Landscape Net

An Application of the Landscape Theory in Machine Learning for Eigenvalue Counting

Li Chen
Joint work with Wei Wang

MIT

May 6, 2020

Code available at
`https://github.com/nehcili/Wave-Localization`

Introduction

Experimental Setting

- $L^2(\mathbb{Z} \cap [1, 1000])$
- $-\Delta_{\mathbb{Z}} + V$ with periodic boundary conditions
- $V(x)$ are iid distributions for $x = 1, 2, 3, ..., 1000$
- $V(x) \geq 0$

Objective: approximate the eigenvalue counting

$$N_V(E) := \# \text{ of eigenvalues of } -\Delta_{\mathbb{Z}} + V \text{ less or equal to } E$$

Introduction

Method: We develop a neural net, the **Landscape Net**, to generalize landscape box counting.

Why?

Introduction

Method: We develop a neural net, the **Landscape Net**, to generalize landscape box counting.

Why?

- David-Filoche-Mayboroda box counting $N^W(E)$

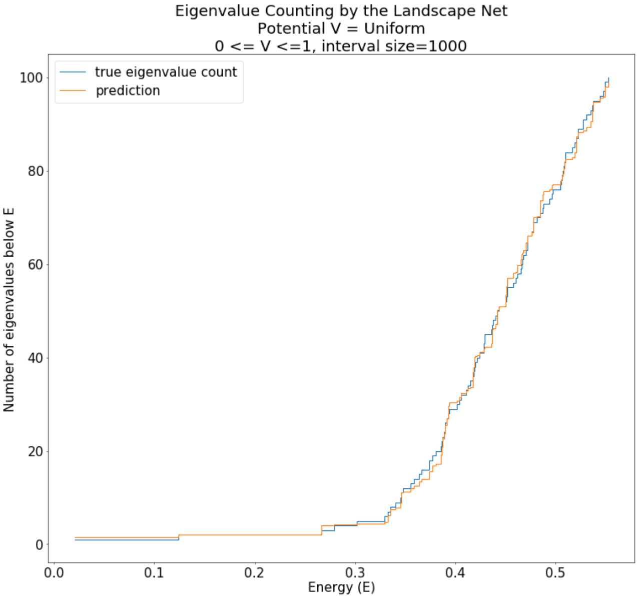
$$c_1 N^W(c_2 E) \leq N_V(E) \leq C_3 N^W(C_4 E)$$

  where $W = 1/u$ is the landscape potential.
  (Discrete case: Arnold-Filoche-Mayboroda-Wang-Zhang)
- More than just inequalities, box counting works well as a regression model with 2 parameters.

$$\text{predicted eigenvalue count} = c_1 N^W(c_2 E)$$

- Under fitting. It only has 2 parameters.

Results


Eigenvalue Counting by the Landscape Net
Potential V = Uniform
0 <= V <=1, interval size=1000
true eigenvalue count
prediction
Number of eigenvalues below E
Energy (E)
0
20
40
60
80
100
0.0
0.1
0.2
0.3
0.4
0.5

Results

- Testing on the bottom 10% of eigenvalues (100/1000 eigenvalues) from 30 samples,

$$\langle(\text{predicted } N_V(E) - \text{true } N_V(E))^2\rangle \approx 3$$

- Time complexity

Data generation*	Training*	Predicting one $N_V(E)$
0.0012s per eigenvalue	$\approx$ 1 hr	0.0025s

*Training data = bottom 10% of eigenvalues (100/1000 eigenvalues) from 300 samples.

Box counting as a neural net layer

Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

$$\boxed{6}\boxed{5}\boxed{2}\boxed{1}\boxed{7}\boxed{3} \quad = \mathrm{W}$$

Box counting as a neural net layer

Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

6	5	2	1	7	3	= W
-2	-1	2	3	-3	1	= E-W

Box counting as a neural net layer

Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

$$\boxed{6}\boxed{5}\boxed{2}\boxed{1}\boxed{7}\boxed{3} \quad = \mathrm{W}$$

$$\boxed{-2}\boxed{-1}\boxed{2}\boxed{3}\boxed{-3}\boxed{1} \quad = \mathrm{E\text{-}W}$$

$$\boxed{-1} \quad = \sup(\mathrm{E\text{-}W}) \text{ on } Q\text{'s}$$

Box counting as a neural net layer

Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

6	5	2	1	7	3	$=$ W
-2	-1	2	3	-3	1	$=$ E-W
		-1	3			$=$ sup(E-W) on Q's

Box counting as a neural net layer

Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

$$\begin{array}{|c|c|c|c|c|c|}\hline 6 & 5 & 2 & 1 & 7 & 3 \\ \hline\end{array} = \text{W}$$

$$\begin{array}{|c|c|c|c|c|c|}\hline -2 & -1 & 2 & 3 & -3 & 1 \\ \hline\end{array} = \text{E-W}$$

$$\begin{array}{|c|c|c|}\hline -1 & 3 & 1 \\ \hline\end{array} = \text{sup(E-W) on } Q\text{'s}$$

Box counting as a neural net layer

Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

6	5	2	1	7	3	$=$ W
-2	-1	2	3	-3	1	$=$ E-W
-1	3	1				$=$ sup(E-W) on Q's
0	1	1				$=$ apply Heaviside Θ

Box counting as a neural net layer

Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

6	5	2	1	7	3	= W
-2	-1	2	3	-3	1	= E-W

-1	3	1	= sup(E-W) on Q's
0	1	1	= apply Heaviside Θ

2	= sum

Box counting as a neural net layer

Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

	$\boxed{6\,|\,5\,|\,2\,|\,1\,|\,7\,|\,3}$		$=$ W
	$\boxed{-2\,|\,-1\,|\,2\,|\,3\,|\,-3\,|\,1}$		$=$ E-W
grader $sup_Q(E - W)$	$\boxed{-1\,|\,3\,|\,1}$		$=$ sup(E-W) on Q's
activation Θ	$\boxed{0\,|\,1\,|\,1}$		$=$ apply Heaviside Θ
(weighted) sum $w_Q = 1$	$\boxed{2}$		$=$ sum

Box counting as a neural net layer

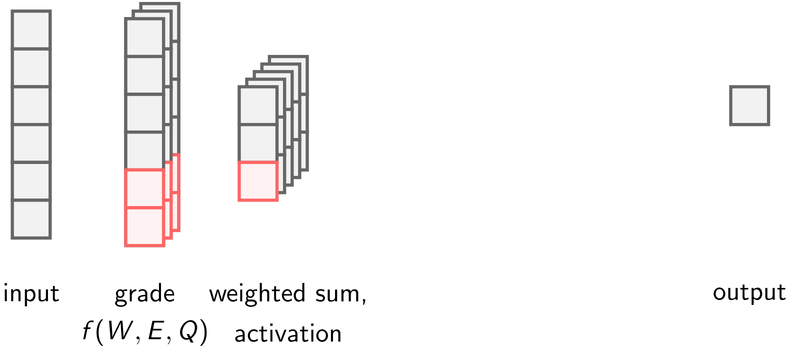
Let us consider $E = 4$ and each box Q has length $= 2$ for illustration purpose.

	$\boxed{6}\boxed{5}\boxed{2}\boxed{1}\boxed{7}\boxed{3}$	= W
	$\boxed{-2}\boxed{-1}\boxed{2}\boxed{3}\boxed{-3}\boxed{1}$	= E-W
grader $f(W, E, Q)$	$\boxed{-1}\boxed{3}\boxed{1}$	= sup(E-W) on Q's
activation Θ, $\max(0, \cdot)$, etc.	$\boxed{0}\boxed{1}\boxed{1}$	= apply Heaviside Θ
weighted sum w_Q	$\boxed{2}$	= sum

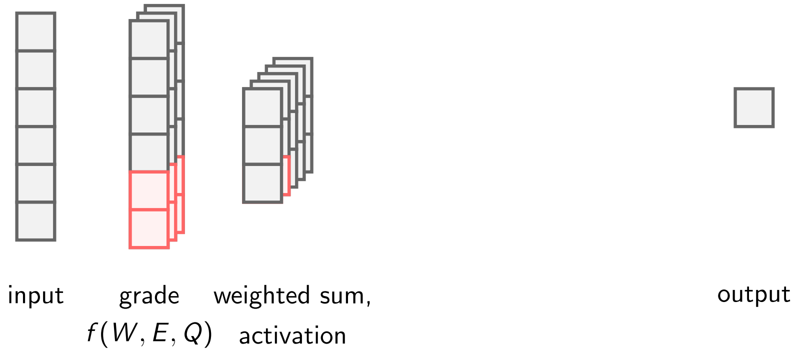
Landscape net architecture

Landscape net architecture

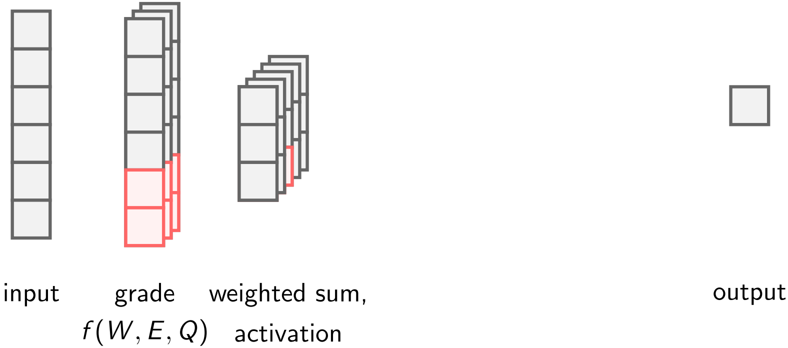
Landscape net architecture



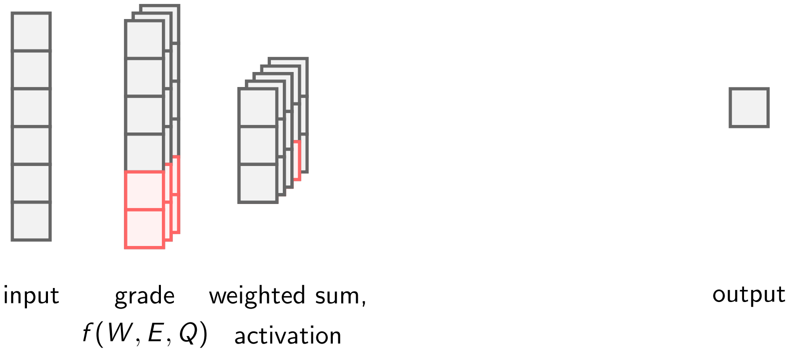
Landscape net architecture



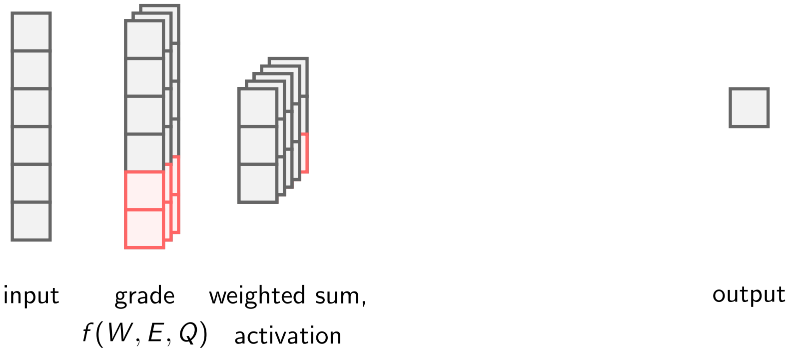
Landscape net architecture



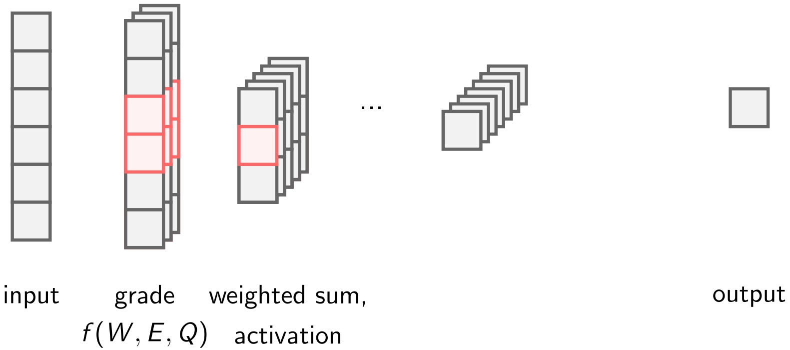
Landscape net architecture



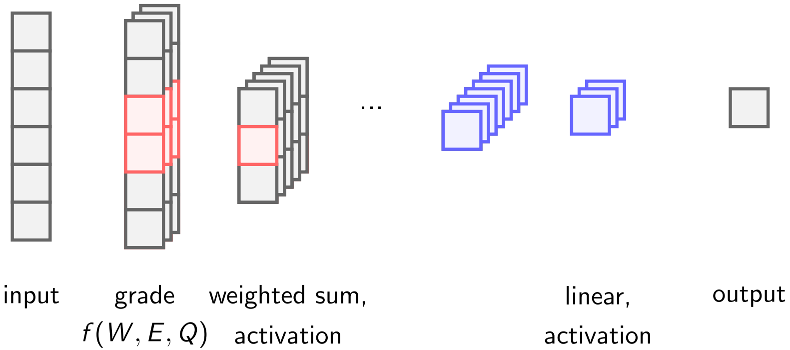
Landscape net architecture



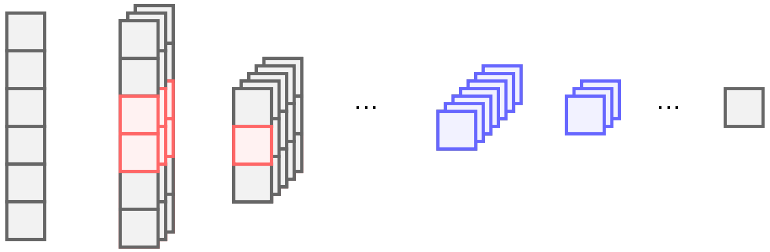
Landscape net architecture



Landscape net architecture



Landscape net architecture



	input	grade	weighted sum,		linear,	output
		$f(W,E,Q)$	activation		activation	
parameters:		10k	60k		3k	

Ansatz for f

- Let ρ be the true density, we would like to have

$$f(W, E, Q) = \int_Q dx\, \rho(x).$$

- In dimension d, the semi-classical expansion ($-\hbar^2\Delta + V$ and $\hbar \to 0$) of ρ is

$$\begin{aligned}\rho(x) =& \hbar^{-d} \int_{\mathbb{R}^d} dp\, 1_{\{p^2 + W(x) - E < 0\}} \text{ (Weyl term)} \\ & + \hbar^{-d+2} \sum_i A_i(V(x), W(x), \nabla W(x)) B_i(E - W(x)) \\ & + O(\hbar^{-d+4}),\end{aligned}$$

for some local functions A_i and B_i.

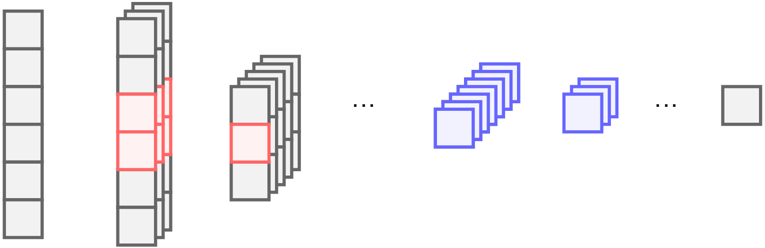
Ansatz for f

- We choose f to have the same form as the first 2 lowest order terms

$$f(W, E, Q) = \sum_{x \in Q} w_x A(V(x), W(x), W'(x)) B(E - W(x))$$

where each A and B are constructed from layers of neural nets and w_x are weights.

Thank you!



input

grade $f(W, E, Q)$

weighted sum, activation

linear, activation

output

parameters: 10k 60k 3k

Semi-classical expansions

For any meromorphic function f which is analytic on a tabular neighborhood of the real line, if $\rho(x) = f(-\hbar^2\Delta + V)(x,x)$ and W is the associated landscape potential, then as $\hbar \to 0$,

$$
\begin{aligned}
\rho = \hbar^{-d} \int_{\mathbb{R}^d} dp\, f(p^2 + W) \\
+\hbar^{-d+2} \Bigg(-\frac{1}{2}\Delta W \int_{\mathbb{R}^d} dp\, f''(p^2 + W) \\
+ 2W^{-1}\Delta W \int_{\mathbb{R}^d} dp\, p^2 f''(p^2 + W) \\
- \frac{1}{3}|\nabla W|^2 \int_{\mathbb{R}^d} dp\, f'''(p^2 + W) \\
+ (2W^{-1}|\nabla W|^2 + \frac{2}{3}\Delta W) \int_{\mathbb{R}^d} dp\, p^2 f'''(p^2 + W) \\
+ \frac{1}{2}|\nabla W|^2 \int_{\mathbb{R}^d} dp\, p^2 f''''(p^2 + W) \Bigg) + O(\hbar^{-d+4})
\end{aligned}
$$

Types of potentials considered

1. constant 1
2. uniform
3. standard exponential
4. poisson
5. power
6. bernoulli
7. chisquare
8. rayleigh
9. laplace
10. normal
11. logistic

Future direction

- Optimization: reduction of parameters, hyperparameter tuning, scalability, training on additional classes of potentials
- Larger domain/higher dimension
- Obtaining eigenvalue counting by using signatures of $H - E$ instead of explicity eigenvalue computation
- Approximating the density ρ itself