j\right| \le Cht^{3/2}\sum_{j=1}^n f_j^2 V_j.$$

Assembling all the above estimates together, we obtain

$$\left|\|u_{t,h}\|_{L^2(\mathcal{M})}^2 - \sum_{i=1}^n u_i^2 V_i\right| \le \frac{Ch}{t^{1/2}}\left(\sum_{i=1}^n u_i^2 V_i + t^{1/2}\sum_{l=1}^m u_l^2 A_l + t^2\|f\|_\infty^2\right).$$

Similarly, we have

$$\left|\|u_{t,h}\|_{L^2(\partial\mathcal{M})}^2 - \sum_{l=1}^m u_l^2 A_l\right| \le \frac{Ch}{t}\left(\sum_{i=1}^n u_i^2 V_i + t^{1/2}\sum_{l=1}^m u_l^2 A_l + t^2\|f\|_\infty^2\right).$$

Using the assumption that $\frac{h}{t^{1/2}}$ is small enough such that $\frac{Ch}{t^{1/2}} \le \frac{1}{2}$, we obtain

$$\begin{aligned}\sum_{i=1}^n u_i^2 V_i + t^{1/2}\sum_{l=1}^m u_l^2 A_l &\le 2\left(\|u_{t,h}\|_{L^2(\mathcal{M})}^2 + t^{1/2}\|u_{t,h}\|_{L^2(\partial\mathcal{M})}^2\right) + Ch\left(t^{3/2}\|f\|_\infty^2\right)\\ &\le C\|u_{t,h}\|_{H^1(\mathcal{M})}^2 + Cht^{3/2}\|f\|_\infty^2,\end{aligned}$$

which implies that

$$\left(\sum_{i=1}^n u_i^2 V_i\right)^{1/2} + t^{1/4}\left(\sum_{l=1}^m u_l^2 A_l\right)^{1/2} \le C\|u_{t,h}\|_{H^1(\mathcal{M})} + C\sqrt{h}\,t^{3/4}\|f\|_\infty.$$

□

REFERENCES

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

[2] M. Belkin and P. Niyogi. Convergence of laplacian eigenmaps. *preprint, short version NIPS 2008*, 2008.

[3] A. Buades, B. Coll, and J.-M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Model. Simul.*, 4:490–530, 2005.

[4] C. J. Burges, Y. LeCun, and C. Cortes. Mnist database.

[5] F. R. K. Chung. *Spectral Graph Theory.* American Mathematical Society, 1997.

[6] P. G. Doyle and J. L. Snell. *Random Walks and Electric Networks.* Mathematical Association of America, Washington, DC, 1984.

[7] Q. Du, M. Gunzburger, R. B. Lehoucq, and K. Zhou. Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Review*, 54:667–696, 2012.

[8] Q. Du, M. Gunzburger, R. B. Lehoucq, and K. Zhou. Analysis and approximation of nonlocal diffusion problems with volume constraints. *SIAM Review*, 54:667–696, 2012.

[9] Q. Du, T. Li, and X. Zhao. A convergent adaptive finite element algorithm for nonlocal diffusion and peridynamic models. *SIAM J. Numer. Anal.*, 51:1211–1234, 2013.

[10] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Model. Simul.*, 7:1005–1028, 2008.

[11] 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, 2005.

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

[13] 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.

[14] Z. Li, Z. Shi, and J. Sun. Point integral method for solving poisson-type equations on manifolds from point clouds with convergence guarantees. *arXiv:1409.2623*.

[15] S. Osher, Z. Shi, and W. Zhu. Low dimensional manifold model for image processing. *Technical report, UCLA, CAM-report 16-04*.

[16] Z. Shi. Enforce the dirichlet boundary condition by volume constraint in point integral method. *arXiv:1506.02343*.

[17] Z. Shi and J. Sun. Convergence of laplacian spectra from point clouds. *arXiv:1506.01788*.

[18] Z. Shi and J. Sun. Convergence of the point integral method for the poisson equation on point cloud. *arXiv:1403.2141*.

[19] A. Singer. From graph to manifold Laplacian: The convergence rate. *Applied and Computational Harmonic Analysis*, pages 128–134, 2006.

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

[21] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf. Learning with local and global consistency. In *Advances in Neural Information Processing Systems 16 [Neural Information Processing Systems, NIPS 2003, December 8-13, 2003, Vancouver and Whistler, British Columbia, Canada]*, pages 321–328, 2003.

[22] X. Zhu. *Semi-supervised Learning with Graphs*. PhD thesis, Pittsburgh, PA, USA, 2005. AAI3179046.

[23] X. Zhu, Z. Ghahramani, and J. D. Lafferty. Semi-supervised learning using gaussian fields and harmonic functions. In *Machine Learning, Proceedings of the Twentieth International Conference ICML 2003), August 21-24, 2003, Washington, DC, USA*, pages 912–919, 2003.