Sparse Reconstruction of the Generalized Parton Distribution

Yuesheng Xu
Old Dominion University, Norfolk, Virginia, USA
Syracuse University, Syracuse, New York, USA (Emeritus)
Presentation at A.I. for Nuclear Physics Workshop, March 4-6, 2020

Joint work with Jie Jiang (A visiting graduate student from Sun Yat-sen University),
Charles Hyde (Department of Physics, ODU)
This work is supported by NSF and JLab.

Overview: Inverse Problems

- Most of data science problems can be formulated as an inverse problem.
- Learning a function (an image or a signal) from measured data is an inverse problem.
- A major challenge of solving an inverse problem is its instability: A small perturbation in data may lead to large changes in solution.
- Most measurement data contain noise. Hence, it is inevitable to for us deal with the instability when solving an inverse problem.
- Inverse problems have to be solved by regularization.
- We shall illustrate the idea of solving inverse problems by an example: reconstruction of the Generalized Parton Distribution.

The Generalized Parton Distribution

- The parton model, proposed by Richard Feynman (1969)[1], describes the inner structure of hadrons, such as protons and neutrons. The aim of the parton model is to analyze high-energy hadron collision.

- The Generalized Parton Distribution (GPD) was proposed[2] to catch kinetic characteristics of partons.
  - GPD is to represent the transverse spatial image of quarks and gluons, as a function of their longitudinal momentum fraction in the proton and the neutron.
  - GPD encodes correlations between the transverse position and longitudinal momentum of partons inside nucleon.

[1] Richard P Feynman. Very high-energy collisions of hadrons. *Physical Review Letters* 23.24 (1969), p. 1415.

[2] Dieter Mueller et al. Wave Functions, Evolution Equations and Evolution Kernels from Light-Ray Operators of QCD. *Fortschritte der Physik/Progress of Physics* 42.2 (1994), pp. 101–141.

The Goal of This Study: Reconstruction of the Generalized Parton Distribution

- Consider the system of equations and constraints for GPD.
- Formulate the construction of GPD as an inverse problem.
- Use orthogonal polynomials such as
  - Chebeshev polynomials
  - Legendre polynomials

  to discretize the equations and constraints.
- Treat the resulting ill-posed problem by sparse regularization.
- Solve the non-smooth optimization problem using the fixed-point proximity algorithm.

Deeply Virtual Compton Scattering (DVCS)

In a nuclear physics experiment, high energy electron e is used to hit the quark h in the proton. It generates an outcoming electron e, a quark h and a high-energy photon γ. This can be described as[3] [4]

$$e(k) + h(P_1) \to e(k') + h(P_2) + \gamma(q_2).$$

1. e – electron, h – quark, γ – photon.
2. k – the four-momentum of the initial electron.
3. P_1 – the four-momentum of the initial nucleon.
4. k' – the four-momentum of the final electron.
5. P_2 – the four-momentum of the final nucleon.
6. q_2 – the four-momentum of the final photon.

[3] Xiangdong Ji. Off-forward parton distributions. *Journal of Physics G: Nuclear and Particle Physics* 24.7 (1998), p. 1181.

[4] Xiangdong Ji. Gauge-invariant decomposition of nucleon spin. *Physical Review Letters* 78.4 (1997), p. 610.

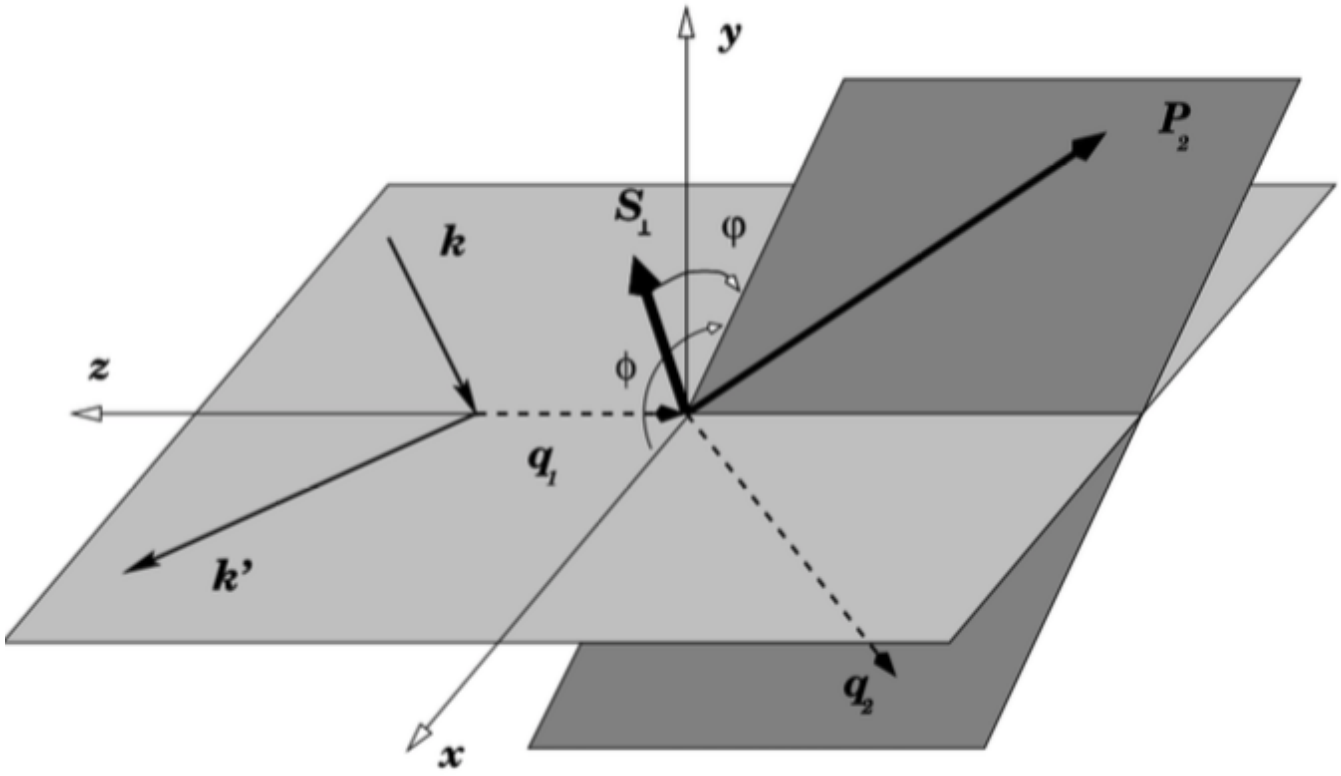


Figure 1: Scattering Figure[5]

$$e(k) + h(P_1) \to e(k') + h(P_2) + \gamma(q_2).$$

[5] Andrei V Belitsky, Dieter Mueller, and A Kirchner. Theory of deeply virtual Compton scattering on the nucleon. *Nuclear Physics B* 629.1-3 (2002), pp. 323–392.

Crucial Physics Variables

1. $q_1 := k - k'$, the four-momentum of the virtual photon.
2. $\Delta := P_2 - P_1 = q_1 - q_2$, overall four-momentum transfer.
3. $Q^2 := -q_1^2 = -(k - k')^2$, the virtuality of the absorbed virtual photon.
4. $x \in [-1, 1]$, longitudinal momentum fraction carried by constituent parton.
5. $\xi := (P_1 - P_2)^+/(P_1 + P_2)^+ \in [-1, 1]$, longitudinal momentum fraction transferred to the parton.
6. $t := \Delta^2$, invariant quantity measuring four-momentum transfer.

- Suppose the interval of the invariant quantity t which we are interested is bounded. We can then use an affine transformation to convert the domain of t into the bounded interval $(-1, 1)$.

Singular Integral Equations for GPDs

The GPD functions $[H, E, \tilde{H}, \tilde{E}]$ are determined by the singular integral equations from the Compton form factors (CFFs) $[\mathcal{H}, \mathcal{E}, \tilde{\mathcal{H}}, \tilde{\mathcal{E}}]$: For $\xi, t \in (-1, 1)$,

$$\mathcal{H}(\xi, t) = \int_{-1}^{1} \left[\frac{1}{x-\xi} + \frac{1}{x+\xi}\right] H(x, \xi, t)\,\mathrm{d}x + i\pi[H(-\xi, \xi, t) - H(\xi, \xi, t)],$$

$$\mathcal{E}(\xi, t) = \int_{-1}^{1} \left[\frac{1}{x-\xi} + \frac{1}{x+\xi}\right] E(x, \xi, t)\,\mathrm{d}x + i\pi[E(-\xi, \xi, t) - E(\xi, \xi, t)],$$

$$\tilde{\mathcal{H}}(\xi, t) = \int_{-1}^{1} \left[\frac{1}{x-\xi} - \frac{1}{x+\xi}\right] \tilde{H}(x, \xi, t)\,\mathrm{d}x - i\pi[\tilde{H}(-\xi, \xi, t) + \tilde{H}(\xi, \xi, t)],$$

$$\tilde{\mathcal{E}}(\xi, t) = \int_{-1}^{1} \left[\frac{1}{x-\xi} - \frac{1}{x+\xi}\right] \tilde{E}(x, \xi, t)\,\mathrm{d}x - i\pi[\tilde{E}(-\xi, \xi, t) + \tilde{E}(\xi, \xi, t)].$$

The CFFs $[\mathcal{H}, \mathcal{E}, \tilde{\mathcal{H}}, \tilde{\mathcal{E}}]$ are obtained from experiments[6].

[6] Andrei V Belitsky, Dieter Mueller, and A Kirchner. Theory of deeply virtual Compton scattering on the nucleon. *Nuclear Physics B* 629.1-3 (2002), pp. 323–392.

Twist-Two Quark GPDs

Let n denote the light-like vector and $P := P_1 + P_2$.

The twist-two quark GPDs are given by[7]

$$\langle P_2|\hat{\psi}(-\kappa n)\gamma\cdot n\psi(\kappa n)|P_1\rangle = \int_{-1}^{1}(h\cdot nH(x,\xi,t)+e\cdot nE(x,\xi,t))e^{-ix\kappa(P\cdot n)}\,\mathrm{d}x,$$

$$\langle P_2|\hat{\psi}(-\kappa n)\gamma\cdot n\gamma_5\psi(\kappa n)|P_1\rangle = \int_{-1}^{1}(\tilde{h}\cdot n\tilde{H}(x,\xi,t)+\tilde{e}\cdot n\tilde{E}(x,\xi,t))e^{-ix\kappa(P\cdot n)}\,\mathrm{d}x.$$

[7] Andrei V Belitsky, Dieter Mueller, and A Kirchner. Theory of deeply virtual Compton scattering on the nucleon. *Nuclear Physics B* 629.1-3 (2002), pp. 323–392.

Reformulation of the Singular Integral Equations

We reformulate the singular equation in terms of the Cauchy singular kernel:
For $\xi, t \in (-1, 1)$,

$$\mathcal{H}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi} \left[H(x,\xi,t) - H(-x,\xi,t)\right] \mathrm{d}x + i\pi[H(-\xi,\xi,t) - H(\xi,\xi,t)],$$

$$\mathcal{E}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi} \left[E(x,\xi,t) - E(-x,\xi,t)\right] \mathrm{d}x + i\pi[E(-\xi,\xi,t) - E(\xi,\xi,t)],$$

$$\tilde{\mathcal{H}}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi} \left[\tilde{H}(x,\xi,t) + \tilde{H}(-x,\xi,t)\right] \mathrm{d}x - i\pi[\tilde{H}(-\xi,\xi,t) + \tilde{H}(\xi,\xi,t)],$$

$$\tilde{\mathcal{E}}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi} \left[\tilde{E}(x,\xi,t) + \tilde{E}(-x,\xi,t)\right] \mathrm{d}x - i\pi[\tilde{E}(-\xi,\xi,t) + \tilde{E}(\xi,\xi,t)].$$

Initial Constraints

Let q be the quark distribution, $\tilde{q}$ the anti-quark distribution, Δq the quark helicity distribution and $\Delta\tilde{q}$ the anti-quark helicity distribution.
We have the initial constraints

$$H(x,0,0) = q(x), \quad 0 \leqslant x < 1,$$

$$\tilde{H}(x,0,0) = \Delta q(x), \quad 0 \leqslant x < 1,$$

$$H(x,0,0) = -\tilde{q}(-x), \quad -1 < x < 0,$$

$$\tilde{H}(x,0,0) = \Delta q(-x), \quad -1 < x < 0.$$

Initial Constraints: An Illustration

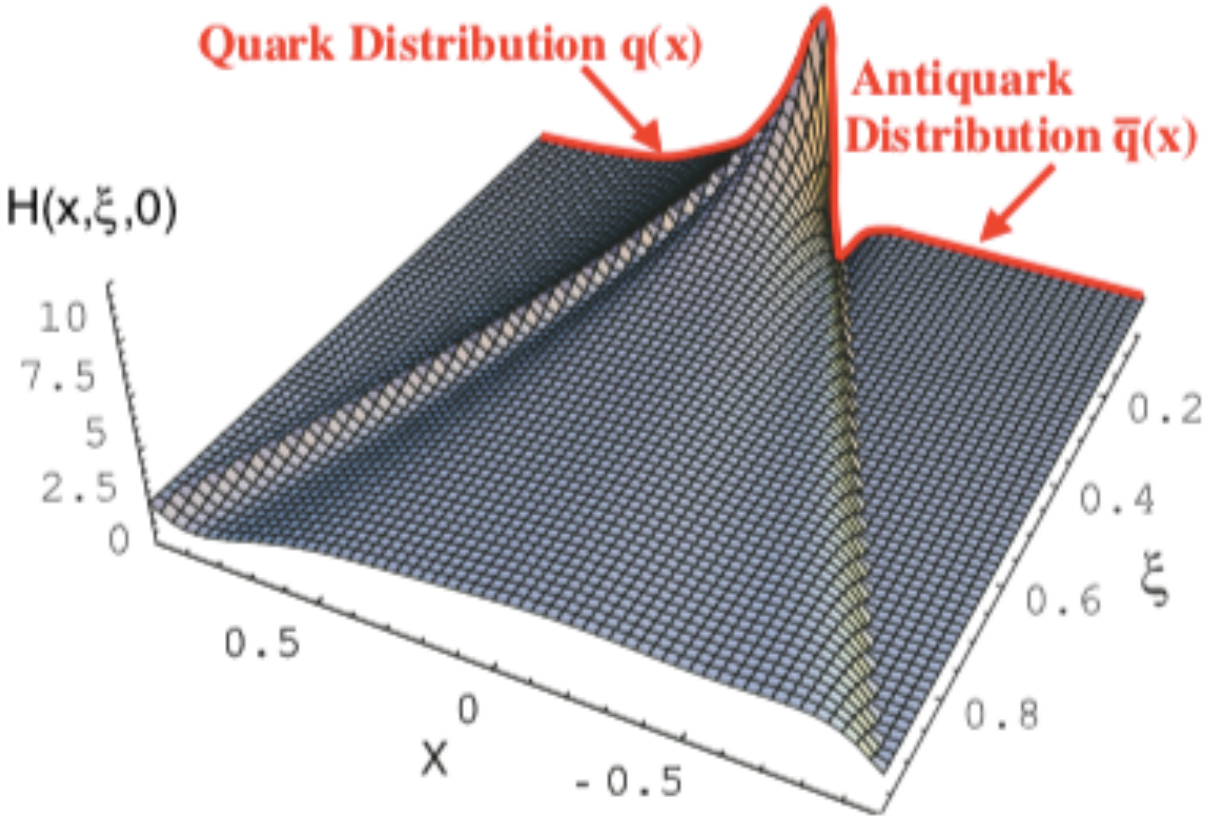


Figure 2: Forward limit for $H(x, \xi, 0)$[8]

[8] K Goeke, Maxim V Polyakov, and M Vanderhaeghen. Hard exclusive reactions and the structure of hadrons. *arXiv preprint hep-ph/0106012* (2001).

Moment Constraints

For $\xi, t \in (-1, 1)$,

$$\int_{-1}^{1} x^{n-1} H(x, \xi, t)\, \mathrm{d}x = \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (2\xi)^{2k} A_{n,k}(t) + mod(n-1, 2)(2\xi)^n C_n(t),$$

$$\int_{-1}^{1} x^{n-1} E(x, \xi, t)\, \mathrm{d}x = \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} (2\xi)^{2k} B_{n,k}(t) - mod(n-1, 2)(2\xi)^n C_n(t),$$

where $A_{n,k}, C_n, B_{n,k}$ are arbitrary functions.

Summary of the Mathematical Model of GPDs

Integral equations:

$$\mathcal{H}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi}\,[H(x,\xi,t) - H(-x,\xi,t)]\,\mathrm{d}x - i\pi[H(\xi,\xi,t) - H(-\xi,\xi,t)], \quad \xi,t \in (-1,1), \tag{5.1}$$

$$\mathcal{E}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi}\,[E(x,\xi,t) - E(-x,\xi,t)]\,\mathrm{d}x - i\pi[E(\xi,\xi,t) - E(-\xi,\xi,t)], \quad \xi,t \in (-1,1), \tag{5.2}$$

$$\tilde{\mathcal{H}}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi}\,\left[\tilde{H}(x,\xi,t) + \tilde{H}(-x,\xi,t)\right]\mathrm{d}x - i\pi[\tilde{H}(\xi,\xi,t) + \tilde{H}(-\xi,\xi,t)], \quad \xi,t \in (-1,1), \tag{5.3}$$

$$\tilde{\mathcal{E}}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi}\,\left[\tilde{E}(x,\xi,t) + \tilde{E}(-x,\xi,t)\right]\mathrm{d}x - i\pi[\tilde{E}(\xi,\xi,t) + \tilde{E}(-\xi,\xi,t)], \quad \xi,t \in (-1,1), \tag{5.4}$$

with constraints

$$H(x,0,0) = q(x), \quad \tilde{H}(x,0,0) = \Delta q(x), \qquad 0 \leqslant x < 1 \tag{5.5}$$

$$H(x,0,0) = -\tilde{q}(-x), \quad \tilde{H}(x,0,0) = \Delta q(-x), \qquad -1 < x < 0 \tag{5.6}$$

$$\int_{-1}^{1} x^{n-1} H(x,\xi,t)\,\mathrm{d}x = \sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} (2\xi)^{2k} A_{n,k}(t) + \mathrm{mod}(n-1,2)(2\xi)^n C_n(t), \tag{5.7}$$

$$\int_{-1}^{1} x^{n-1} E(x,\xi,t)\,\mathrm{d}x = \sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} (2\xi)^{2k} B_{n,k}(t) - \mathrm{mod}(n-1,2)(2\xi)^n C_n(t). \tag{5.8}$$

The Cauchy Singular Integral Operator and Chebyshev Polynomials

Let T_n and U_n, respectively, be the Chebyshev polynomials of order n, of the first and second kind. That is,

$$T_n(x) = \cos n\theta, \quad U_n(x) = \frac{\sin(n+1)\theta}{\sin\theta}, \quad \cos\theta = x, \quad \theta \in (0, 2\pi).$$

T_n and U_n are connected by the Cauchy Singular Integral Operator[9]:

$$\int_{-1}^{1} \frac{U_n(t)\sqrt{1-t^2}}{t-x}\,\mathrm{d}t = -\pi T_{n+1}(x), \quad x \in (-1, 1)$$

$$\int_{-1}^{1} \frac{T_n(t)}{(t-x)\sqrt{1-t^2}}\,\mathrm{d}t = \begin{cases} 0, & n = 0, \\ \pi U_{n-1}(x), & n \geqslant 1, \end{cases} \quad x \in (-1, 1).$$

[9] David Elliott. Orthogonal polynomials associated with singular integral equations having a Cauchy kernel. *SIAM Journal on Mathematical Analysis* 13.6 (1982), pp. 1041–1052.

The Cauchy Singular Integral Operator and Legendre Polynomials

- The Legendre polynomials P_n and the second kind Legendre function Q_n are connected by the Cauchy singular integral operator in the way:

$$\int_{-1}^{1} \frac{P_n(t)}{t-x}\,\mathrm{d}t = -2Q_n(x), \quad x \in (-1,1).$$

- The connection of the Cauchy singular integral operator with orthogonal polynomials motivates us to use orthogonal polynomials for approximation of GPDs.

Approximate Subspaces of $[H, E]$

Let $I_n := \left\{2i : i = 0, 1, \ldots, \left\lfloor \frac{n-1}{2} \right\rfloor\right\}$, $V_n := \mathrm{span}\{P_k : k \in I_n\}$.
The moment constraints (5.7)-(5.8) are equivalent to that for $\xi, t \in (-1, 1)$,

$$\int_{-1}^{1} x^{n-1} H(x, \xi, t) \mathrm{d}x \in V_n, \quad n = 1, 2, \ldots \tag{7.1}$$

$$\int_{-1}^{1} x^{n-1} E(x, \xi, t) \mathrm{d}x \in V_n, \quad n = 1, 2, \ldots \tag{7.2}$$

$$\int_{-1}^{1} x^{n-1} \left(H(x, \xi, t) + E(x, \xi, t)\right) \mathrm{d}x \in V_{n-1}, \quad \text{if } \mathrm{mod}(n, 2) = 0. \tag{7.3}$$

Representation of $[H, E]$

For $x, \xi, t \in (-1, 1)$, define

$$\Psi_{j,i,l}(x, \xi, t) := \omega(x) P_j(\xi) W_i(x) P_l(t),$$

where $\{W_i\}$ is an orthogonal basis for $L^2([-1, 1], \omega)$, and let

$$H(x, \xi, t) := \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} \alpha_{j,i,l} \Psi_{j,i,l}(x, \xi, t), \tag{7.4}$$

$$E(x, \xi, t) := \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} \beta_{j,i,l} \Psi_{j,i,l}(x, \xi, t). \tag{7.5}$$

Lemma 1

The following statements are equivalent:

(1) Equation (7.1) holds.

(2)

$$\sum_{i=0}^{\infty} G_{n,i} \alpha_{j,i,l} = 0, \quad \textit{for all} \quad j \in \mathbb{N}/I_n, l, n, \tag{7.6}$$

where $G_{n,i} := \int_{-1}^{1} \omega(x) x^{n-1} W_i(x) \mathrm{d}x$.

(3) For odd j, $\alpha_{j,i,l} = 0$, $i \geqslant 0$; *for even* $j \geqslant 2$, $\alpha_{j,i,l} = 0$, $i \leqslant j - 2$.

Proposition 1

If H, E have the form like (7.4) (7.5) respectively, then equations (5.7) and (5.8) are equivalent to

$$H(x,\xi,t) = \sum_{j=1}^{\infty}\sum_{i=2j-1}^{\infty}\sum_{l=0}^{\infty}\alpha_{2j,i,l}\Psi_{2j,i,l}(x,\xi,t) + \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\alpha_{0,i,l}\Psi_{0,i,l}(x,\xi,t),$$

$$E(x,\xi,t) = \sum_{j=1}^{\infty}\sum_{i=2j-1}^{\infty}\sum_{l=0}^{\infty}\beta_{2j,i,l}\Psi_{2j,i,l}(x,\xi,t) + \sum_{i=0}^{\infty}\sum_{l=0}^{\infty}\beta_{0,i,l}\Psi_{0,i,l}(x,\xi,t),$$

for $x, \xi, t \in (-1,1)$, and

$$\alpha_{n,n-1,l} + \beta_{n,n-1,l} = 0, \quad \textit{for all} \;\; n \geqslant 2, \;\; \textit{mod}\,(n,2) = 0. \tag{7.7}$$

According to Proposition 1, both H and E are in

$$\mathbb{X} := \left\{ f : f(x,\xi,t) := \sum_{j=1}^{\infty} \sum_{i=2j-1}^{\infty} \sum_{l=0}^{\infty} \gamma_{2j,i,l} \Psi_{2j,i,l}(x,\xi,t), \right.$$
$$\left. + \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} \gamma_{0,i,l} \Psi_{0,i,l}(x,\xi,t), f \in L^2(\Omega) \right\},$$

where $\Omega = (-1,1)^3$. This motivates us to choose the approximation subspace for $[H, E]$ as

$$\mathbb{X}_N := \left\{ f : f(x,\xi,t) := \sum_{j=1}^{N/2} \sum_{i=2j-1}^{2j+N-1} \sum_{l=0}^{N} \gamma_{2j,i,l} \Psi_{2j,i,l}(x,\xi,t) \right.$$
$$\left. + \sum_{i=0}^{N} \sum_{l=0}^{N} \gamma_{0,i,l} \Psi_{0,i,l}(x,\xi,t), f \in L^2(\Omega) \right\}, \text{ for } \mathrm{mod}(N,2) = 0.$$

Clearly, $s_N := \dim \mathbb{X}_N = (N+1)^2(1+0.5N)$. We use $\{\Phi_j : j \in \mathbb{N}_{s_N}\}$ to denote a basis of $\mathbb{X}_N$.

Discretization of the Singular Integral Equations

We use a collocation method[10] to discretize the equations. Comparing the real part of (5.1) and (5.2) yields, respectively,

$$\mathcal{R}e\,\mathcal{H}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi}\left[H(x,\xi,t) - H(-x,\xi,t)\right] \mathrm{d}x, \quad \xi,t \in (-1,1) \tag{8.1}$$

$$\mathcal{R}e\,\mathcal{E}(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi}\left[E(x,\xi,t) - E(-x,\xi,t)\right] \mathrm{d}x, \quad \xi,t \in (-1,1). \tag{8.2}$$

Since equations (8.1) and (8.2) have a similar form, we consider a unified equation below

$$f(\xi,t) = \int_{-1}^{1} \frac{1}{x-\xi}\left(u(x,\xi,t) - u(-x,\xi,t)\right) \mathrm{d}x, \tag{8.3}$$

where $u \in \mathbb{X}$. For all $g \in \mathbb{X}$, define operator $\mathcal{K}$ by

$$(\mathcal{K}g)(\xi,t) := \int_{-1}^{1} \frac{1}{x-\xi}\left(g(x,\xi,t) - g(-x,\xi,t)\right) \mathrm{d}x.$$

Equation (8.3) becomes

$$\mathcal{K}u = f.$$

[10] Zhongying Chen, Charles A Micchelli, and Yuesheng Xu. *Multiscale methods for Fredholm integral equations*. Vol. 28. Cambridge University Press, 2015.

A Collocation Scheme

Define $\mathbb{Y} := \mathcal{K}\mathbb{X}$, $\mathbb{Y}_N := \mathcal{K}\mathbb{X}_N$, $\mathbb{X}'_N := \mathrm{Null}(\mathcal{K})^\perp \cap \mathbb{X}_N$, where $\mathrm{Null}(\mathcal{K})$ denotes the null space of $\mathcal{K}$. Then $\mathcal{K}|_{\mathbb{X}'_N}$ is bijective from $\mathbb{X}'_N$ onto $\mathbb{Y}_N$.
Let $d_N := \dim \mathbb{X}'_N$.
The collocation scheme for equation (8.3) is to find $u_N \in \mathbb{X}'_N$ such that

$$\int_{-1}^{1} \frac{1}{x-\xi_j}(u_N(x,\xi_j,t_j) - u_N(-x,\xi_j,t_j))\,\mathrm{d}x = f(\xi_j,t_j), \text{for all } j \in \mathbb{N}_m, \tag{8.4}$$

where $\{(\xi_j,t_j) : j \in \mathbb{N}_m\}$ are distinct points.
Let $u_N := \sum_{j\in\mathbb{N}_{d_N}} u_j\phi_j$, where $\{\phi_j : j \in \mathbb{N}_{d_N}\}$ is a basis for $\mathbb{X}'_N$. Define

$$\mathbf{K} := [(\mathcal{K}\phi_j)(\xi_i,t_i) : i \in \mathbb{N}_m, j \in \mathbb{N}_{d_N}],$$

$$\mathbf{u} := [u_j : j \in \mathbb{N}_{d_N}], \quad \mathbf{f} := [f(\xi_i,t_i) : i \in \mathbb{N}_m].$$

Equation (8.4) can be written in the matrix form

$$\mathbf{K}\mathbf{u} = \mathbf{f}. \tag{8.5}$$

Changes of Bases

We choose a basis $\{\phi_{d_N+1}, \phi_{d_N+2}, \ldots, \phi_{s_N}\}$ for $\mathrm{Null}(\mathcal{K}) \cap \mathbb{X}_N$. Then, $\{\phi_i : i \in \mathbb{N}_{s_N}\}$ forms a basis for $\mathbb{X}_N$. We wish to change to the basis $\{\Phi_i : i \in \mathbb{N}_{s_N}\}$. There is an invertible matrix $\mathbf{B}$ such that

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{s_N} \end{bmatrix} = \mathbf{B} \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_{s_N} \end{bmatrix}.$$

Let $\hat{\mathbf{u}} := \mathbf{B}\mathbf{u}$. Equation (8.5) becomes

$$\hat{\mathbf{K}}\hat{\mathbf{u}} = \mathbf{f}, \tag{8.6}$$

where

$$\hat{\mathbf{K}} := [(\mathcal{K}\Phi_j)(\xi_i, t_i) : i \in \mathbb{N}_m, j \in \mathbb{N}_{s_N}],$$

and $\hat{\mathbf{u}}$ is the coefficients of u under the basis $\{\Phi_i : i \in \mathbb{N}_{s_N}\}$.

Equations for $[H, E]$: the Real Part

Suppose we have already given data

$$\{\mathcal{H}(\xi_i, t_i) : i \in \mathbb{N}_m\},\ \{\mathcal{E}(\xi_i^*, t_i^*) : i \in \mathbb{N}_p\},\ \{H(x_i, 0, 0) : i = \mathbb{N}_k\}.$$

We choose $m = p = k$. The values $\{x_i, x_i^*, \xi_i, \xi_i^*, t_i : i = 1, 2, \ldots, m\}$ are sampled from the uniform distribution in $(0, 1)$ independently.
First, we consider the real part of equations (5.1) and (5.2). H and E satisfy equations in the form (8.3). Substituting $\mathcal{R}e\,\mathcal{H}$ for f and H for u in (8.3), we have the following systems similar to (8.6):

$$\hat{\mathbf{K}}^H \hat{\mathbf{u}}^H = \mathbf{f}^H, \tag{8.7}$$

$$\hat{\mathbf{K}}^E \hat{\mathbf{u}}^E = \mathbf{f}^E, \tag{8.8}$$

where $\hat{\mathbf{K}}^H, \hat{\mathbf{K}}^E, \mathbf{f}^H, \mathbf{f}^E, \hat{\mathbf{u}}^H, \hat{\mathbf{u}}^E$ are defined in the same way as those in (8.6).

Equations for $[H, E]$: the Imaginary Part

Considering the imaginary part of equations (5.1) and (5.2) yields

$$\mathbf{G}^H \hat{\mathbf{u}}^H = \mathbf{h}^H, \quad \mathbf{G}^E \hat{\mathbf{u}}^E = \mathbf{h}^E, \tag{8.9}$$

where

$$\mathbf{h}^H := [-\mathcal{I}m\,\mathcal{H}(\xi_i, t_i)/\pi : i \in \mathbb{N}_m], \ \mathbf{h}^E := [-\mathcal{I}m\,\mathcal{E}(\xi_i^*, t_i^*)/\pi : i \in \mathbb{N}_p],$$

$$\mathbf{G}^H := [(\mathcal{W}\Phi_j)(\xi_i, t_i) : i \in \mathbb{N}_m, j \in \mathbb{N}_{s_N}],$$

$$\mathbf{G}^E := [(\mathcal{W}\Phi_j)(\xi_i^*, t_i^*) : i \in \mathbb{N}_p, j \in \mathbb{N}_{s_N}],$$

$$(\mathcal{W}u)(\xi, t) := u(\xi, \xi, t) - u(-\xi, \xi, t), \quad \text{for any } u \in \mathbb{X}.$$

Matrix Form of the Initial Constraints

The initial constraints described in (5.5) and (5.6) are translated to the matrix form

$$\mathbf{F}\hat{\mathbf{u}}^H = \mathbf{g}, \tag{8.10}$$

where

$$\mathbf{F} := [\Psi_j(x_i, 0, 0) : i \in \mathbb{N}_k, j \in \mathbb{N}_{s_N}],$$

$$\mathbf{g} := [H(x_i, 0, 0) : i = \mathbb{N}_k].$$

The Discrete System for GPD

Let

$$\mathbf{G} := \begin{bmatrix} \mathrm{diag}(\hat{\mathbf{K}}^H, \hat{\mathbf{K}}^E) \\ \mathrm{diag}(\mathbf{G}^H, \mathbf{G}^E) \\ \mathbf{F}, \mathbf{0}_{k \times s_N} \end{bmatrix}, \quad \mathbf{N} := \begin{bmatrix} \mathbf{f}^H \\ \mathbf{f}^E \\ \mathbf{h}^H \\ \mathbf{h}^E \\ \mathbf{g} \end{bmatrix}, \quad \hat{\mathbf{u}} := \begin{bmatrix} \hat{\mathbf{u}}^H \\ \hat{\mathbf{u}}^E \end{bmatrix},$$

where $\mathbf{0}_{k \times s_N}$ is a zero matrix whose dimension is $k \times s_N$.

- GPD can be solved by the discrete system

$$\mathbf{G}\hat{\mathbf{u}} = \mathbf{N}. \tag{9.1}$$

  Moreover, $\hat{\mathbf{u}}^E$ and $\hat{\mathbf{u}}^H$ must satisfy the relationship (7.7).
- System (9.1) is ill-posed. It requires regularization.
- Regularization uses prior information on a potential solution to transfer the ill-posed system to a well-posed problem.

L_2 Regularization

Under the hypothesis that the true solution has certain minimum energy, we propose the L_2 regularization:

$$\min_{\hat{\mathbf{u}}} \left\{ \frac{1}{2} \|\mathbf{G}\hat{\mathbf{u}} - \mathbf{N}\|_2^2 + \lambda \|\hat{\mathbf{u}}\|_2 \right\}, \tag{9.2}$$

where $\lambda > 0$ is a regularization parameter.
The L_2 regularization problem (9.2) has a closed-form solution

$$\hat{\mathbf{u}} = (\mathbf{G}^T\mathbf{G} + \lambda\mathcal{I})^{-1}(\mathbf{G}^T\mathbf{N}).$$

If the solution of (9.2) is given by $\hat{\mathbf{u}} = \begin{bmatrix} \hat{\mathbf{u}}^H \\ \hat{\mathbf{u}}^E \end{bmatrix}$ with $\hat{\mathbf{u}}^H := [\hat{u}_1^H, \ldots, \hat{u}_{s_N}^H]^T$ and $\hat{\mathbf{u}}^E := [\hat{u}_1^E, \ldots, \hat{u}_{s_N}^E]^T$, then

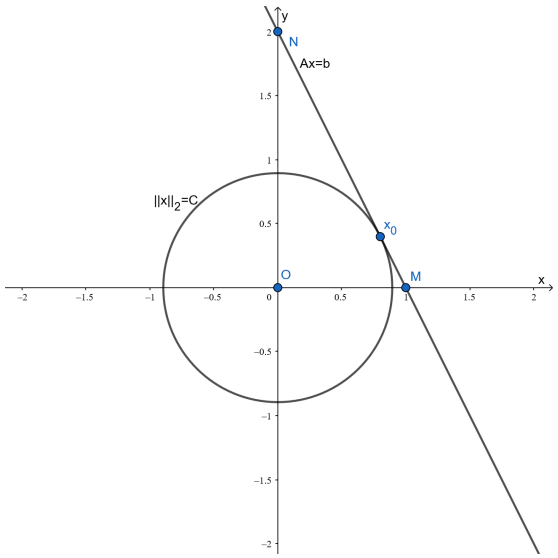
$$H_N := \sum_{j \in \mathbb{N}_{s_N}} \hat{u}_j^H \Phi_j, \quad E_N := \sum_{j \in \mathbb{N}_{s_N}} \hat{u}_j^E \Phi_j.$$

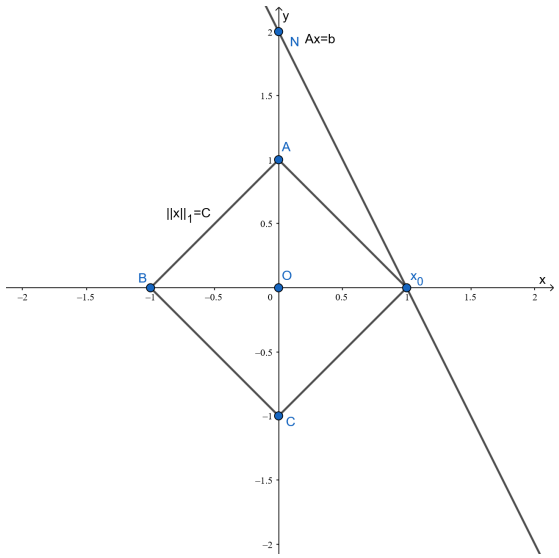
give solutions for H, E in $\mathbb{X}_N$.

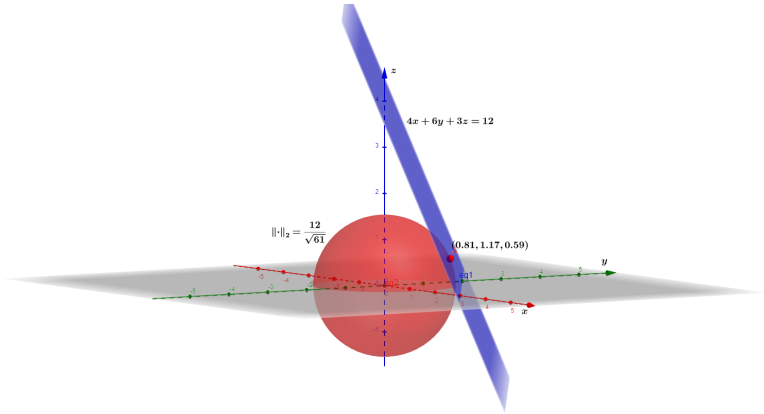
Sparse Solutions

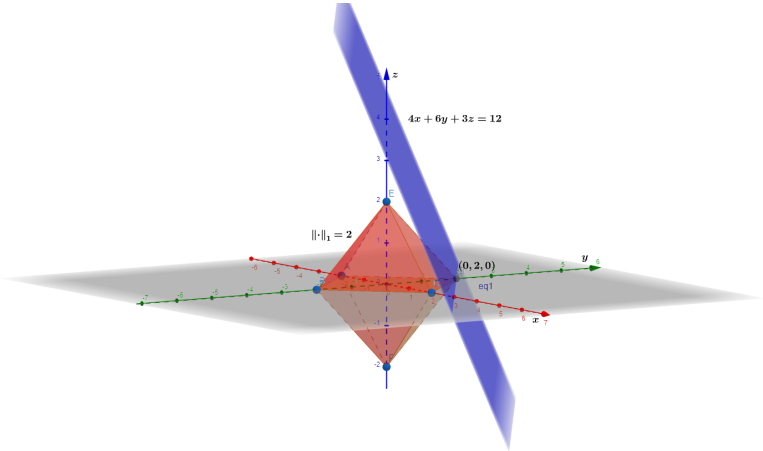
- Certain energy concentration in physics may require sparse representation of solutions.
- Sparse solutions are desirable also due to their efficiency in computation.
- The L_1-norm promotes sparsity. It helps us concentrate energy in a few directions.
- The L_2-norm does not promote sparsity. It spreads energy in all directions.

The L_2-Norm vs the L_1-Norm









L_1 Regularization

- Aiming at reconstructing sparse solutions, we propose the L_1 regularization[11] [12]

$$\min_{\hat{\mathbf{u}}} \left\{ \frac{1}{2} \|\mathbf{G}\hat{\mathbf{u}} - \mathbf{N}\|_2^2 + \lambda \|\hat{\mathbf{u}}\|_1 \right\}, \tag{9.3}$$

  where $\lambda > 0$ is a regularization parameter.
- The L_1 regularization leads to non-smooth minimization problems.
- Since their objective functions are not differentiable, gradient-type algorithms are not applicable. Solving non-smooth minimization problems requires special efforts.

[11] Emmanuel J Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory* 52.12 (2006), pp. 5406–5425.

[12] David L Donoho. For most large underdetermined systems of linear equations the minimal 1-norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences* 59.6 (2006), pp. 797–829.

Fixed-Point Proximity Algorithm I

Following the work[13], we propose to use a fixed-point proximity algorithm[14] to solve (9.3).

- We convert the minimization problem (9.3) to an equivalent fixed-point equation defined via the proximity operator of the l_1-norm $\|\cdot\|_1$.
- The fixed-point equation can be solved by the Picard iteration scheme with convergence guaranteed.
- The proximity operator of $\|\cdot\|_1$ has a closed form

$$\mathrm{prox}_{\lambda\|\cdot\|_1}(b) = \max\{|b| - \lambda, 0\}\,\mathrm{sign}(b),$$

which leads to fast computation.

[13] Charles A Micchelli, Lixin Shen, and Yuesheng Xu. Proximity algorithms for image models: denoising. *Inverse Problems* 27.4 (2011), p. 045009.

[14] Qia Li et al. Multi-step fixed-point proximity algorithms for solving a class of optimization problems arising from image processing. *Advances in Computational Mathematics* 41.2 (2015), pp. 387–422.

Fixed-Point Proximity Algorithm II

- Choose initial point $\hat{\mathbf{u}}^0$ and $\mathbf{y}^0$.
- Conduct the iteration:

$$\begin{cases} \hat{\mathbf{u}}^{k+1} = \hat{\mathbf{u}}^k - \beta(\nabla F(\hat{\mathbf{u}}^k) + \mathbf{y}^k), \\ \mathbf{y}^{k+1} = \rho(\mathcal{I} - \mathrm{prox}_{\frac{\lambda}{\rho}\|\cdot\|_1})(\frac{1}{\rho}\mathbf{y}^k + (\hat{\mathbf{u}}^{k+1} - \hat{\mathbf{u}}^k)) \end{cases}$$

where β, ρ are positive parameters, $\mathcal{I}$ is identity matrix,

$$\nabla F(\hat{\mathbf{u}}) = \mathbf{G}^T(\mathbf{G}\hat{\mathbf{u}} - \mathbf{N}).$$

- Stop the iteration when the relative error $\frac{\|\hat{\mathbf{u}}^{k+1} - \hat{\mathbf{u}}^k\|_2}{\|\hat{\mathbf{u}}^{k+1}\|_2}$ is smaller than a given number, and output $\hat{\mathbf{u}}^{k+1}$ as an approximate solution.
- When ρ is chosen in certain range, this iteration scheme converges linearly.

Simulation Data and Settings

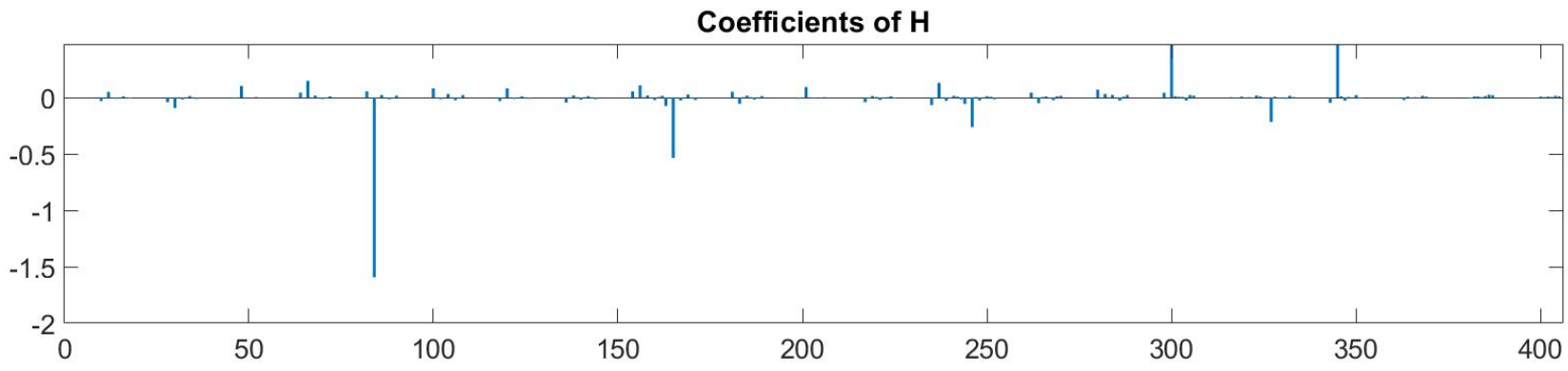
- The objective of the numerical experiments is to reconstruct the GPD functions H and E by solving equations (5.1)-(5.8) via our proposed method.
- We use simulation data to test our method.
  - Choose testing functions H and E that satisfy the moment constraints.
  - Sample the values $\{x_i, x_i^*, \xi_i, \xi_i^*, t_i : i = 1, 2, \ldots, 100\}$ uniformly from $(0, 1)$.
  - Compute

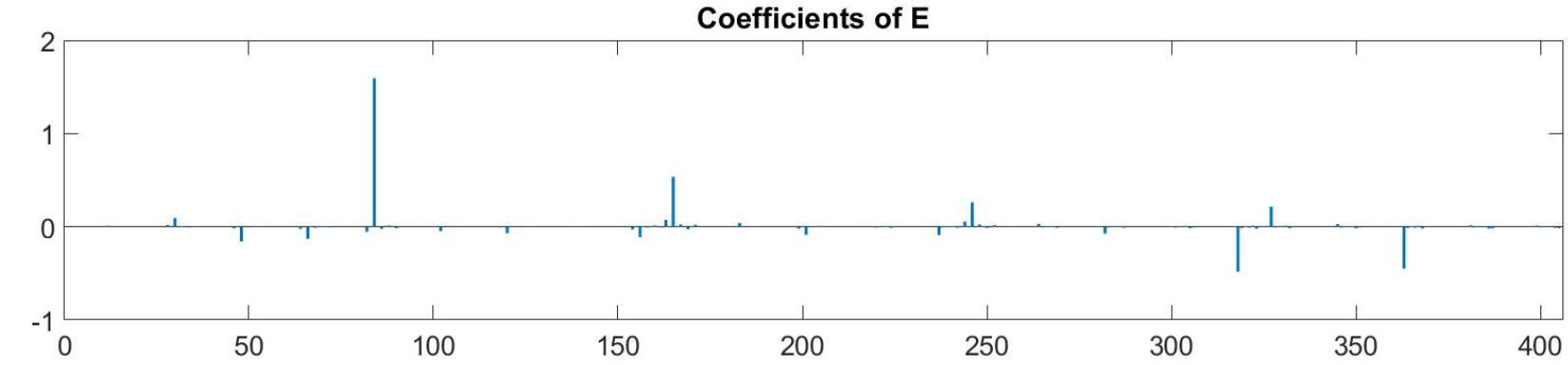
$$\mathcal{H}(\xi_i, t_i), \mathcal{E}(\xi_i^*, t_i^*), H(x_i, 0, 0), \quad i = 1, 2, \ldots, 100.$$

    and add gaussian noise $N(0, 0.01^2)$ to them.
- Construct approximate space $\mathbb{X}_N$.
- Solve (9.3) to obtain approximate functions of H_N and E_N by using the fixed-point proximity algorithm.

Experiment 1: Non-Sparse H and E

We choose functions H and E in $\mathbb{X}_8'$, and use $\mathbb{X}_6$ as an approximation space. The coefficients of H and E are showed as follows:



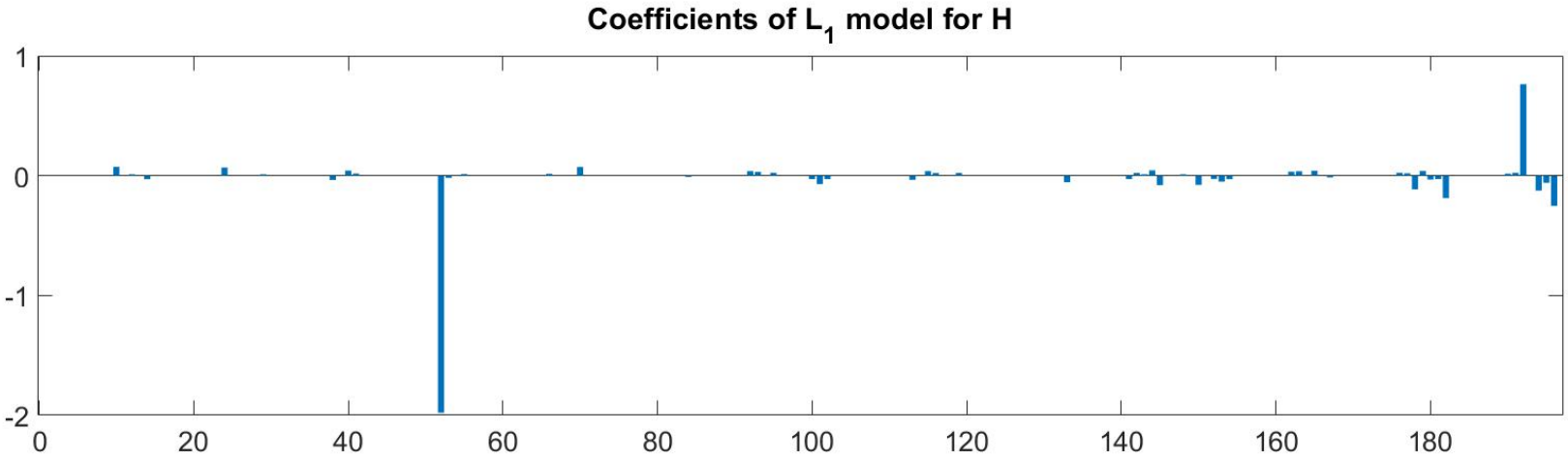


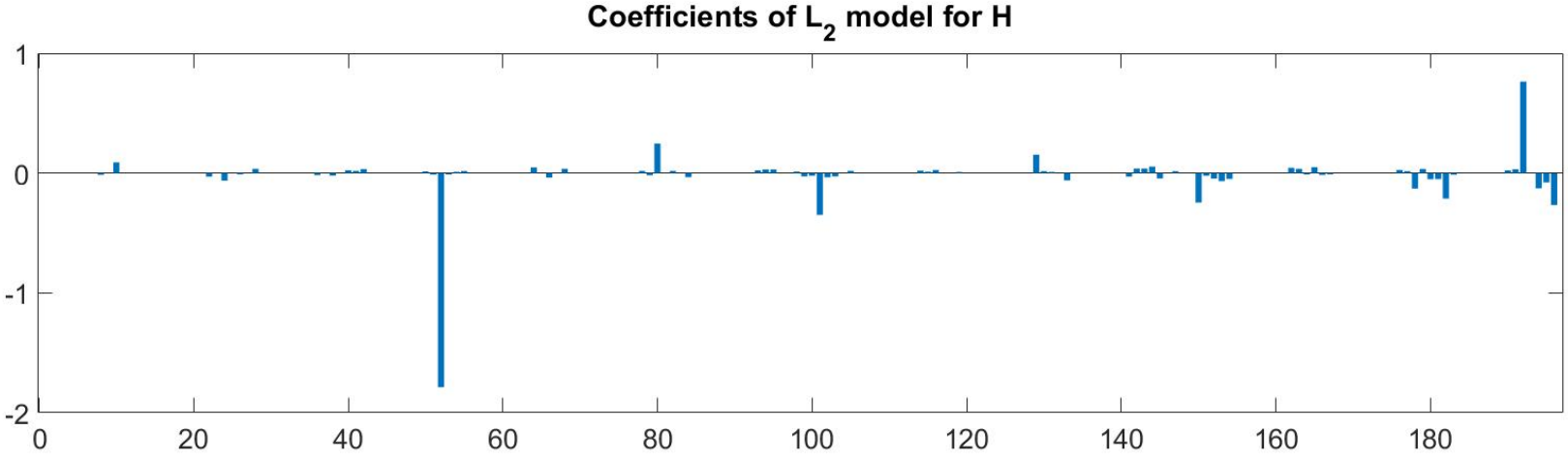
Experiment 1: Numerical Results

- The best regularization parameter for both the L_1 and L_2 models is 0.02.
- When the absolute value of a coefficient is less than 0.01, we let it be zero.
- The relative error is defined by $\frac{\|H-H_N\|_2^2}{\|H\|_2^2}$ for H and H_N and likewise for E and E_N.

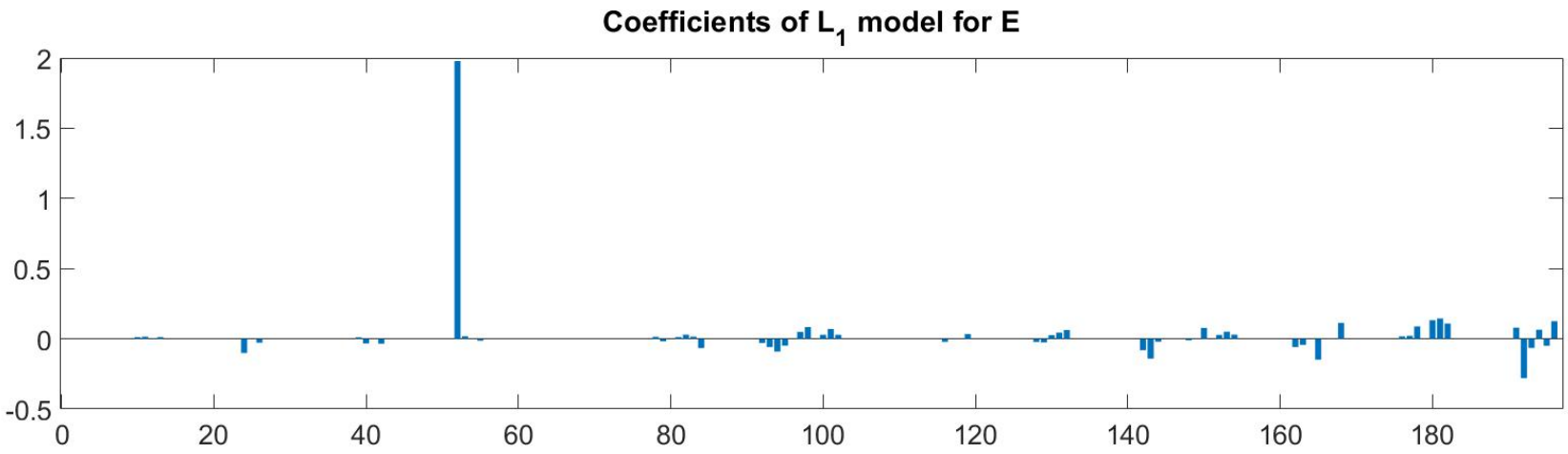
Models	relative errors		# of nonzero terms	
	H	E	H	E
The L_1 model	0.356	0.337	47	54
The L_2 model	0.235	0.239	78	83

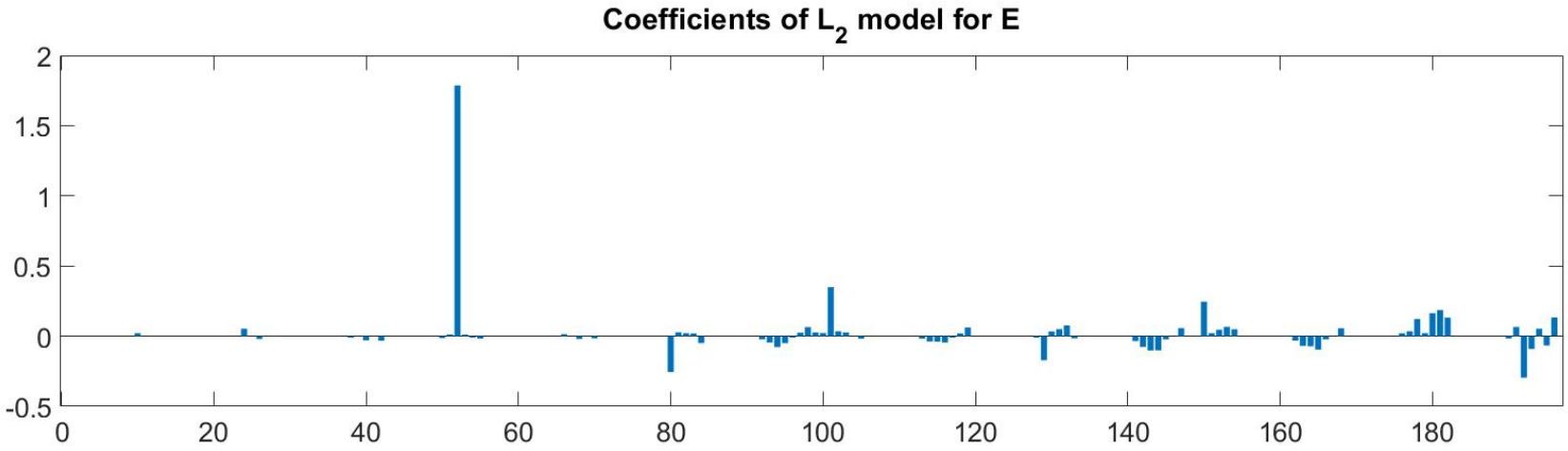
Experiment 1: Coefficients of the Reconstructed H_N



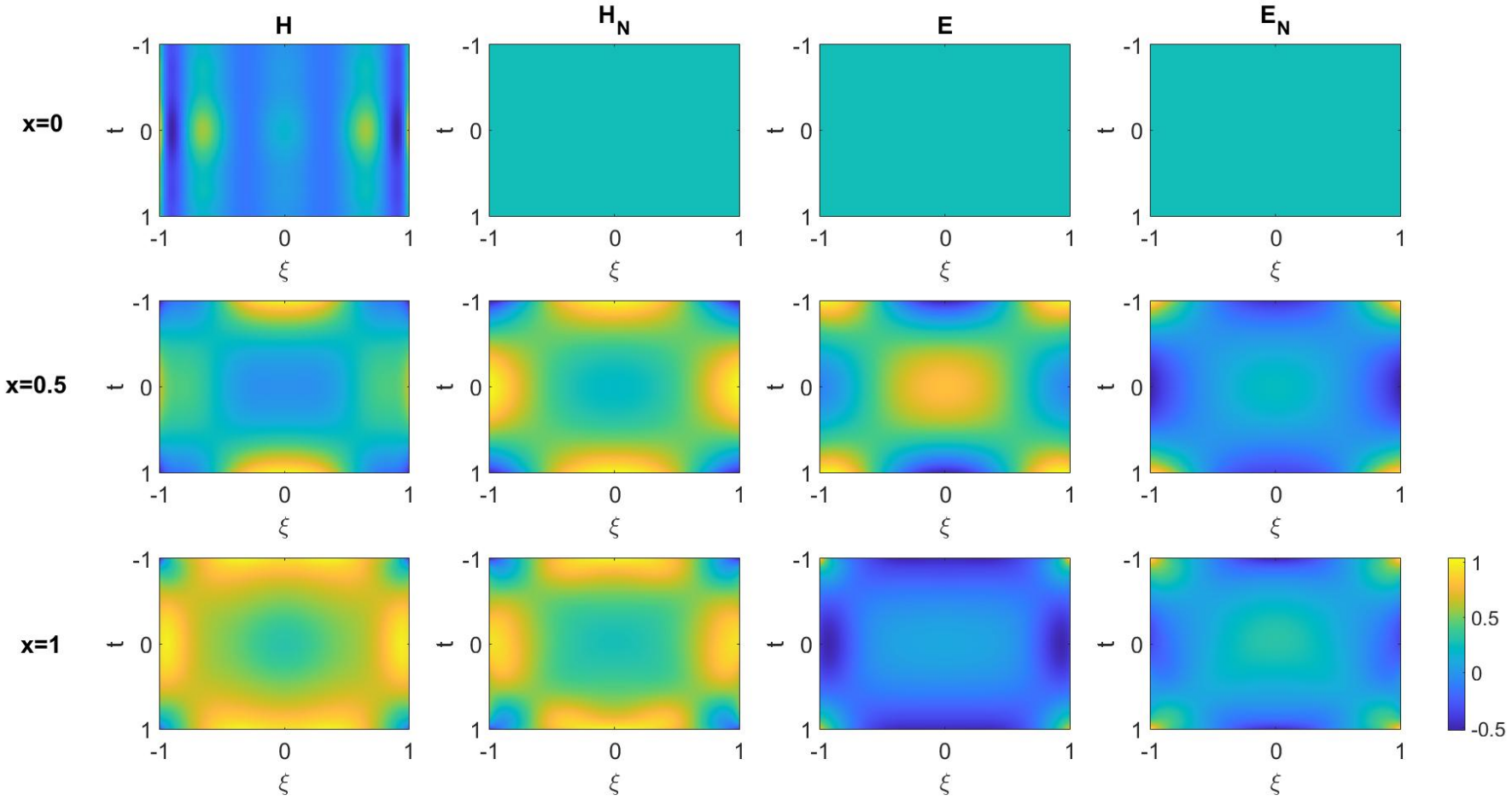


Experiment 1: Coefficients of the Reconstructed E_N

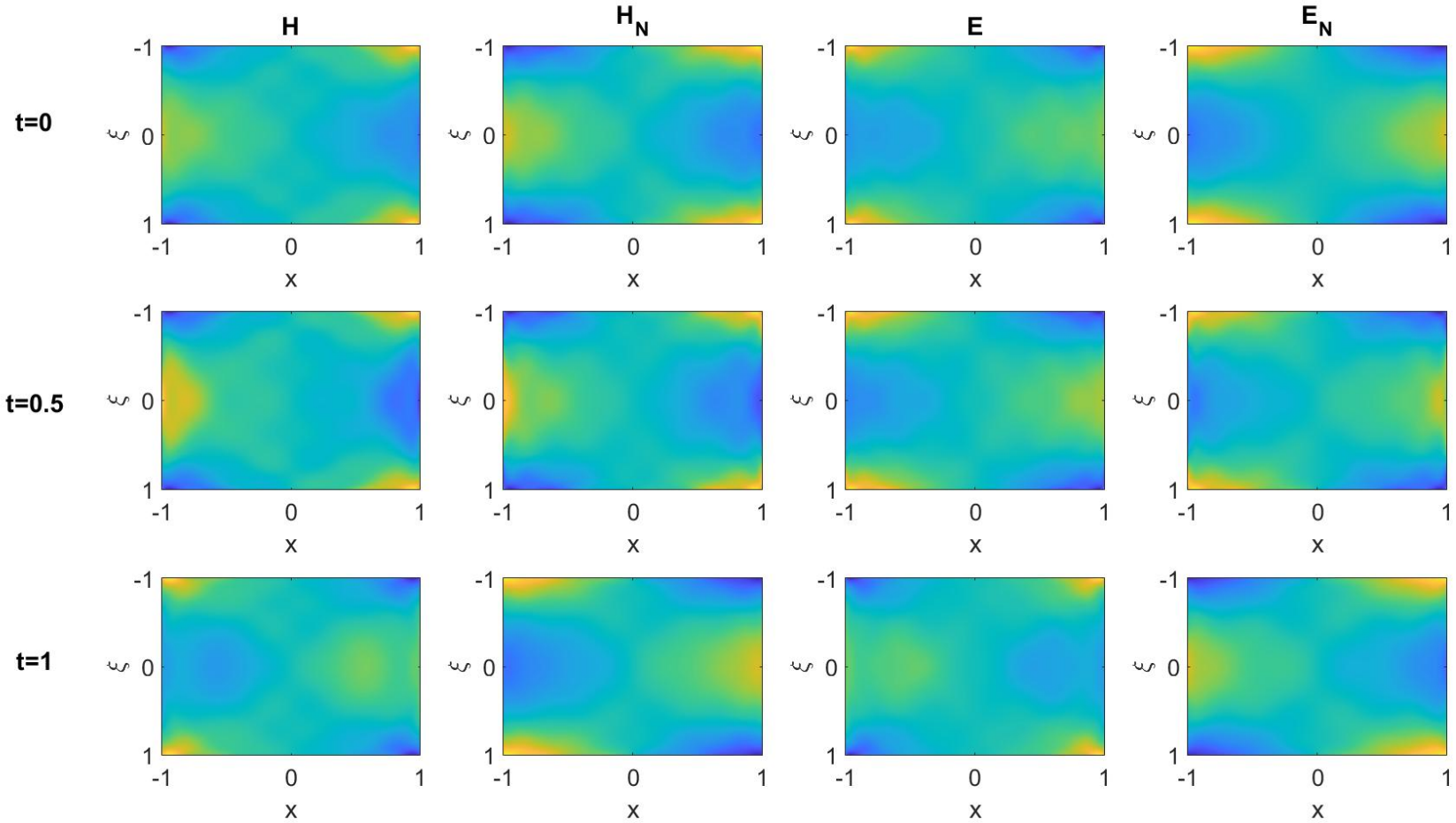




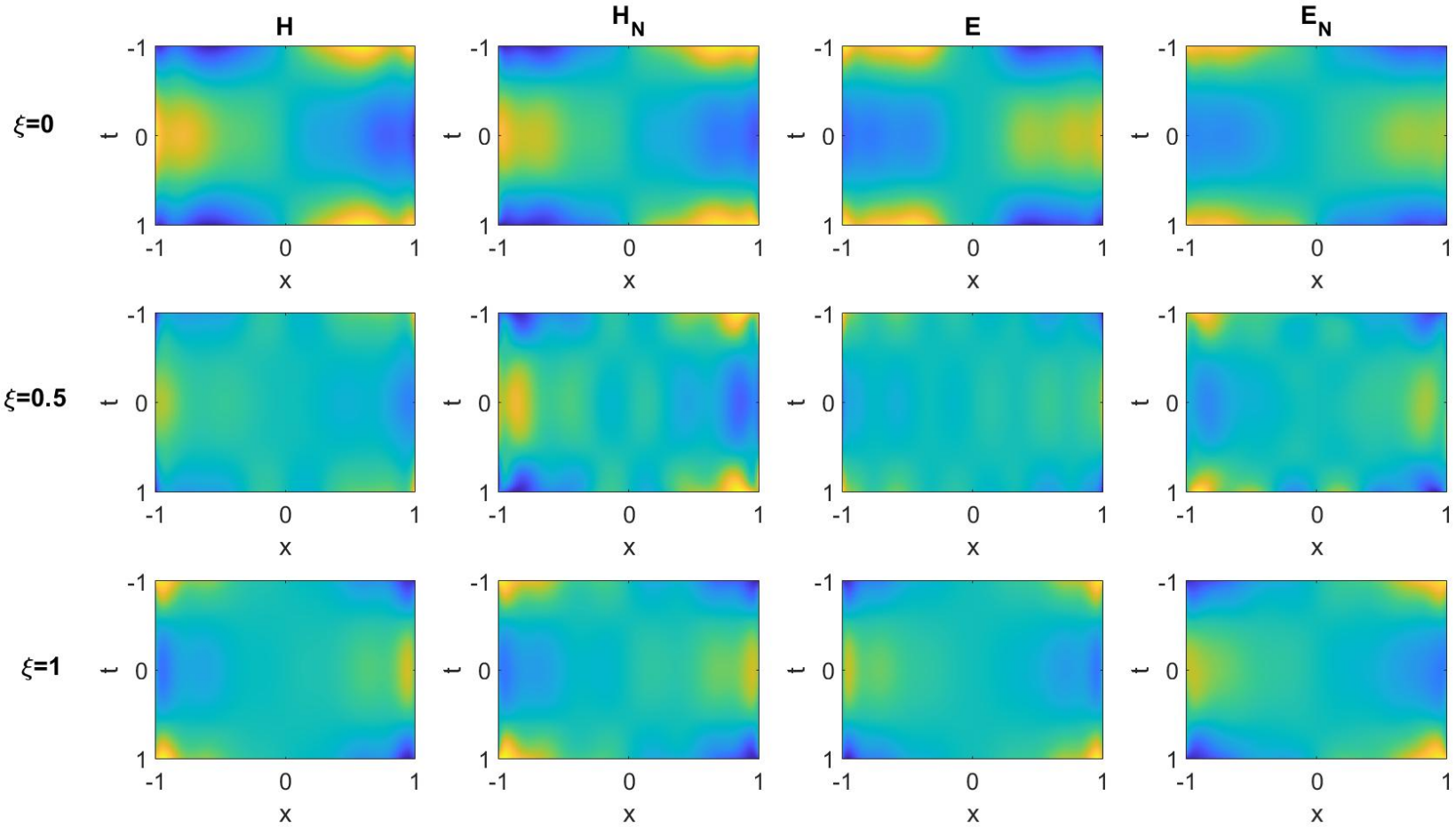
Experiment 1: The L_1 model, x-Slices



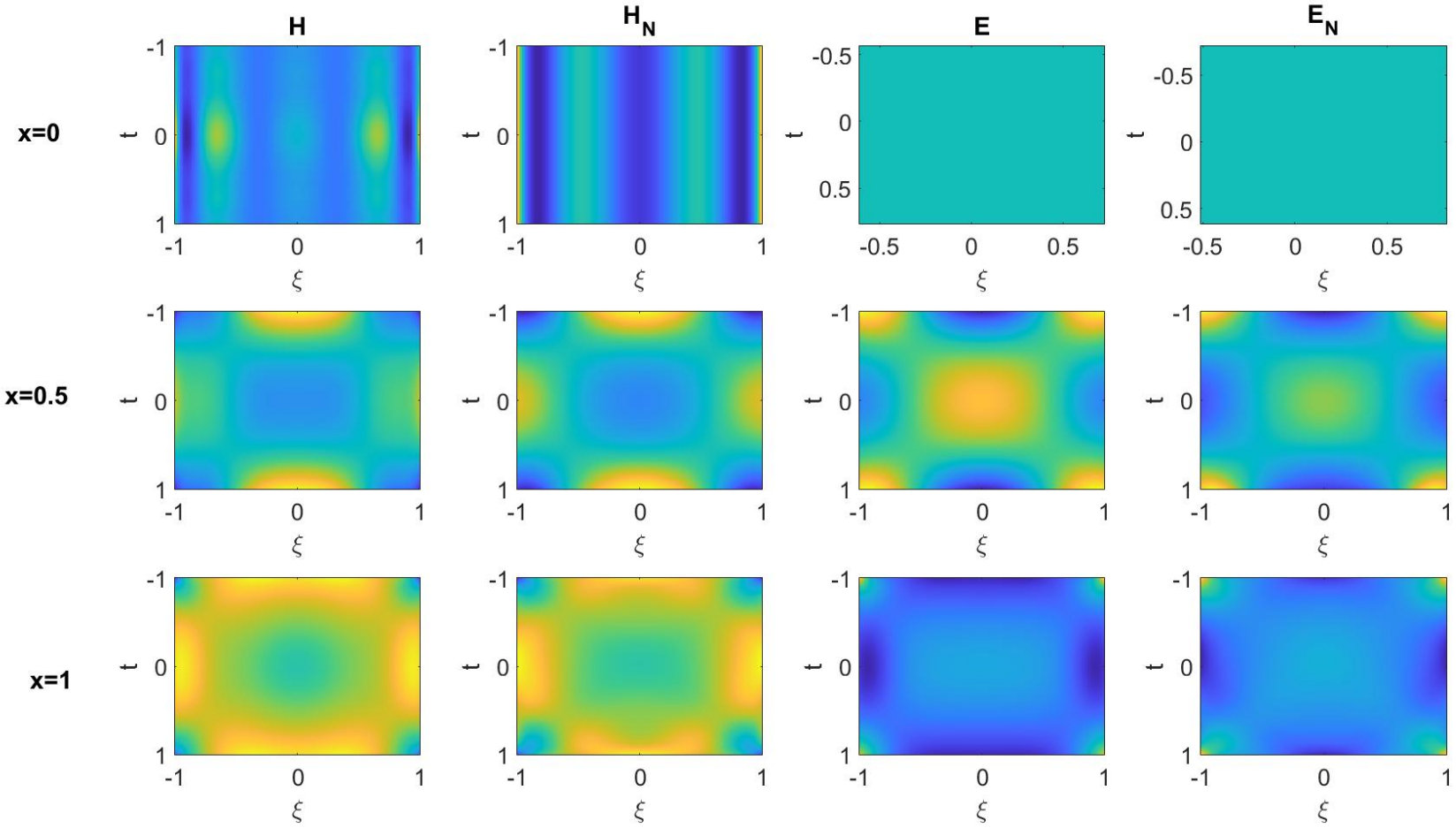
Experiment 1: The L_1 model, t-Slices



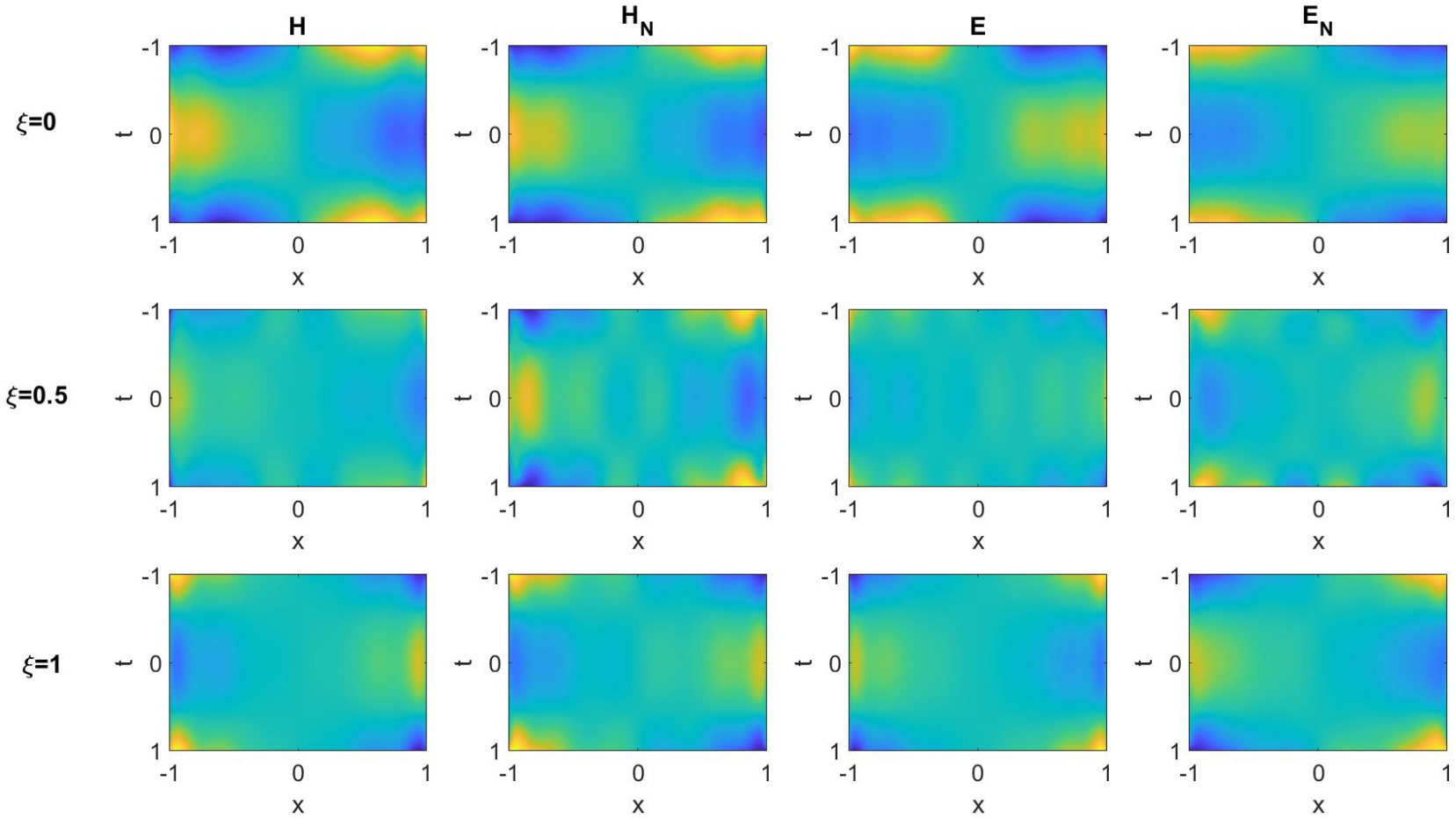
Experiment 1: The L_1 model, ξ-Slices



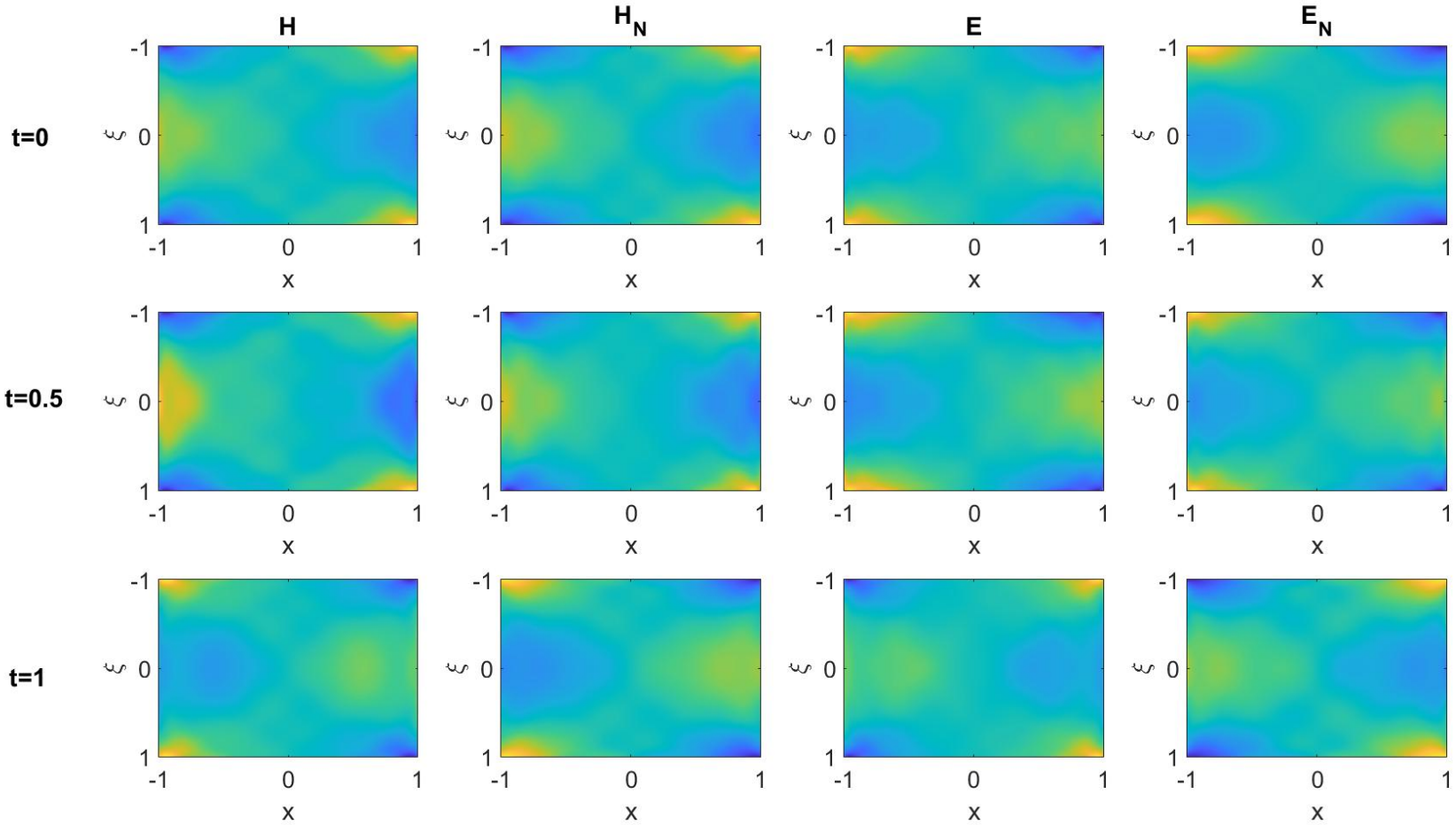
Experiment 1: The L_2 model, x-Slices



Experiment 1: The L_2 model, ξ-Slices



Experiment 1: The L_2 model, t-Slices



Experiment 2: Sparse H and E in $\mathbb{X}_6$

- Choose

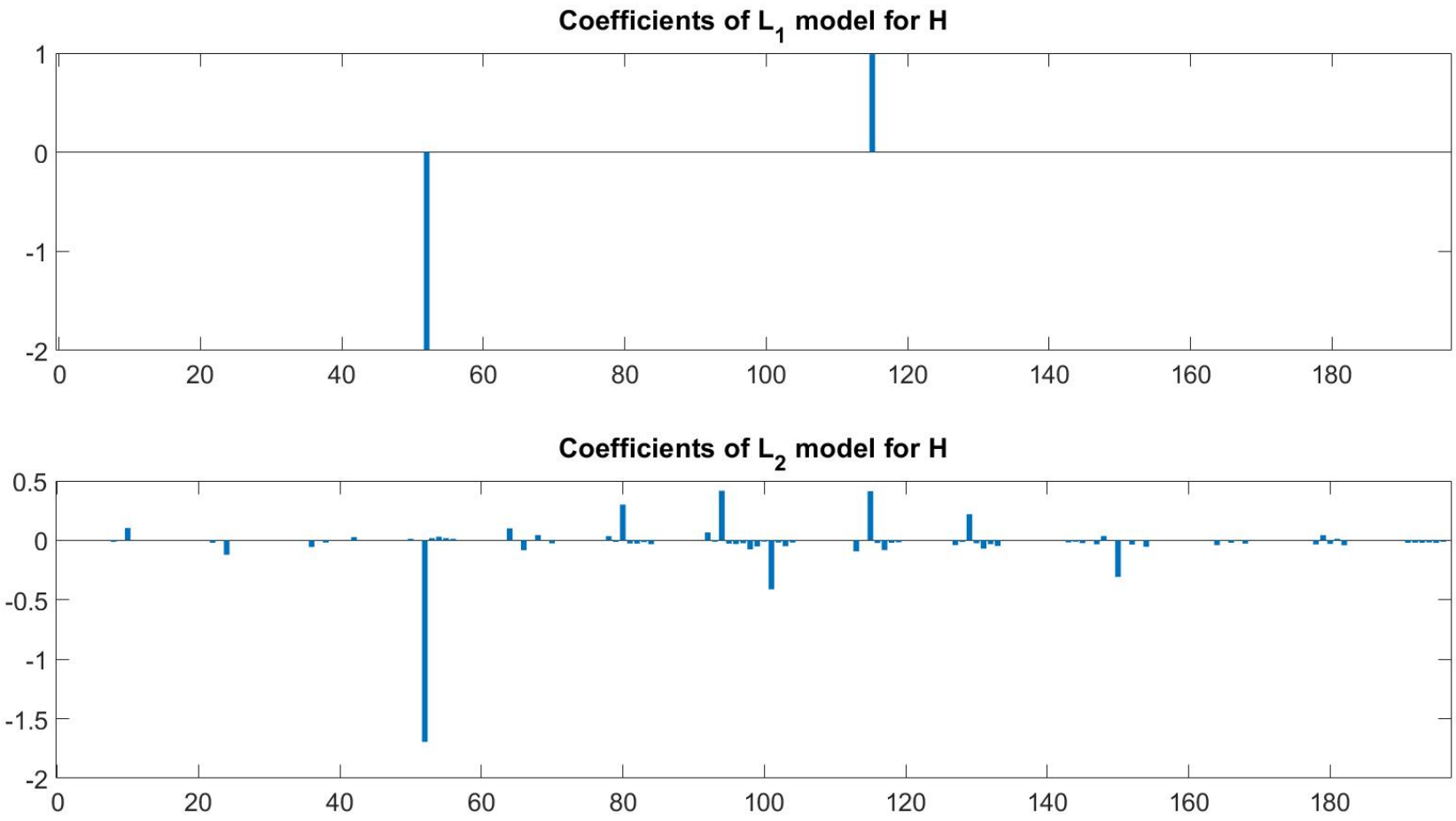
$$H(x,\xi,t) = P_4(\xi)P_5(x)P_2(t) - 2P_2(\xi)P_1(x)P_2(t),\ x,\xi,t \in (0,1),$$

$$E(x,\xi,t) = -P_0(\xi)P_3(x)P_2(t) + 2P_2(\xi)P_1(x)P_2(t),\ x,\xi,t \in (0,1).$$

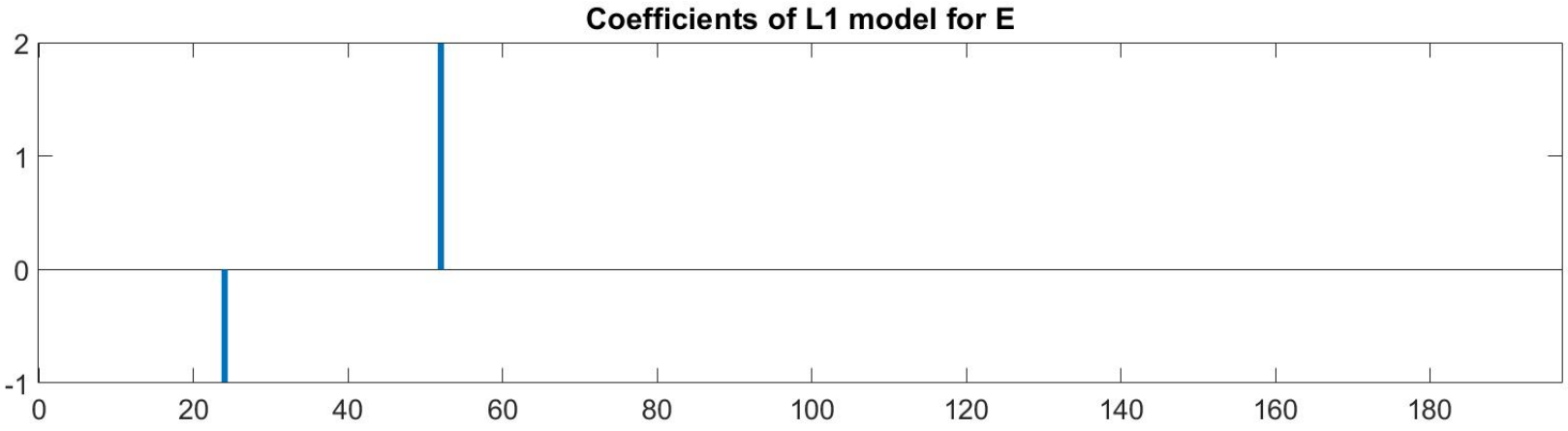
- Choose $\mathbb{X}_6$ as the approximation space.
- Best regularization parameters for the L_1 and L_2 models is respectively 0.012 and 0.05.

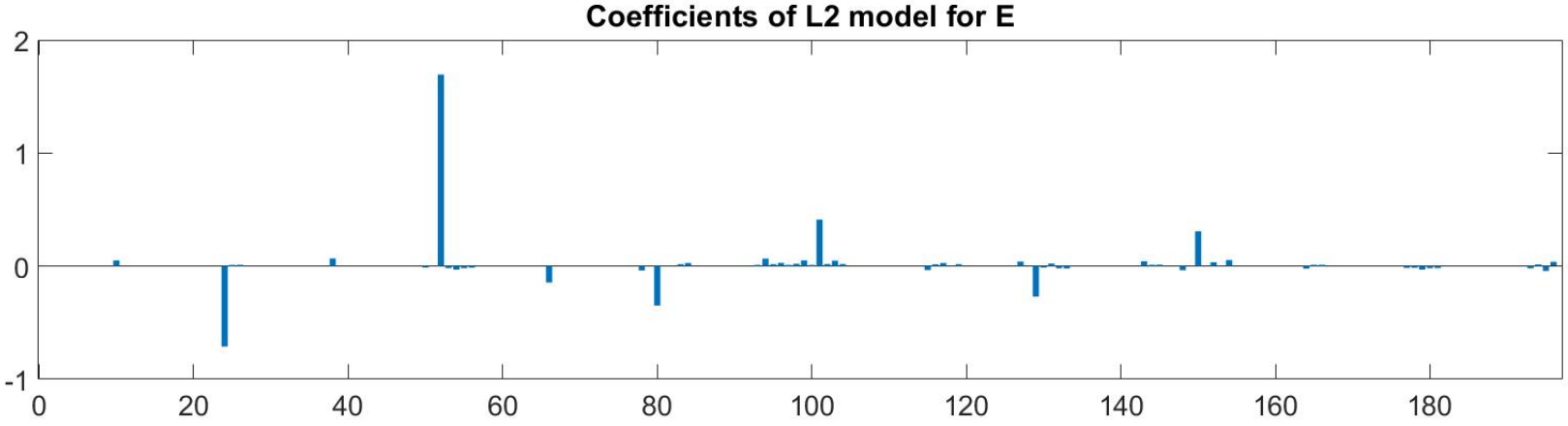
Models	relative errors		# of nonzero terms	
	H	E	H	E
L_1 model	0.0015	0.0017	2	2
L_2 model	0.377	0.363	93	71

Experiment 2: Coefficients of H_N

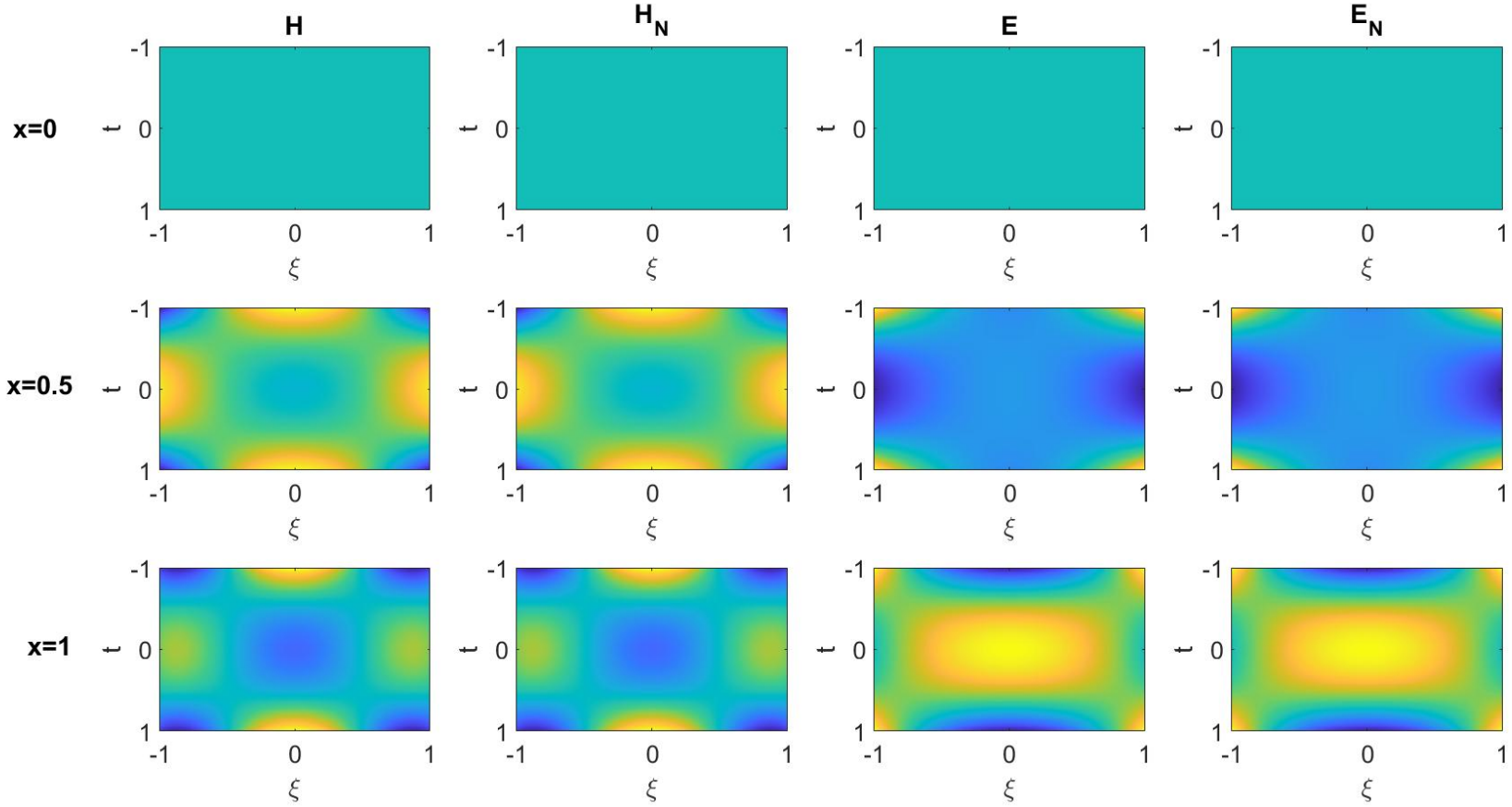


Experiment 2: Coefficients of E_N

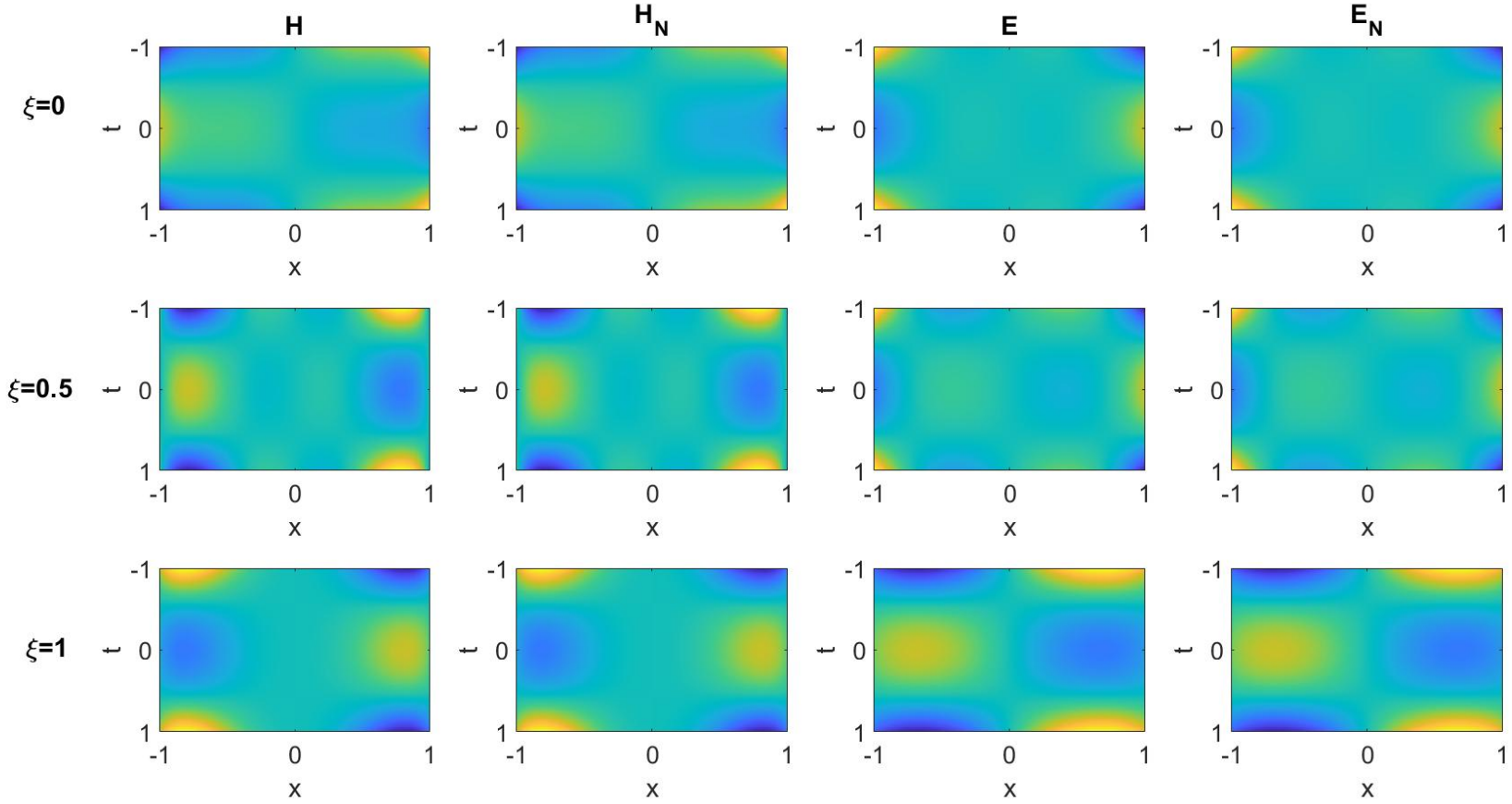




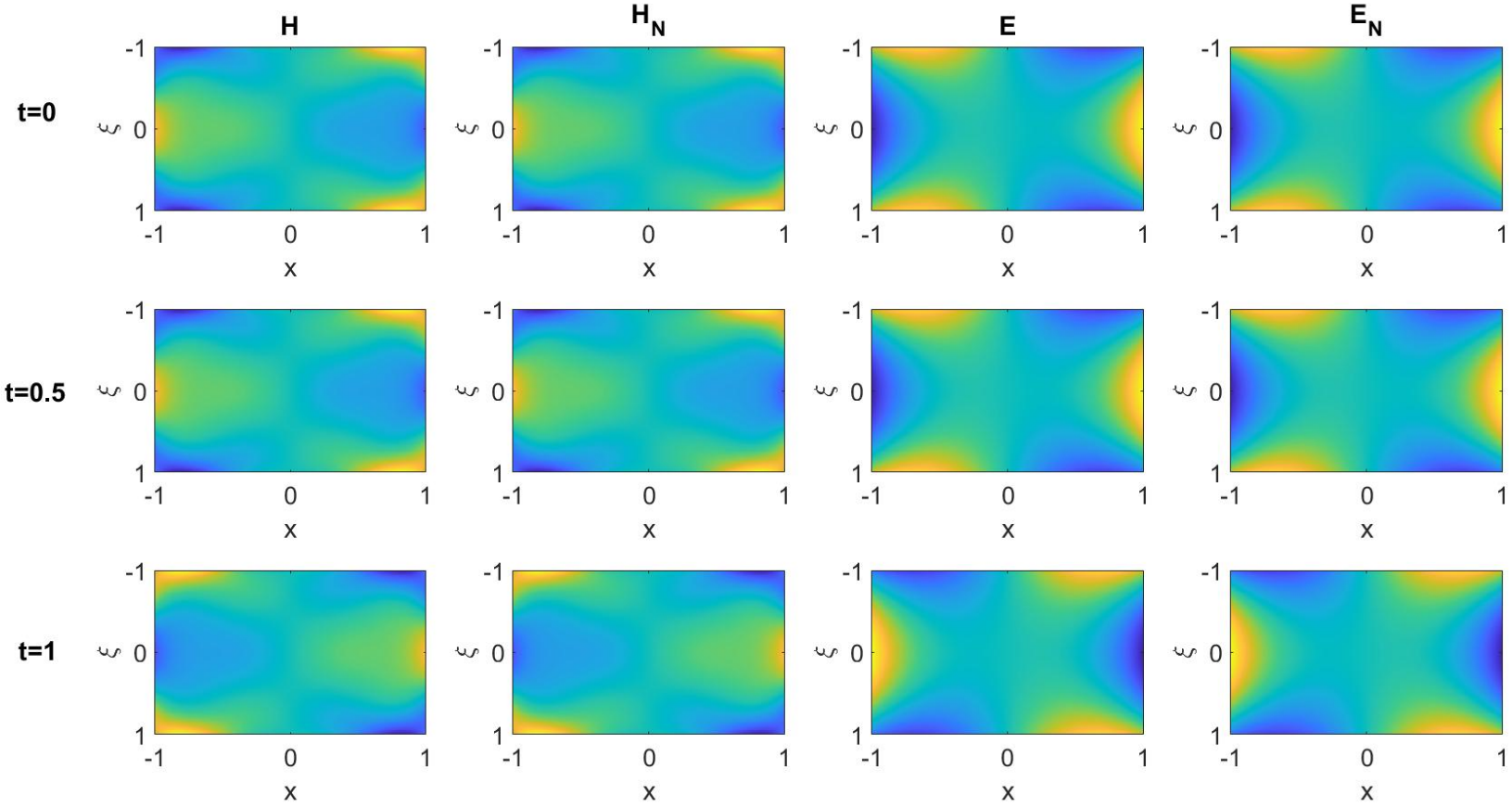
Experiment 2: the L_1-Model, x-Slices



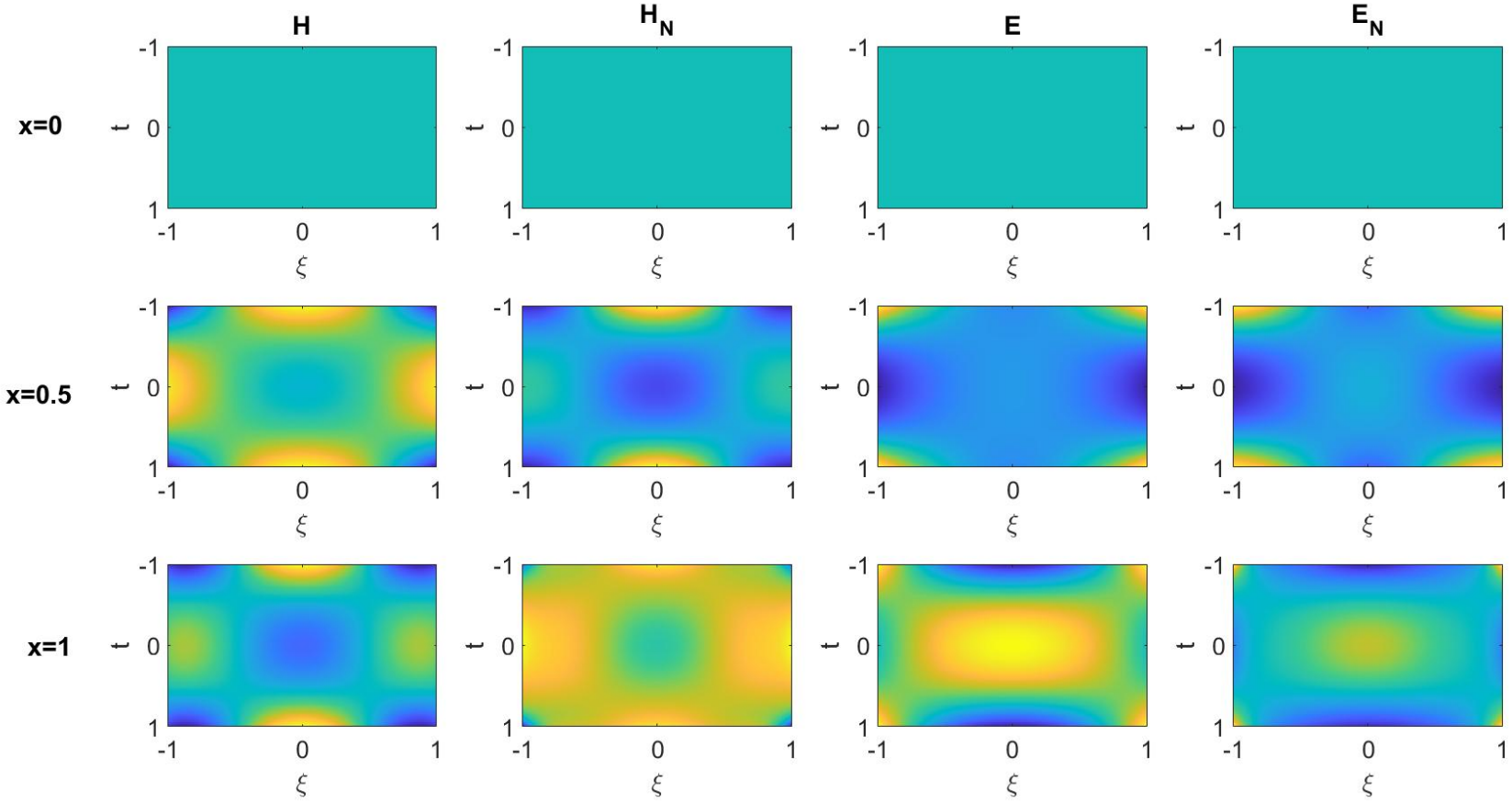
Experiment 2: the L_1-Model, ξ-Slices



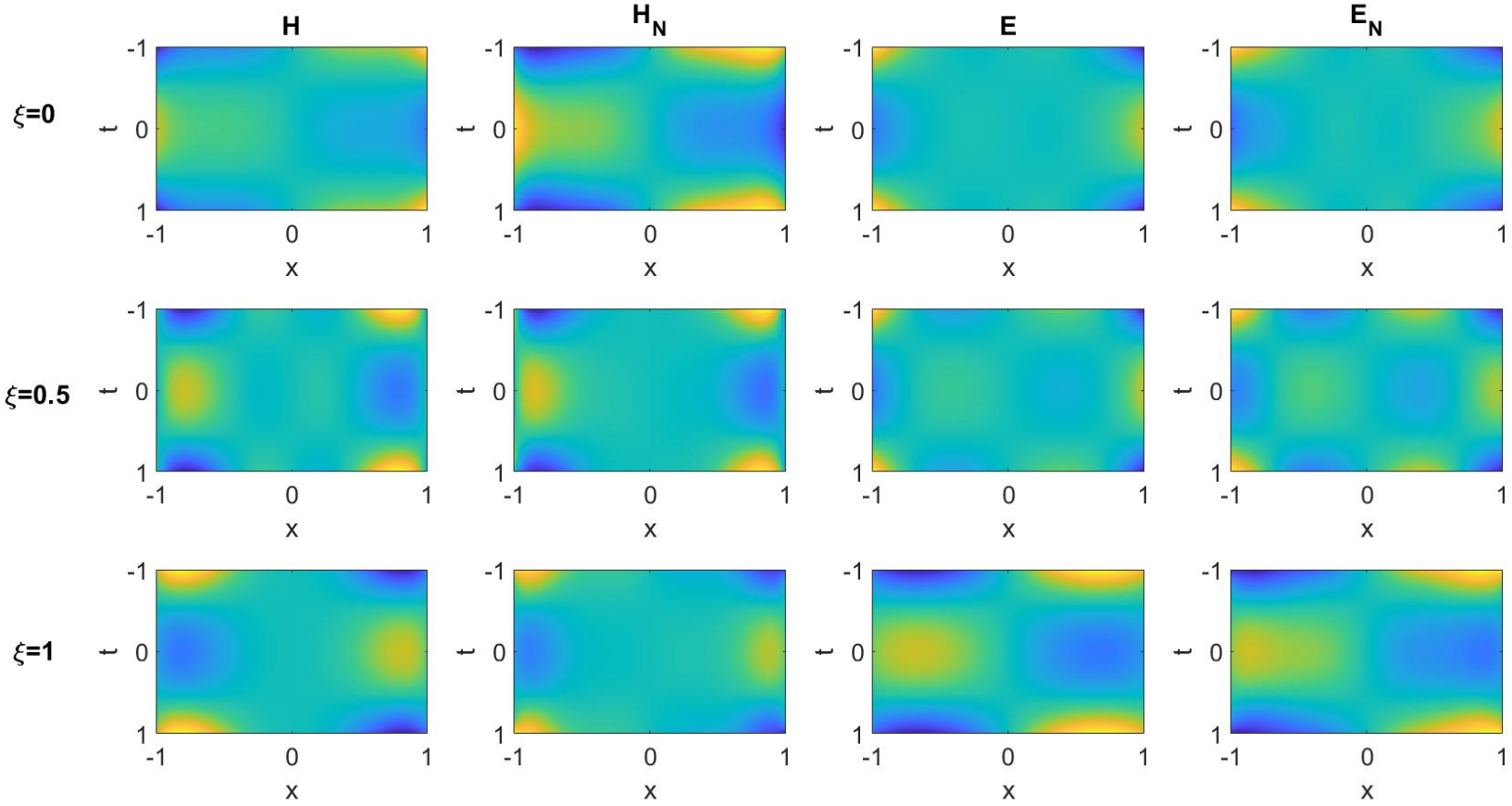
Experiment 2: the L_1-Model, t-Slices



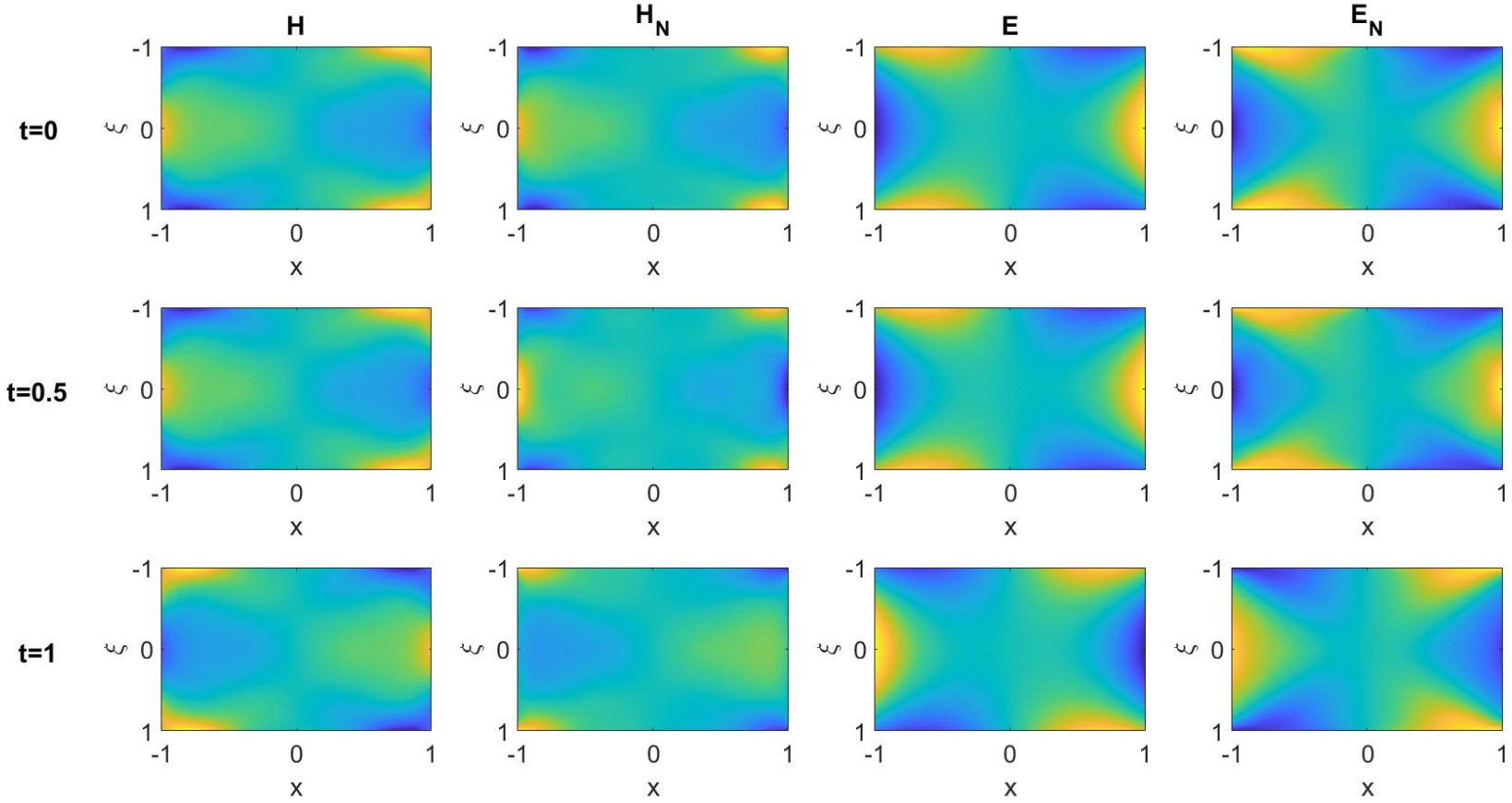
Experiment 2: the L_2-Model, x-Slices



Experiment 2: the L_2-Model, ξ-Slices



Experiment 2: the L_2-Model, t-Slices



Experiment 3: H and E from $\mathbb{X}$

- We choose H and E from $\mathbb{X}$:

$$H(x,\xi,t) = P_4(\xi)U_7(x)P_2(t)\sqrt{1-x^2} - 2P_2(\xi)P_1(x)P_2(t),$$
$$E(x,\xi,t) = -P_0(\xi)U_3(x)P_2(t)\sqrt{1-x^2} + 2P_2(\xi)P_1(x)P_2(t),$$

  where $x, \xi, t \in (0,1)$, and U_n denotes the second kind Chebyshev polynomial of order n.
- Clearly, H and E are not in $\mathbb{X}_N$, for all $n = 2, 4, \ldots$.

Experiment 3: Settings and Data

- We sample $\{x_i, x_i^*, \xi_i, \xi_i^*, t_i : i = 1, 2, \ldots, 150\}$ uniformly from $(0, 1)$.
- We then add gaussian noise $N(0, 0.01^2)$ to $\{(\mathcal{H}(\xi_i, t_i), \mathcal{E}(\xi_i^*, t_i^*), H(x_i, 0, 0)) : i = 1, 2, \ldots, 150\}$. We let $\mathcal{D}$ to denote this dataset.
- We choose 100 data points in $\mathcal{D}$ as our training set, and use the remaining 50 data points in $\mathcal{D}$

$$\left\{ \left(\hat{\mathcal{H}}(\xi_i, t_i), \hat{\mathcal{E}}(\xi_i^*, t_i^*), \hat{H}(x_i, 0, 0) \right) : i = 101, 102, \ldots, 150 \right\},$$

  as the testing set.
- We reconstruct H_6 and E_6 in $\mathbb{X}_6$ using our method.
- For evaluation, we use the test relative error defined by

$$\frac{\sum_{i=101}^{150} \left(|\hat{\mathcal{H}}(\xi_i, t_i) - \mathcal{H}_6(\xi_i, t_i)|^2 + |\hat{H}(x_i, 0, 0) - H_6(x_i, 0, 0)|^2 \right)}{\sum_{i=101}^{150} \left(|\hat{\mathcal{H}}(\xi_i, t_i)|^2 + |\hat{H}(x_i, 0, 0)|^2 \right)},$$

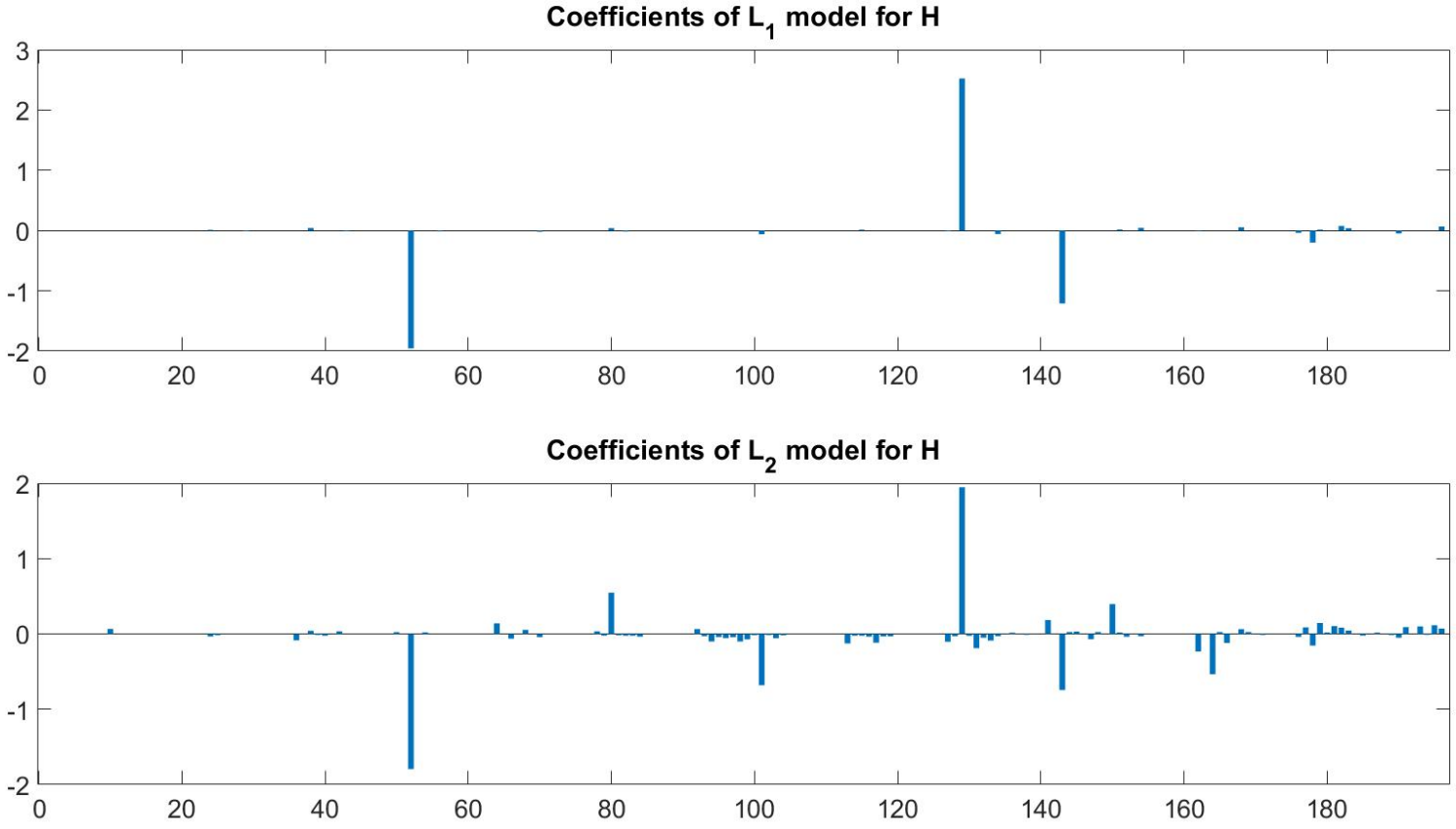
  where $\mathcal{H}_6$ is obtained by substituting H_6 into (5.1).

Experiment 3: Numerical Results

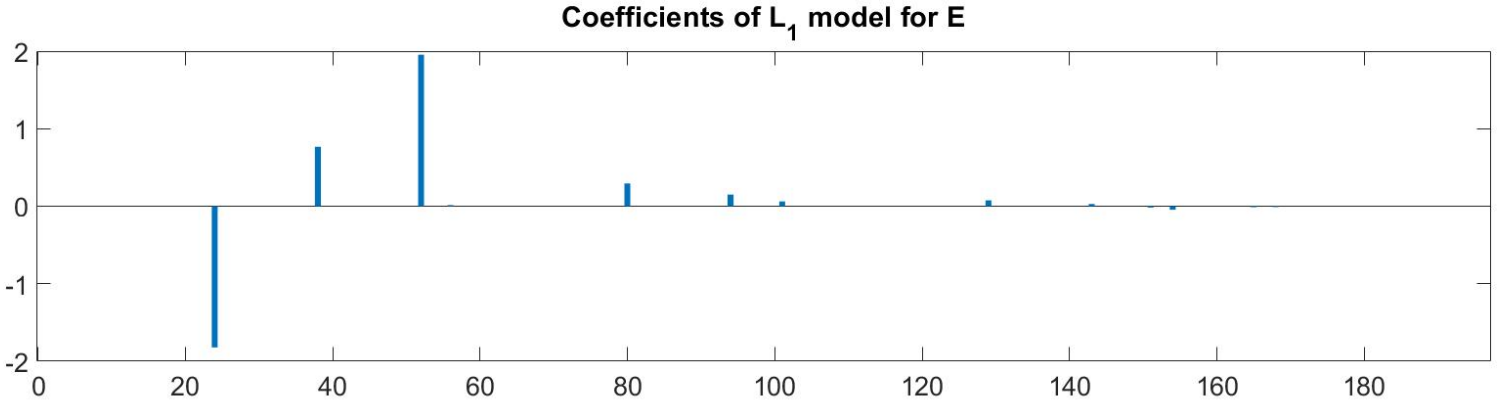
The best regularization parameters for both the L_1 and L_2 models are 0.01.

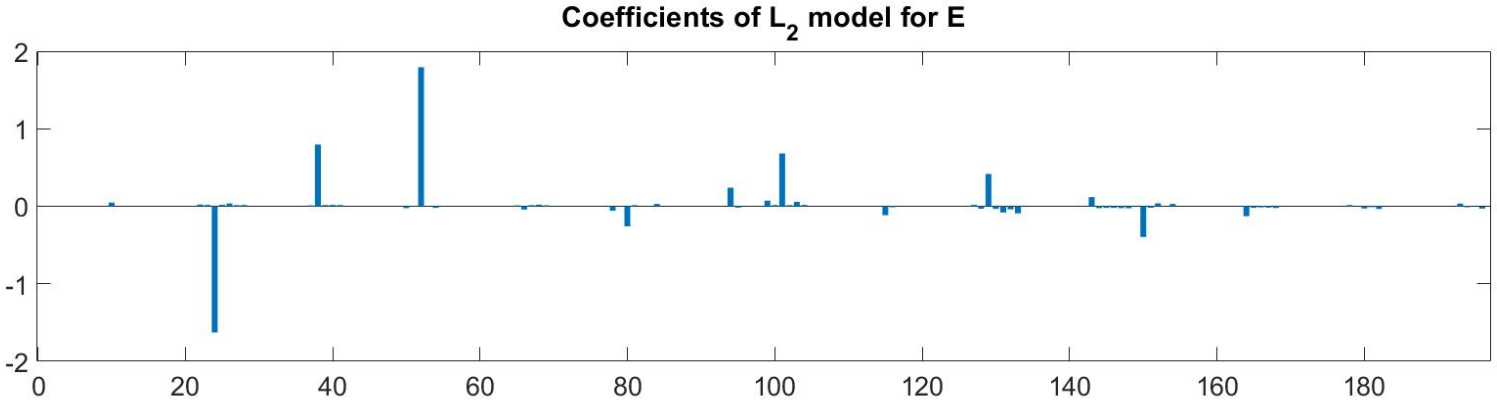
Models	test relative errors		# of nonzero terms	
	H	E	H	E
L_1 model	0.053	0.015	37	16
L_2 model	0.079	0.037	73	65

Experiment 3: Coefficients of H_N

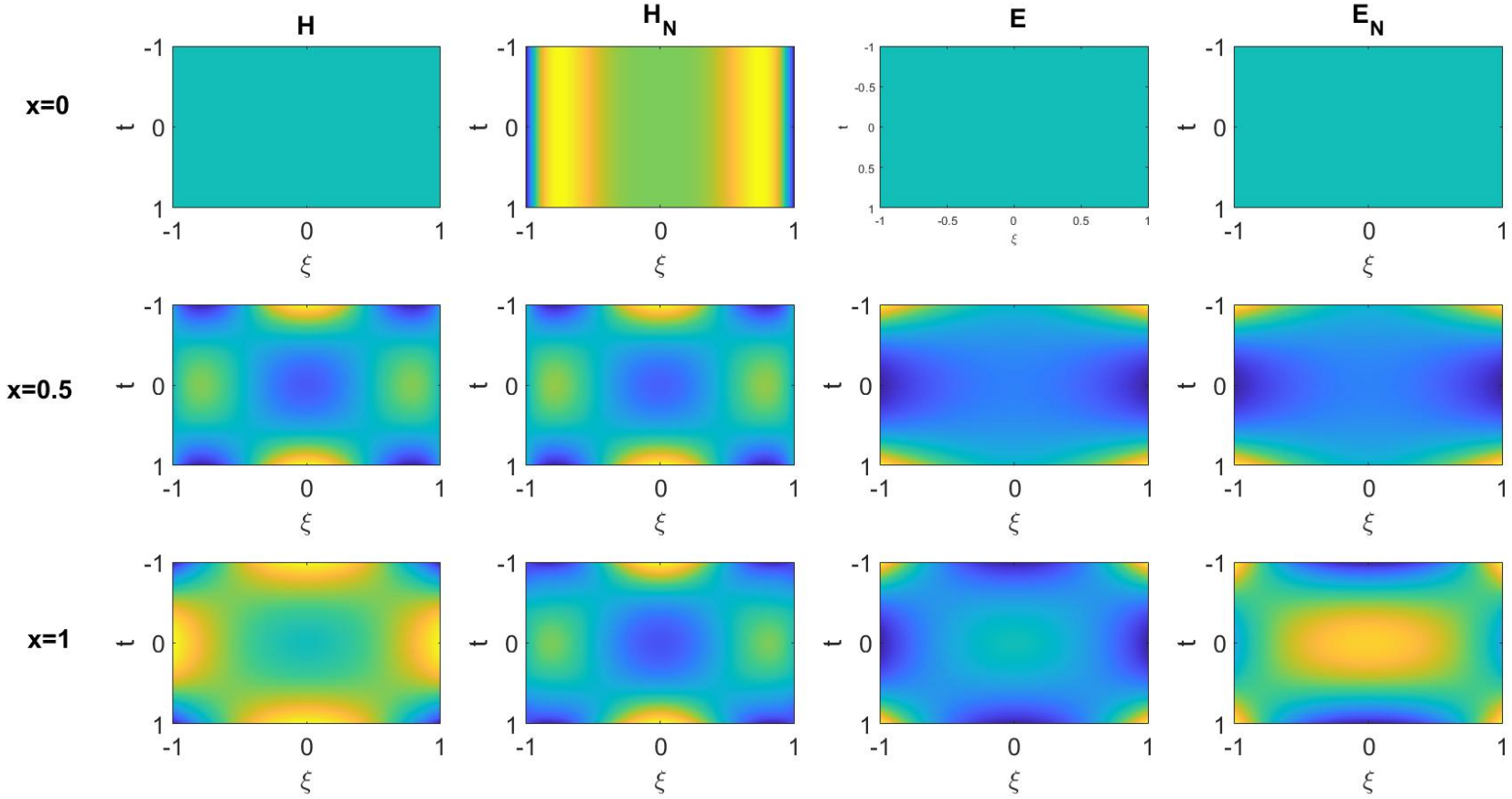


Experiment 3: Coefficients of E_N

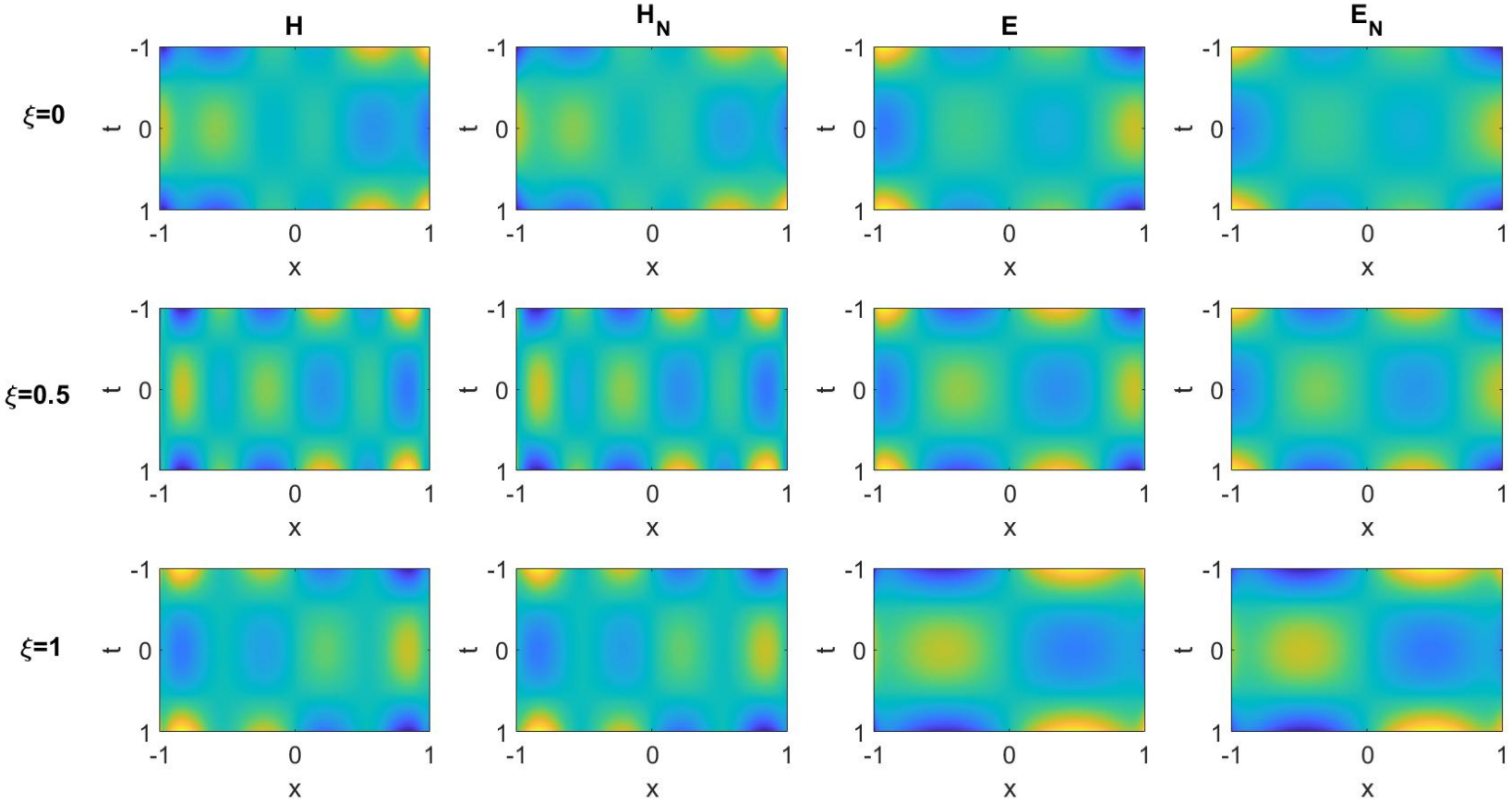




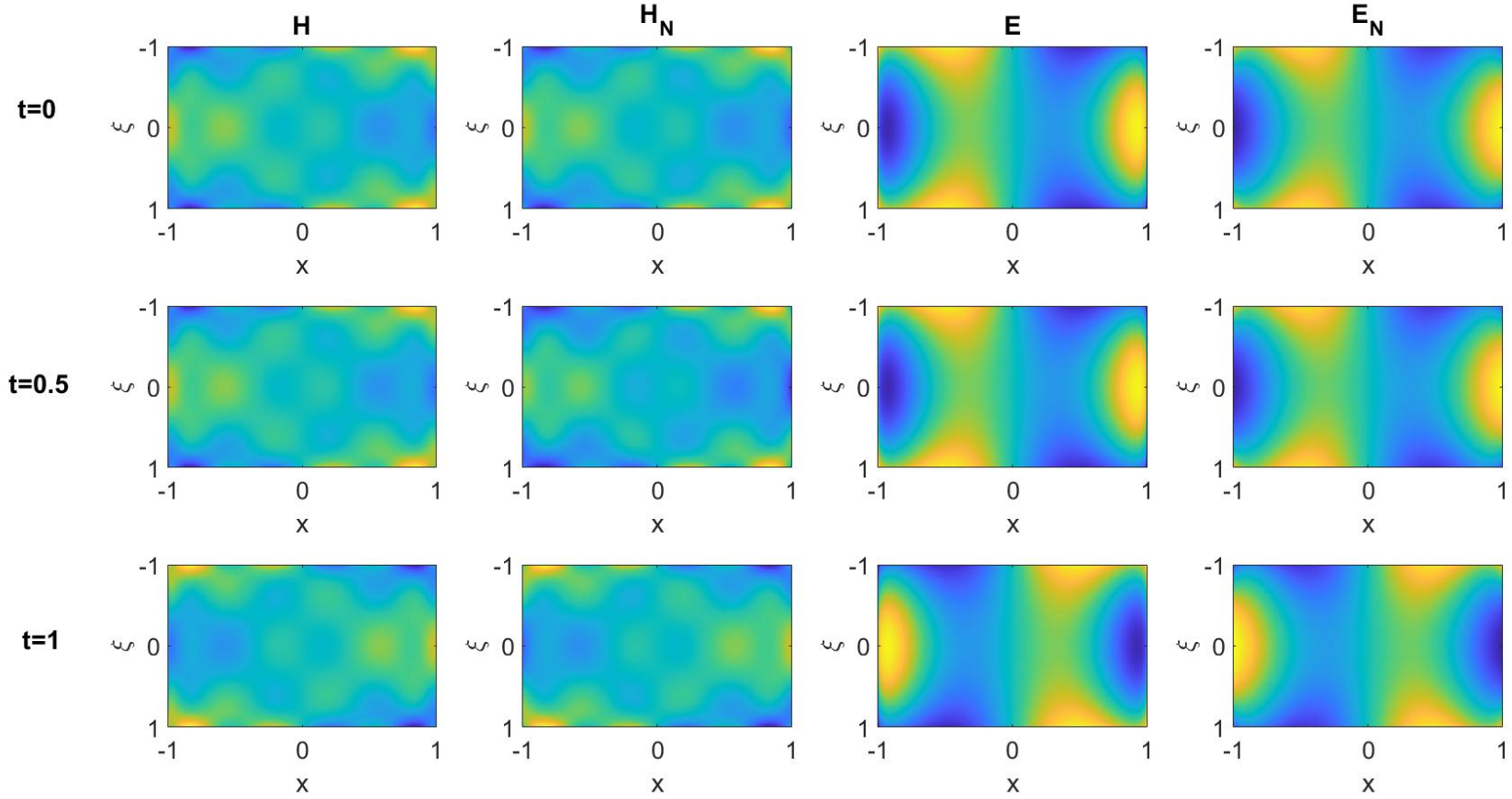
Experiment 3: the L_1 Model, x-Slices



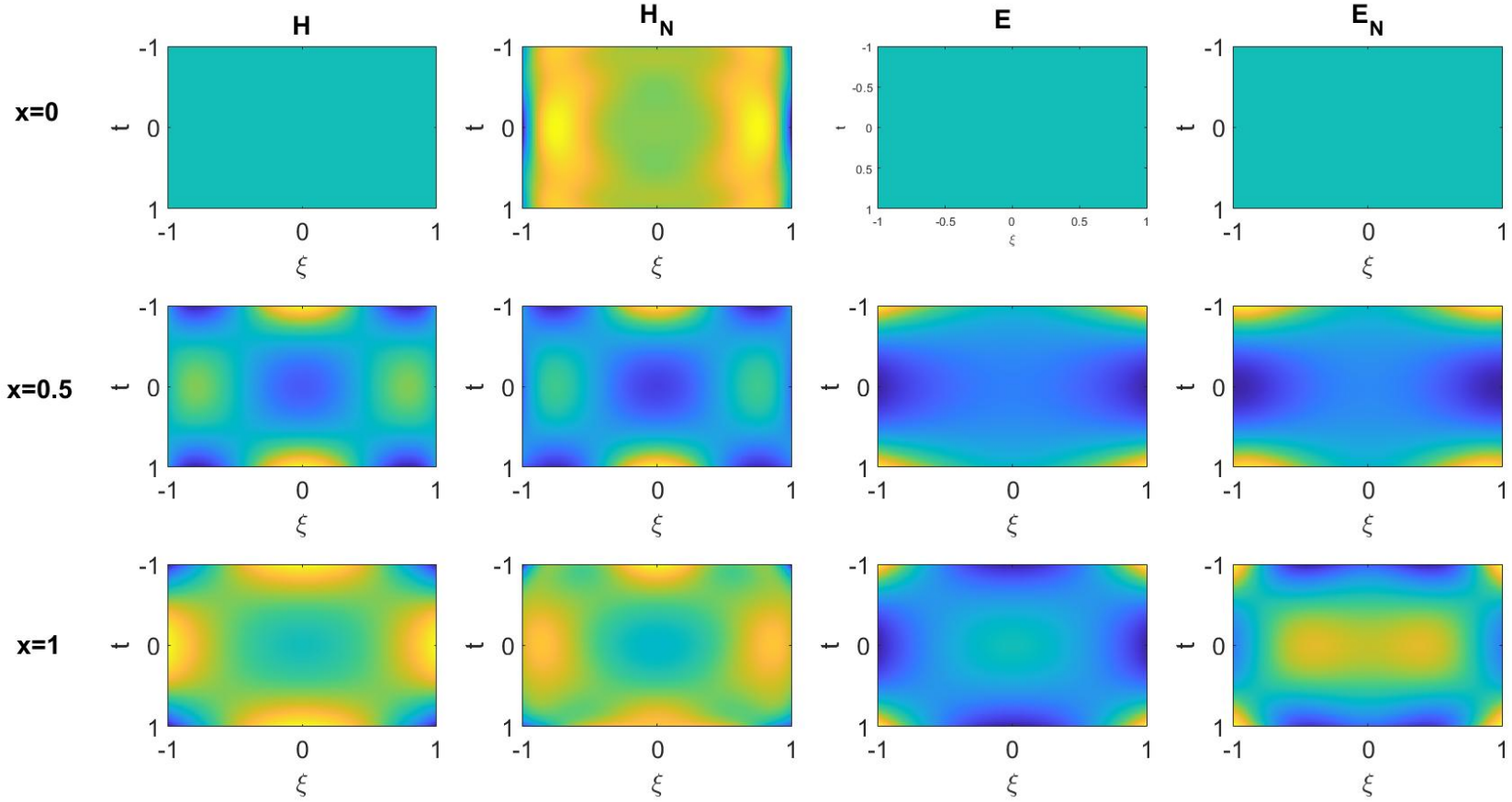
Experiment 3: the L_1 Model, ξ-Slices



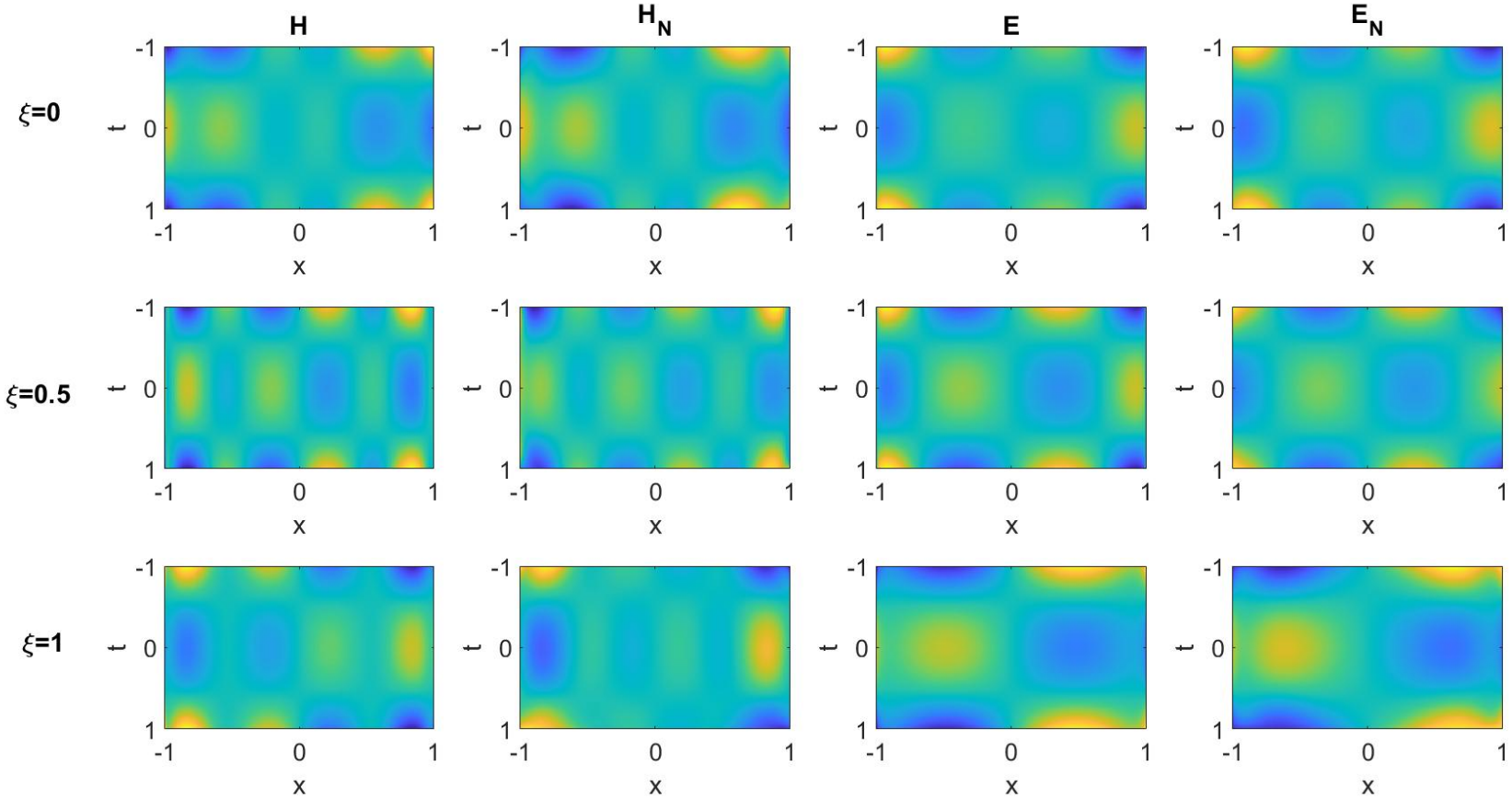
Experiment 3: the L_1 Model, t-Slices



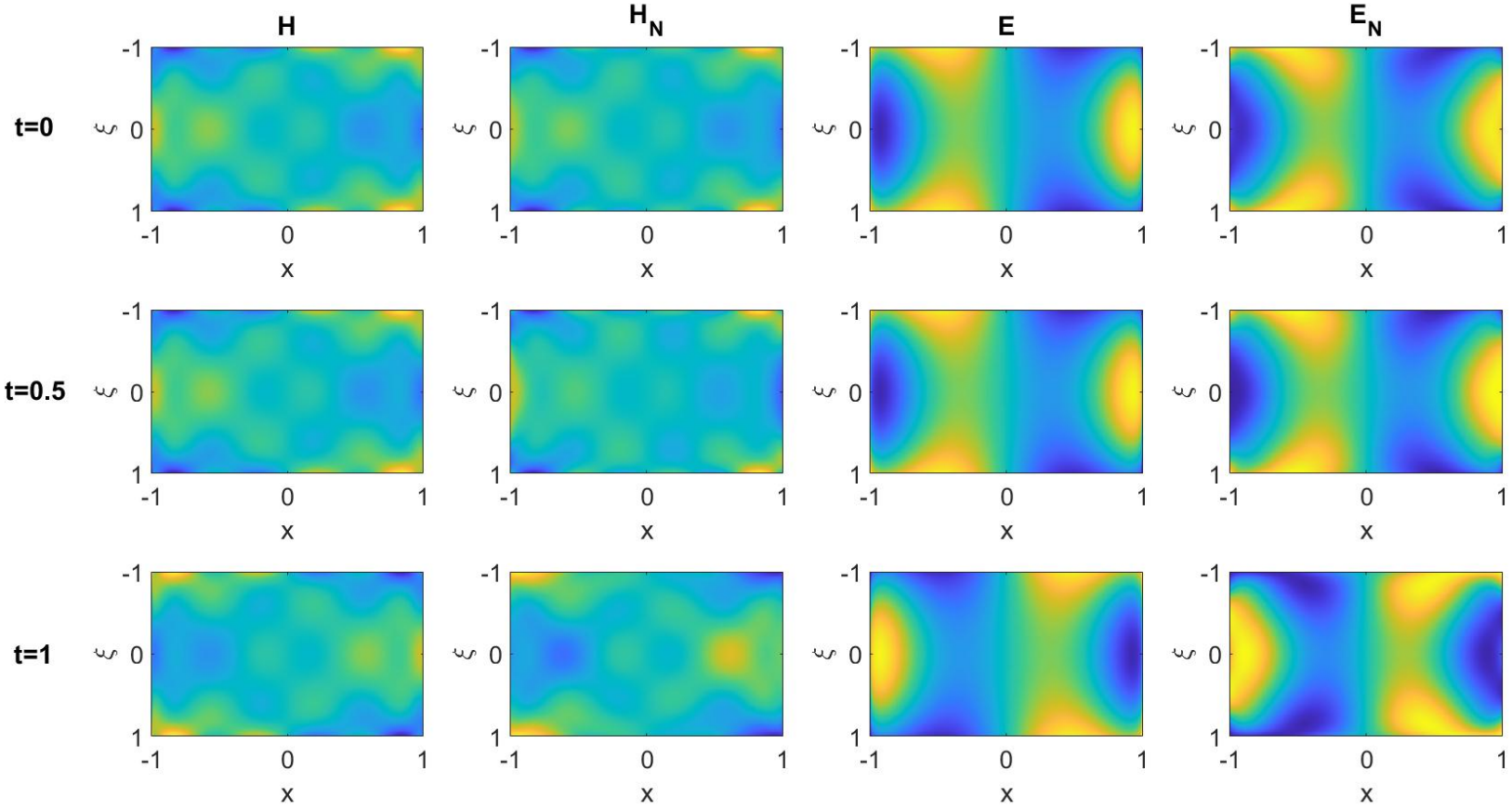
Experiment 3: the L_2 Model, x-Slices



Experiment 3: the L_2 Model, ξ-Slices



Experiment 3: the L_2 Model, t-Slices



Future Work

- We shall consider other constraints in the GPD system, in addition to the moment and initial constraints.
- We shall study the noise distribution of real nuclear physics data and use it to formulate appropriate fidelity term.
- We shall investigate prior information of the solution of GPD and use it to choose appropriate regularization terms.
- We shall implement our proposed method for real nuclear physics data.

Thanks!