Deep learning networks and gravitational wave signal recognition

Zhoujian Cao

Department of Astronomy, BNU

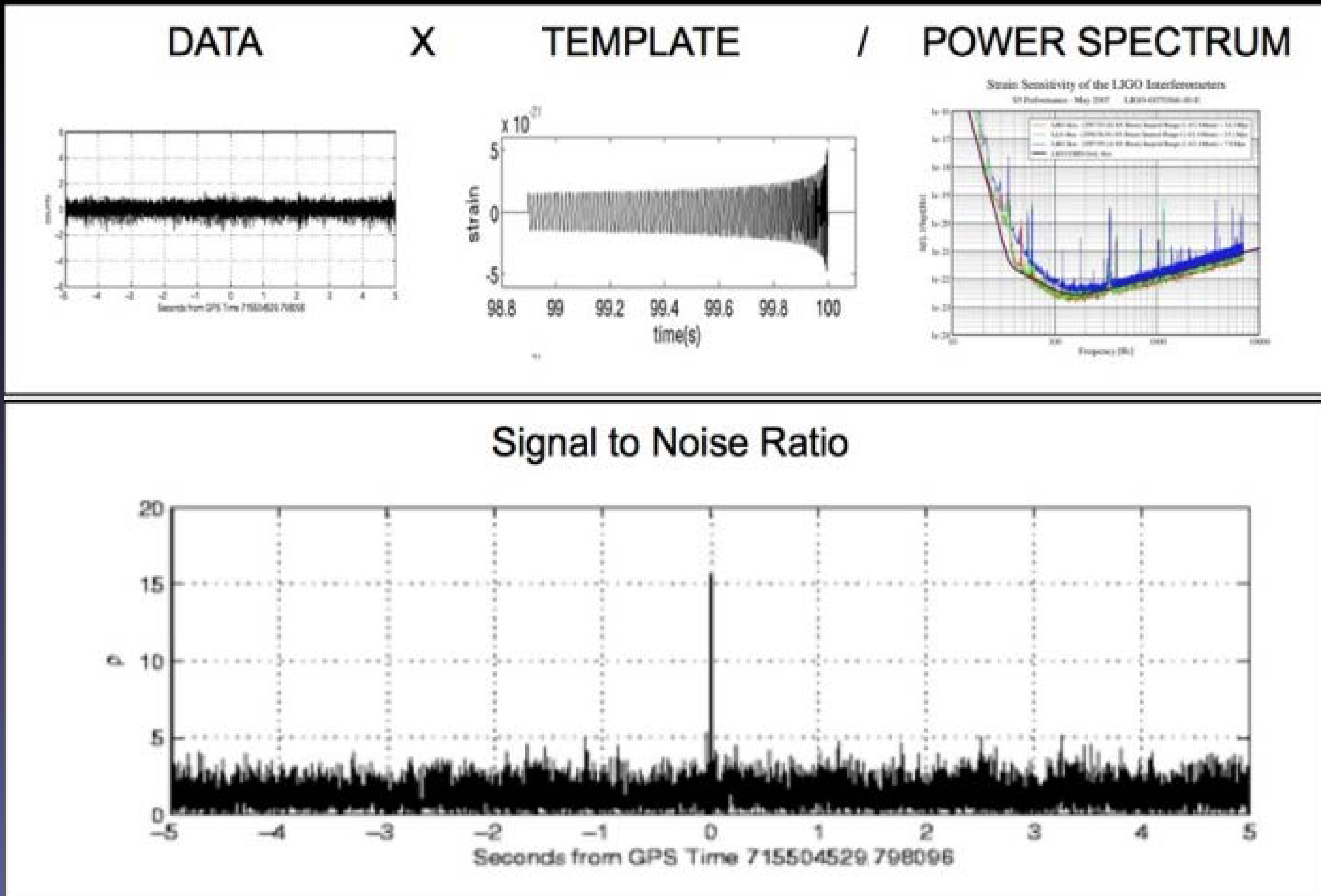
2019-2-15

5th KAGRA International Workshop @ Italy

Event	FAR [y^{-1}] PyCBC	GstLAL	cWB	Network SNR PyCBC	GstLAL	cWB
GW150914	$< 1.53 \times 10^{-5}$	$< 1.00 \times 10^{-7}$	$< 1.63 \times 10^{-4}$	23.6	24.4	25.2
GW151012	0.17	7.92×10^{-3}	–	9.5	10.0	–
GW151226	$< 1.69 \times 10^{-5}$	$< 1.00 \times 10^{-7}$	0.02	13.1	13.1	11.9
GW170104	$< 1.37 \times 10^{-5}$	$< 1.00 \times 10^{-7}$	2.91×10^{-4}	13.0	13.0	13.0
GW170608	$< 3.09 \times 10^{-4}$	$< 1.00 \times 10^{-7}$	1.44×10^{-4}	15.4	14.9	14.1
GW170729	1.36	0.18	0.02	9.8	10.8	10.2
GW170809	1.45×10^{-4}	$< 1.00 \times 10^{-7}$	–	12.2	12.4	–
GW170814	$< 1.25 \times 10^{-5}$	$< 1.00 \times 10^{-7}$	$< 2.08 \times 10^{-4}$	16.3	15.9	17.2
GW170817	$< 1.25 \times 10^{-5}$	$< 1.00 \times 10^{-7}$	–	30.9	33.0	–
GW170818	–	4.20×10^{-5}	–	–	11.3	–
GW170823	$< 3.29 \times 10^{-5}$	$< 1.00 \times 10^{-7}$	2.14×10^{-3}	11.1	11.5	10.8

arXiv: 1811.12907

Data analysis and template



LIGO GW templates

```
--approximant APPROX
Supported TD approximants:
  TaylorT1 (default)
  TaylorT2
  TaylorT3
  TaylorT4
  TaylorEt
  EccentricTD
  IMRPhenomA
  IMRPhenomB
  IMRPhenomC
  IMRPhenomD
  IMRPhenomPv2
  EOBNRv2
  EOBNRv2HM
  SEOBNRv1
  SEOBNRv2
  SEOBNRv3
  SEOBNRv4
  SpinTaylorT4
  SpinTaylorT2
  SpinTaylorT1
  PhenSpinTaylor
  PhenSpinTaylorRD
  SpinDominatedWf
  HGimri
  NR_hdf5
Supported FD approximants:
  IMRPhenomA
  IMRPhenomB
  IMRPhenomC
  IMRPhenomD
  IMRPhenomP
  IMRPhenomPv2
  EOBNRv2_ROM
  EOBNRv2HM_ROM
  SEOBNRv1_ROM_EffectiveSpin
  SEOBNRv1_ROM_DoubleSpin
  SEOBNRv2_ROM_EffectiveSpin
  SEOBNRv2_ROM_DoubleSpin
  TaylorF2
  SpinTaylorF2
  TaylorR2F4
  SpinTaylorT4Fourier
  SpinTaylorT2Fourier
  TaylorF2RedSpin
  TaylorF2RedSpinTidal
```

State of art for EOBNR

- Total mass: $0<M<\infty$
- Mass ratio: $1<q<20$

GW150914 (1.24), LVT151012 (1.77), GW151226 (1.89), GW170104 (1.61), GW170608 (1.71), GW170814 (1.21)

- Spin: $-1<s<0.9$

GW150914 (<0.24), LVT151012 (<0.3), GW151226 (<0.35), GW170104 (<0.42), GW170608 (<0.3), GW170814 (<0.18)

- Precession simplified (BCV)

GW150914, LVT151012 & GW151226 assume no precession, GW170104, GW170608 & GW170814 gain little constrain on precession

- $e = 0$

All announced BBH GW events assumed

GW new Astronomy

- Boundary of GR
- Completely unknown astrophysical objects

Unknown theory; unknown sources !

Crazy templates !

Extremely sensitive detectors

Deep learning and GW

Real time GW signal monitor based DL

- ➢ Requirement
- ✓ Fast
- ✓ Beyond known template

- ➢ Possible solution
- ✓ Transfer time consuming to training stage
- ✓ Generalization of trained data

Deep learning and GW

Physics Letters B 778 (2018) 64–70

Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

ELSEVIER

Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data

Daniel George[a,b,*], E.A. Huerta[b]

[a] Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, United States
[b] NCSA, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, United States

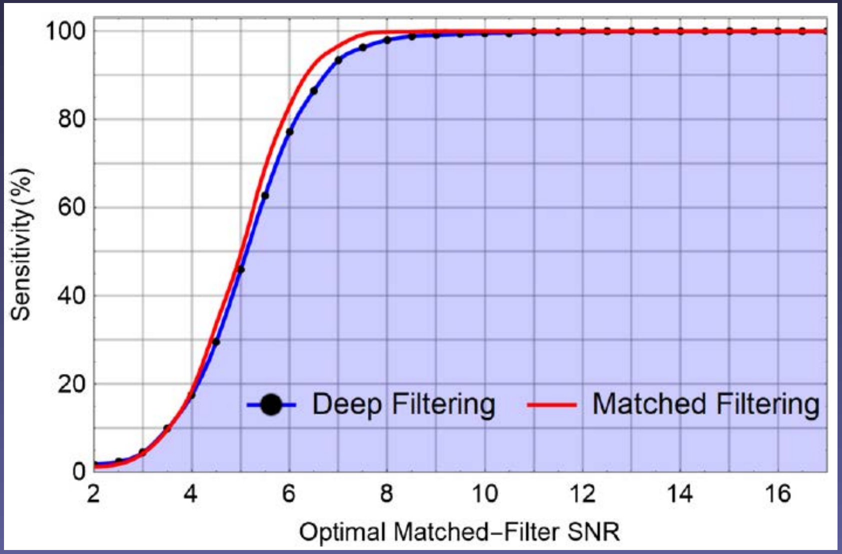
PHYSICAL REVIEW LETTERS **120**, 141103 (2018)

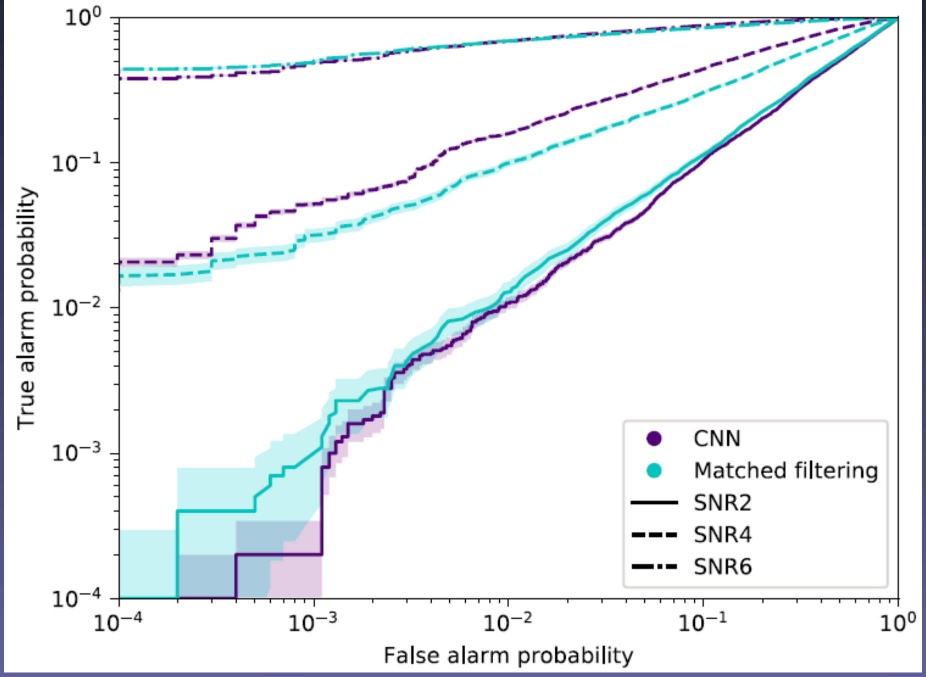
Editors' Suggestion

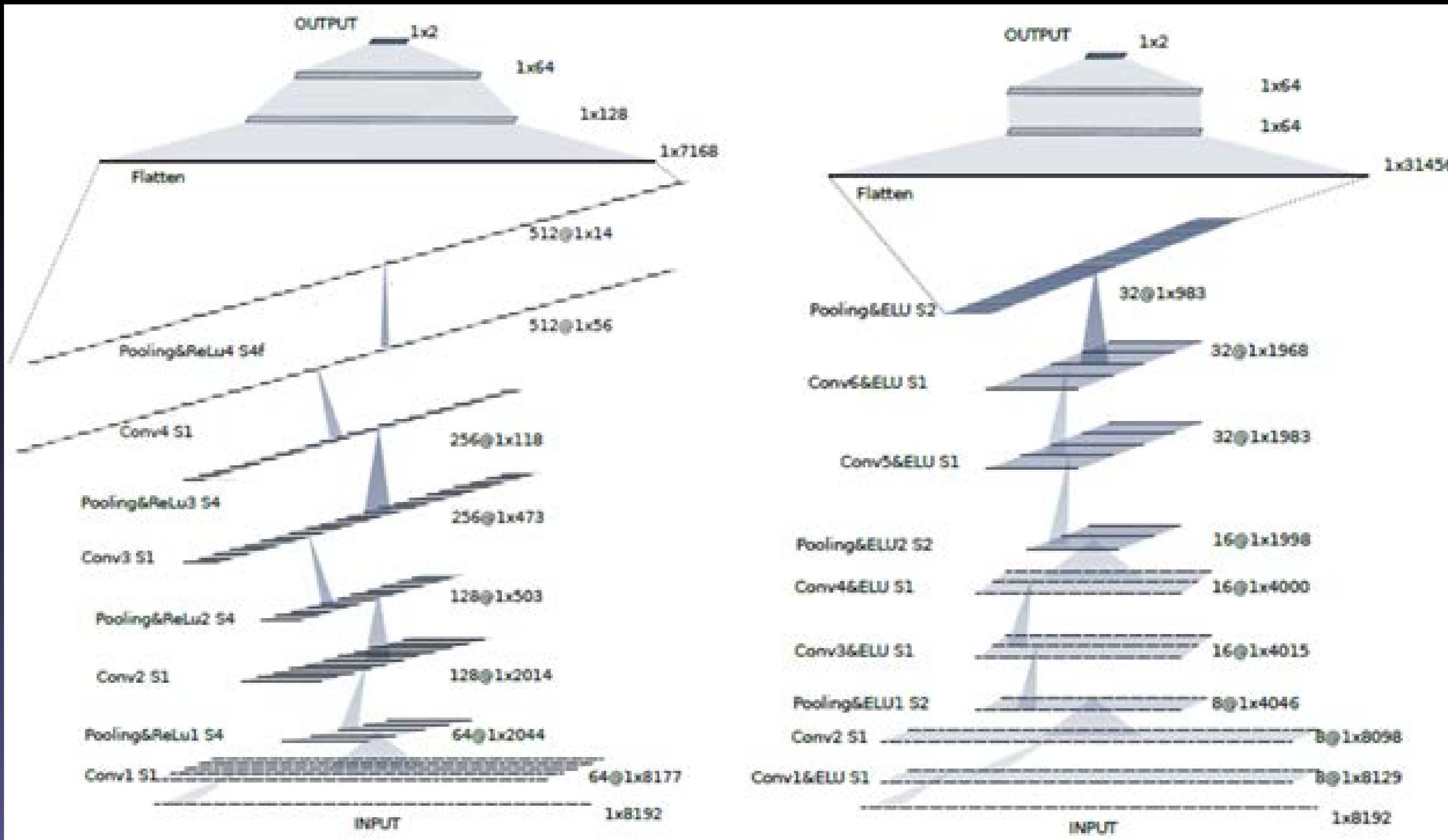
Matching Matched Filtering with Deep Networks for Gravitational-Wave Astronomy

Hunter Gabbard,* Michael Williams, Fergus Hayes, and Chris Messenger
SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom

(Received 16 December 2017; revised manuscript received 12 February 2018; published 6 April 2018)

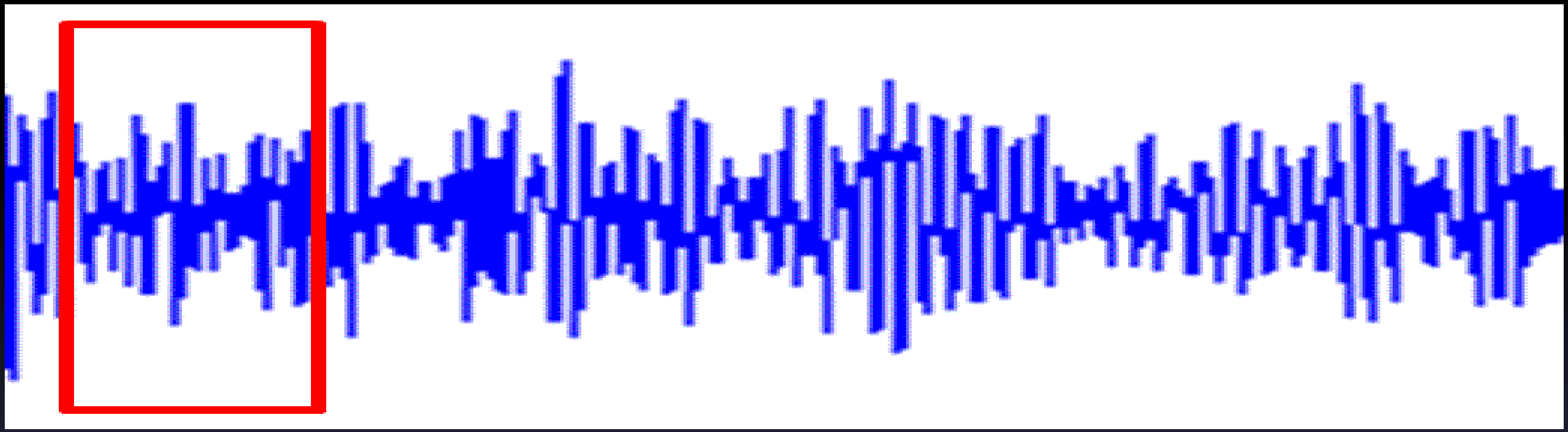




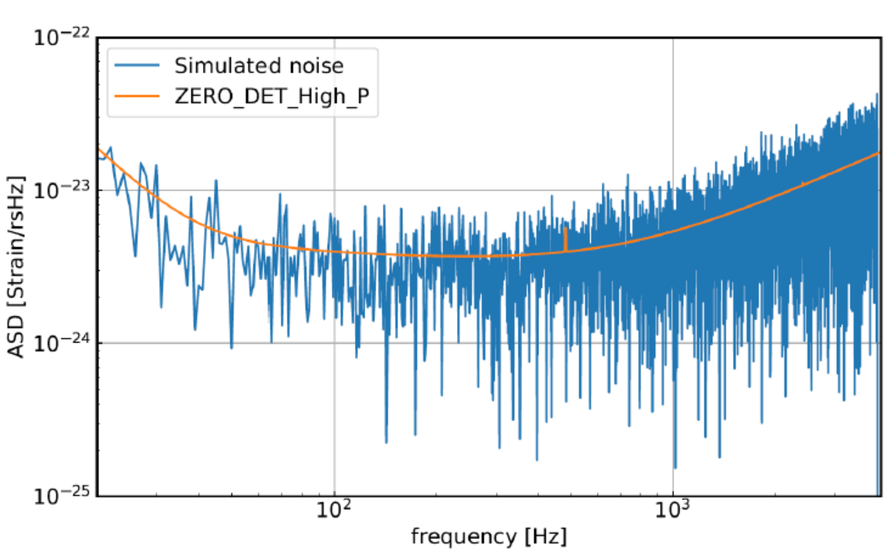


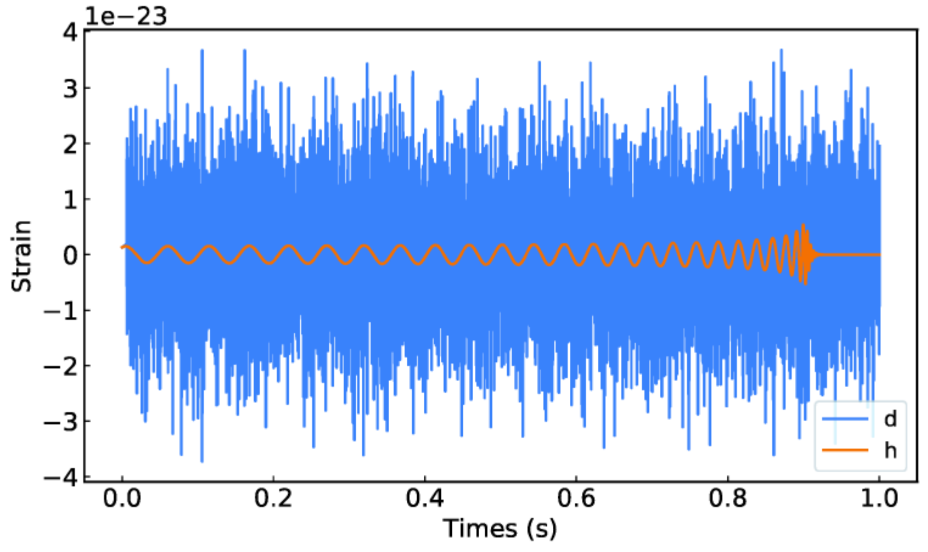
UIUC

Glasgow

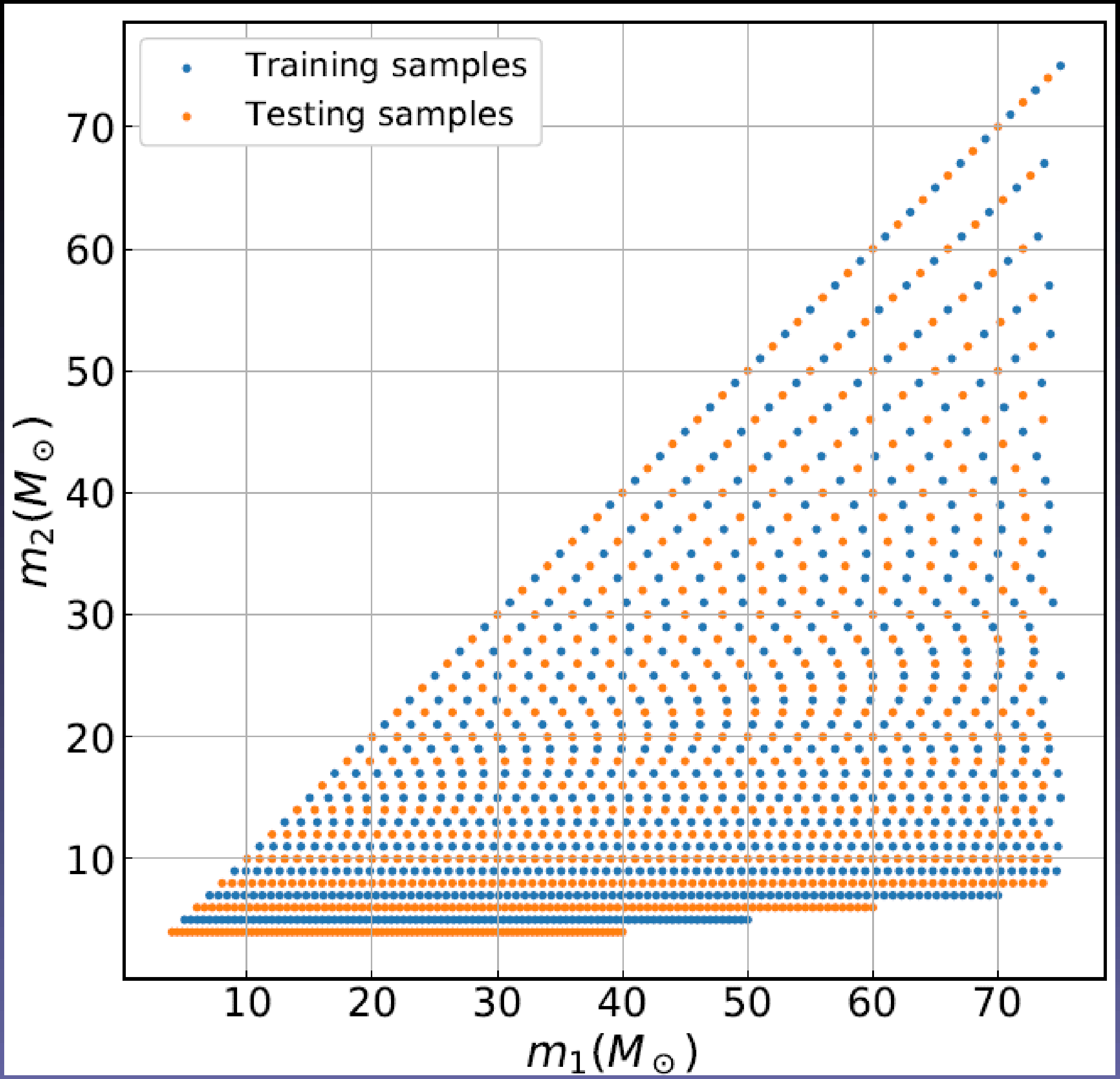


As a real time monitor we consider fixed time duration data segment, sample rate 8192





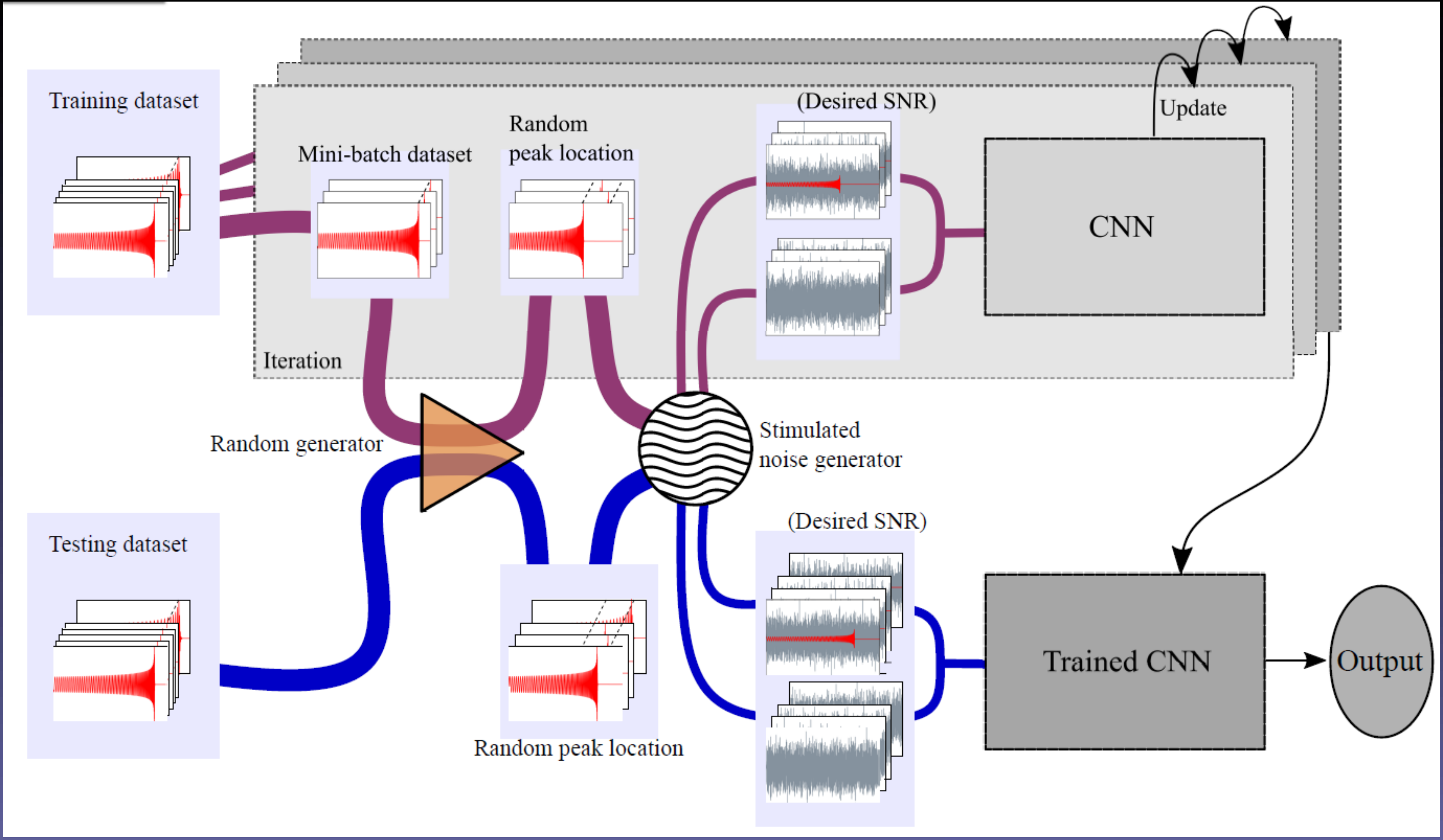
Data samples



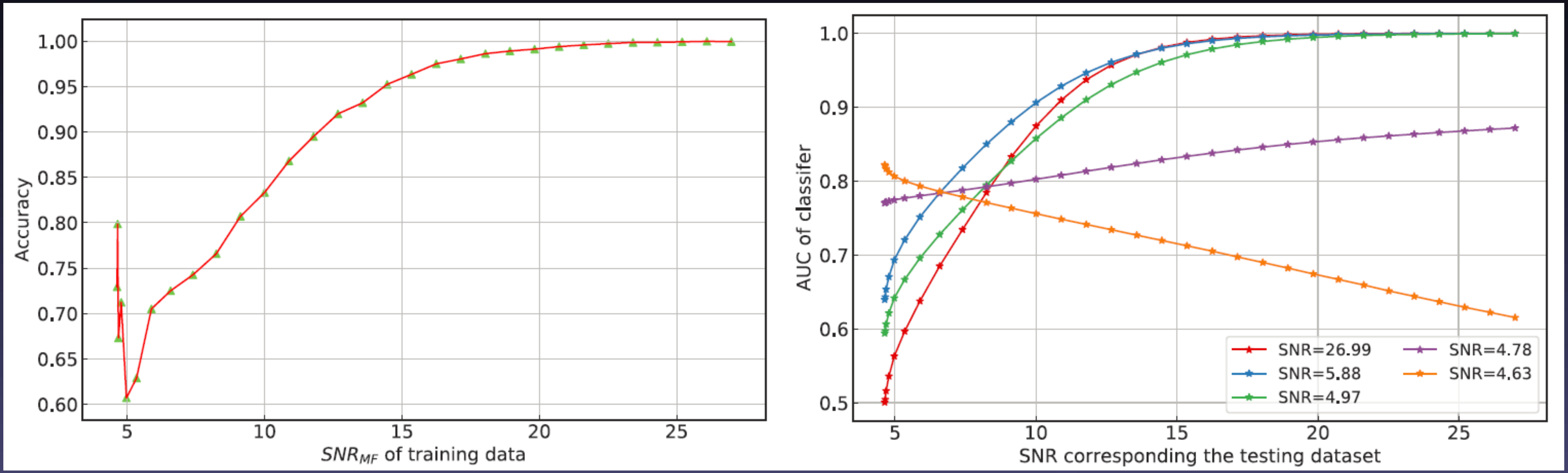
m1 and q uniformly distributed

SEOBNR waveform model

Data sample construction

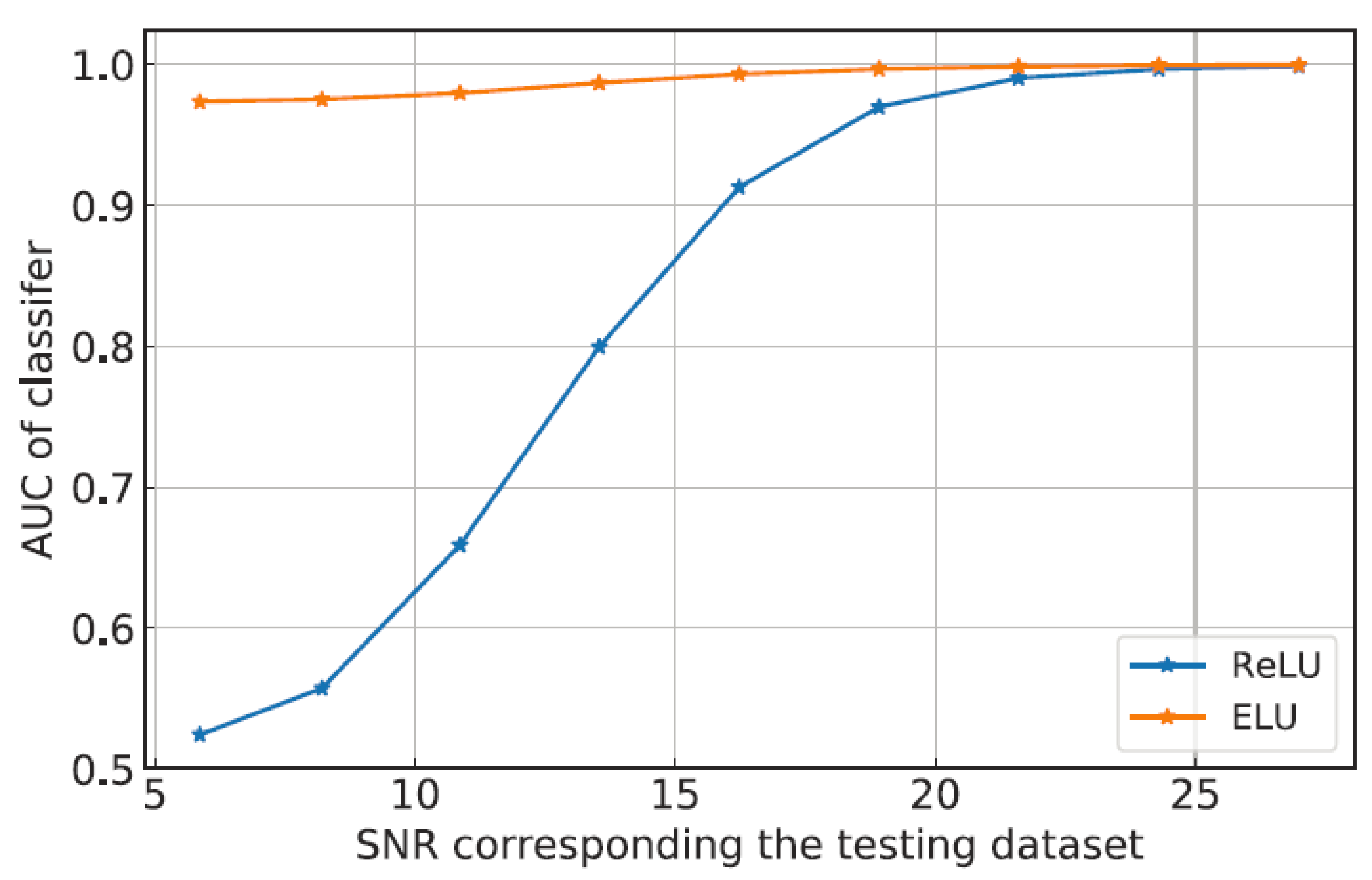


Effect of training data's strength



Optimal SNR of training data is about 5

Effect of activation function

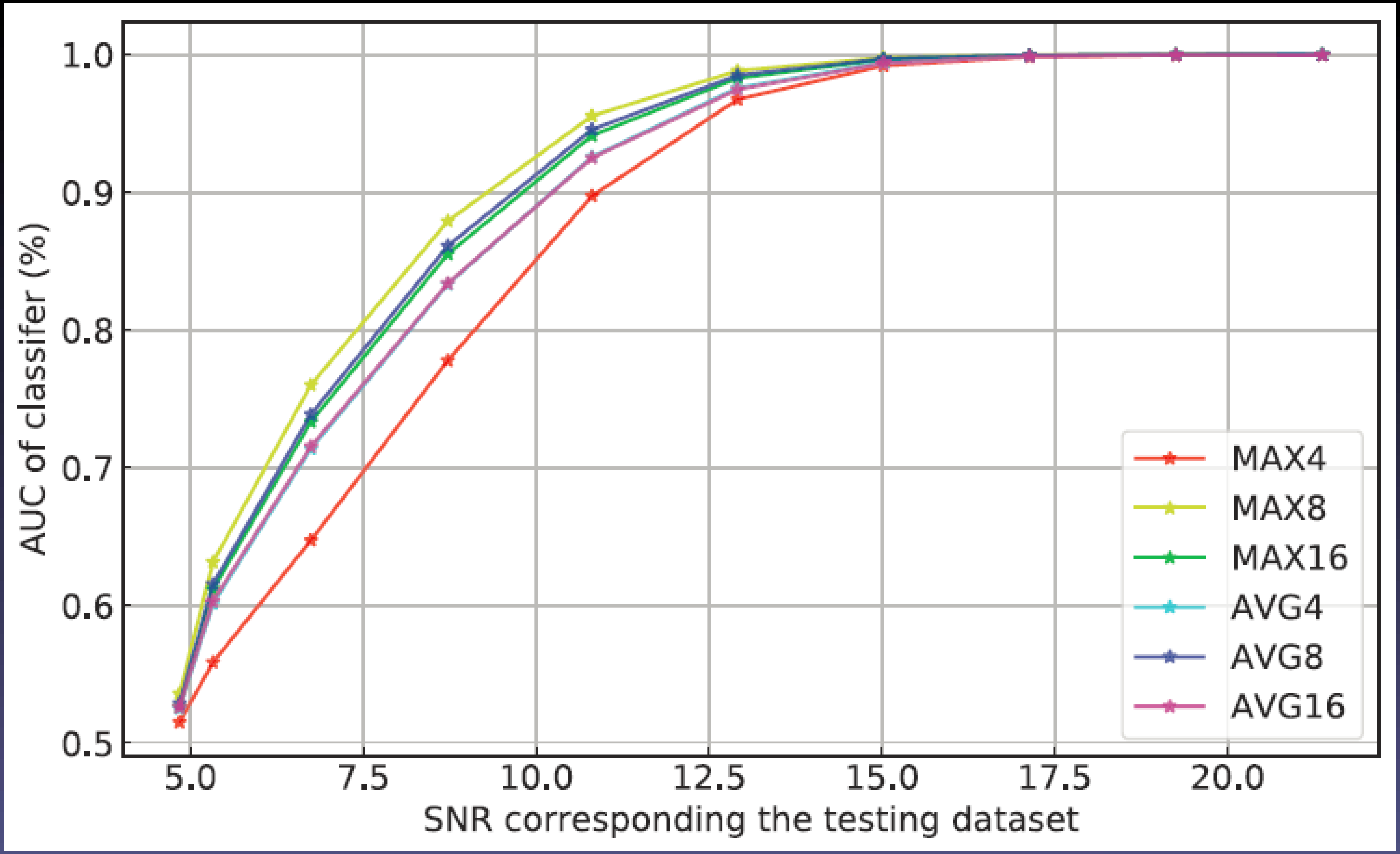


For GW data analysis, ELU is much better than ReLU

Effect of dilation parameter in CNN

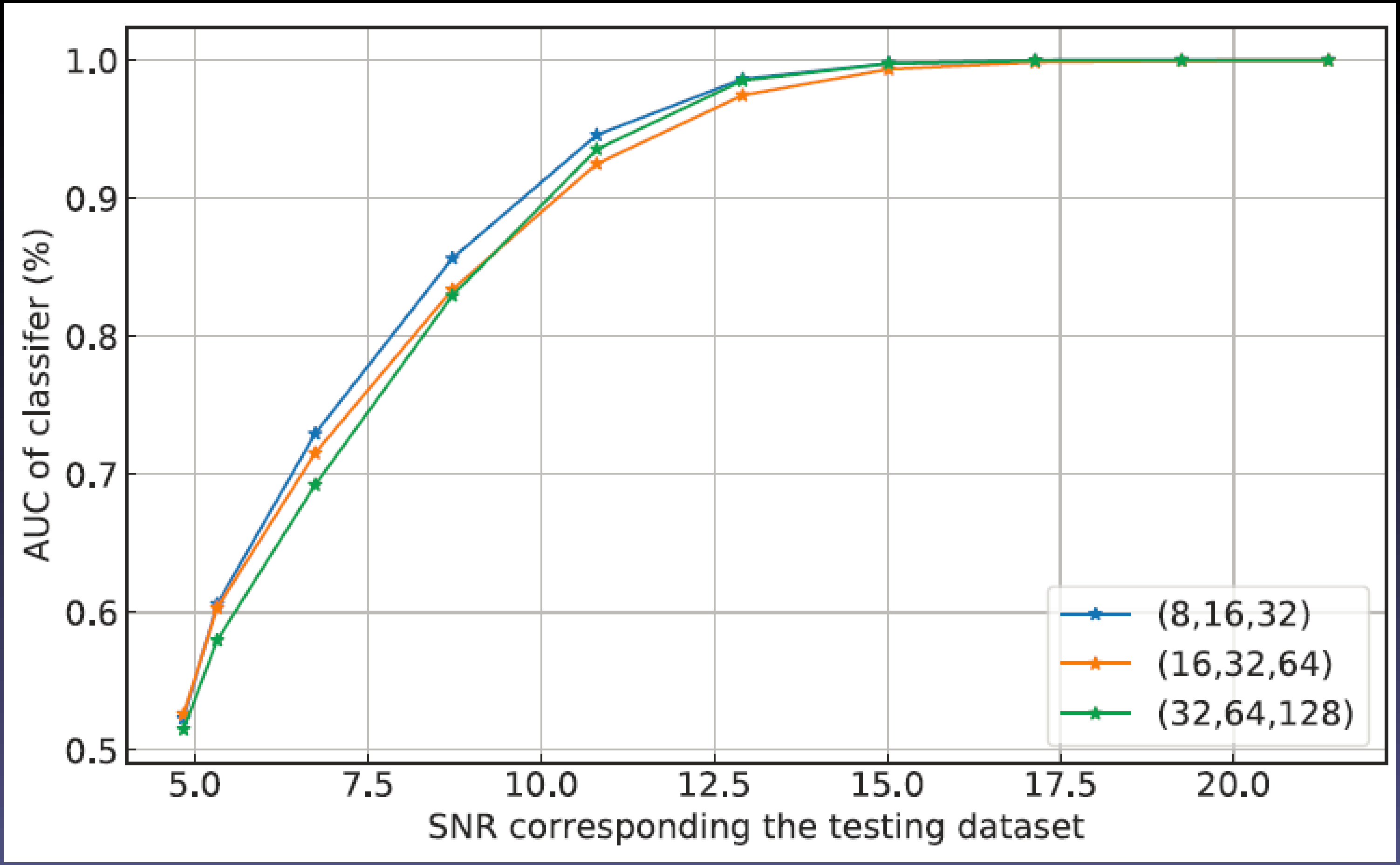
s=1 corresponds to normal convolution, if s < 4 different s result in roughly the same result

Effect of pooling



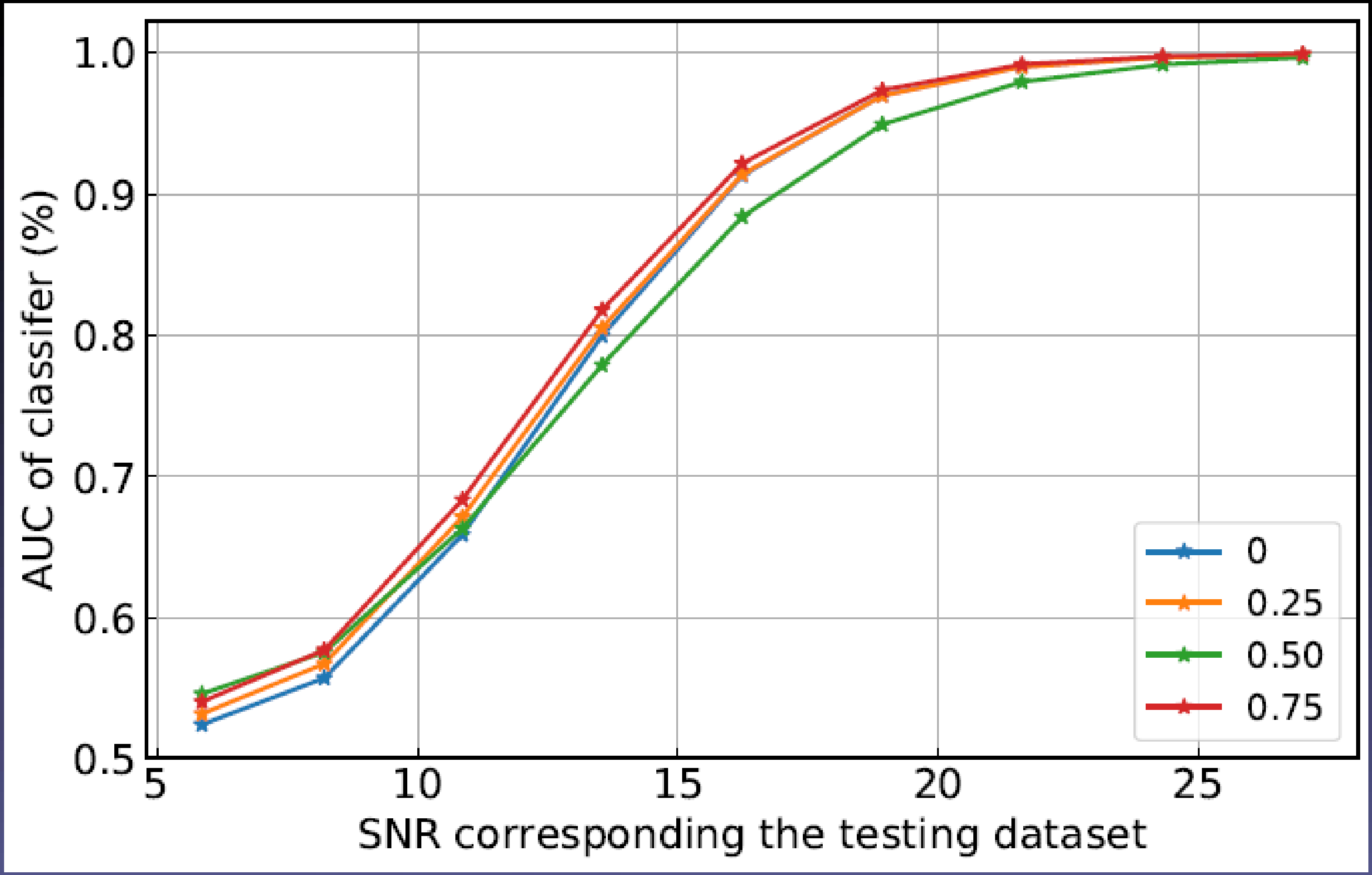
Optimal pooling size: 8; maximal pooling is a little better than average pooling

Effect of convolutional kernel size



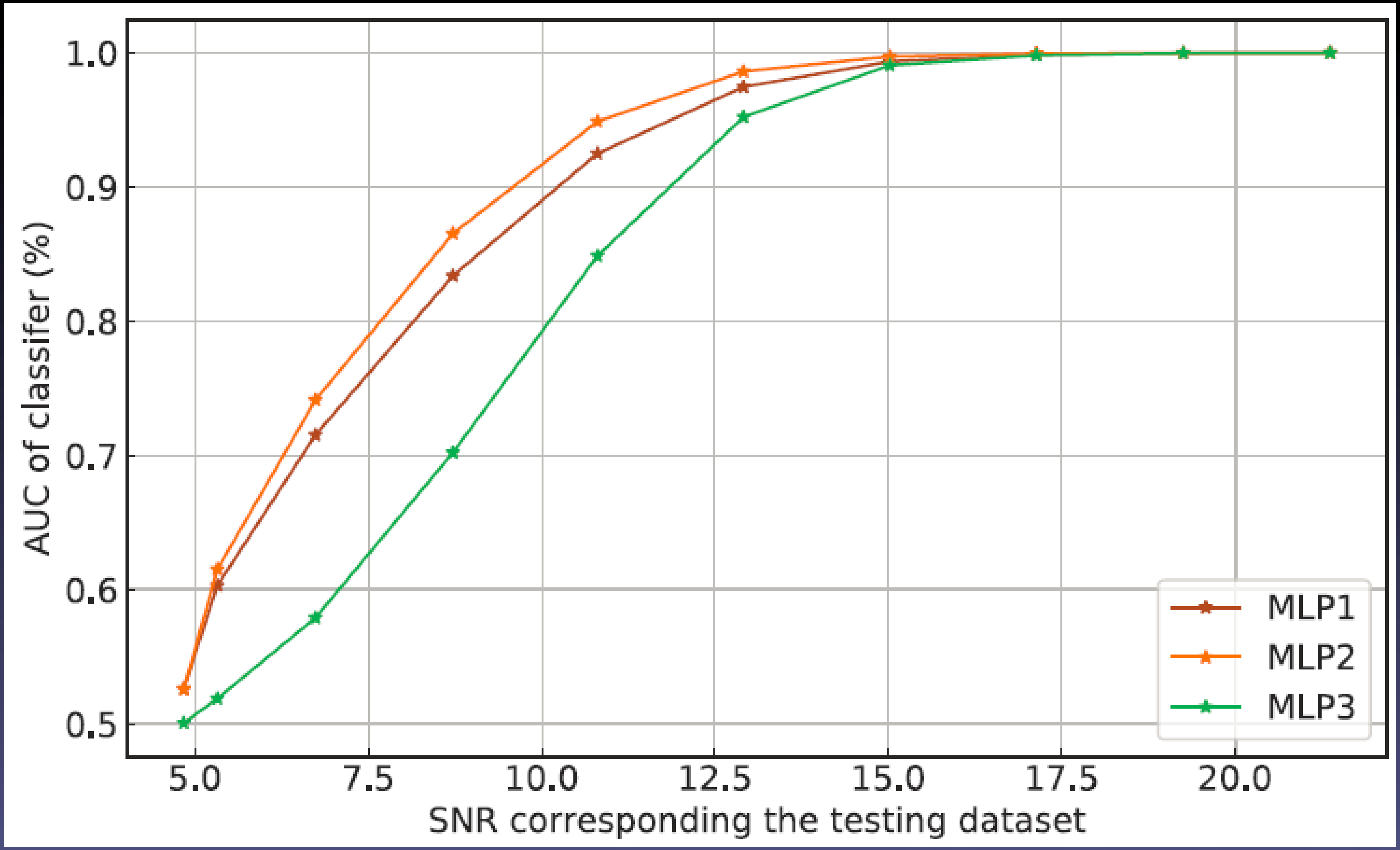
The optimal one (8,16,32), 8 correspondes to the best pooling size which should be the GW signal characteristic size

Effect of drop out rate in FCL



UIUC used 0; Glasgow used 0.5 (worst). Optimal setting: 0.25~0.75 (a little better)

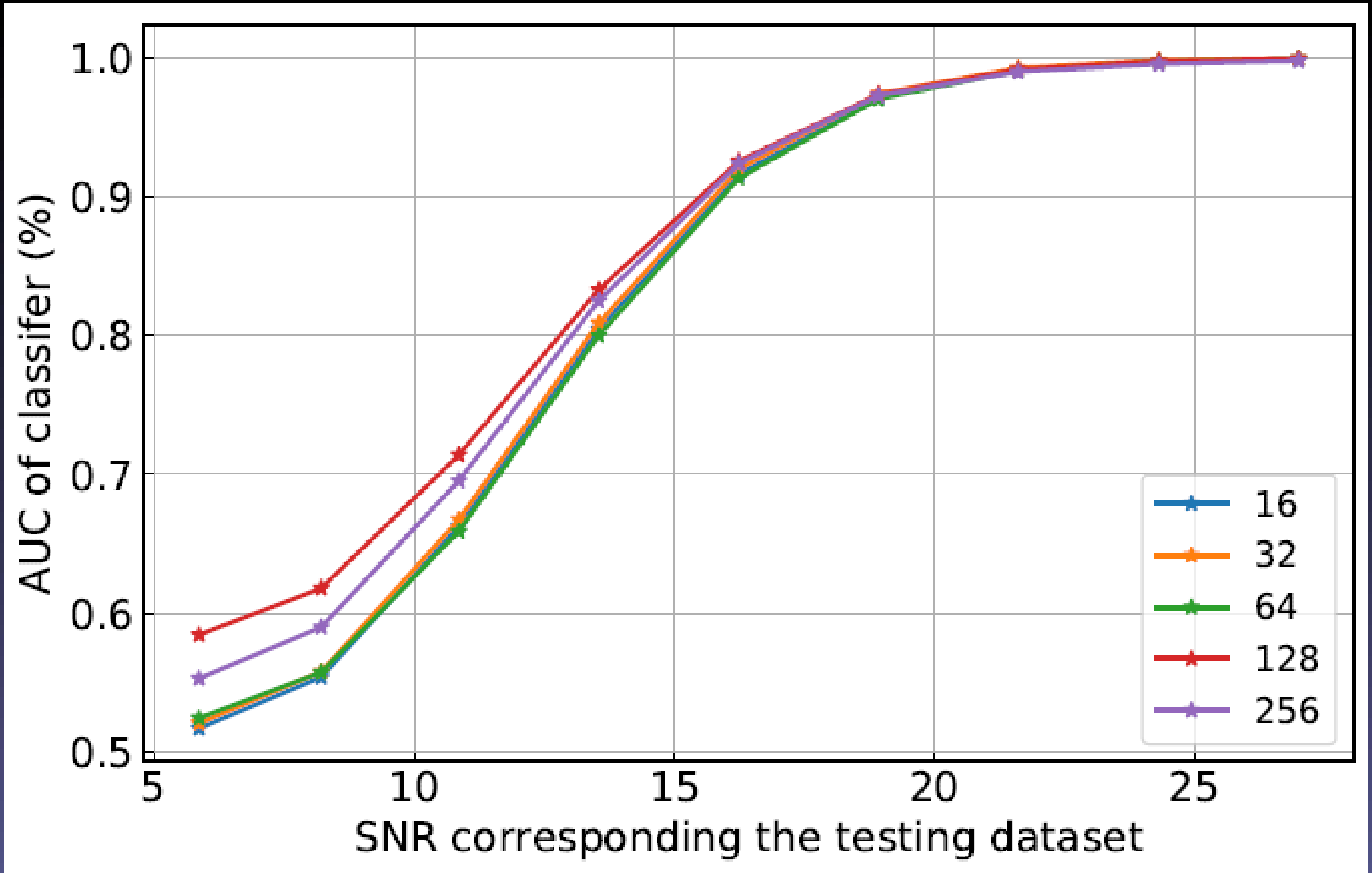
Effect of FCL layer number



Optimal Fully Connected Layer number: 2

Both UIUC and Glasgow used 3

Effect of FCL size

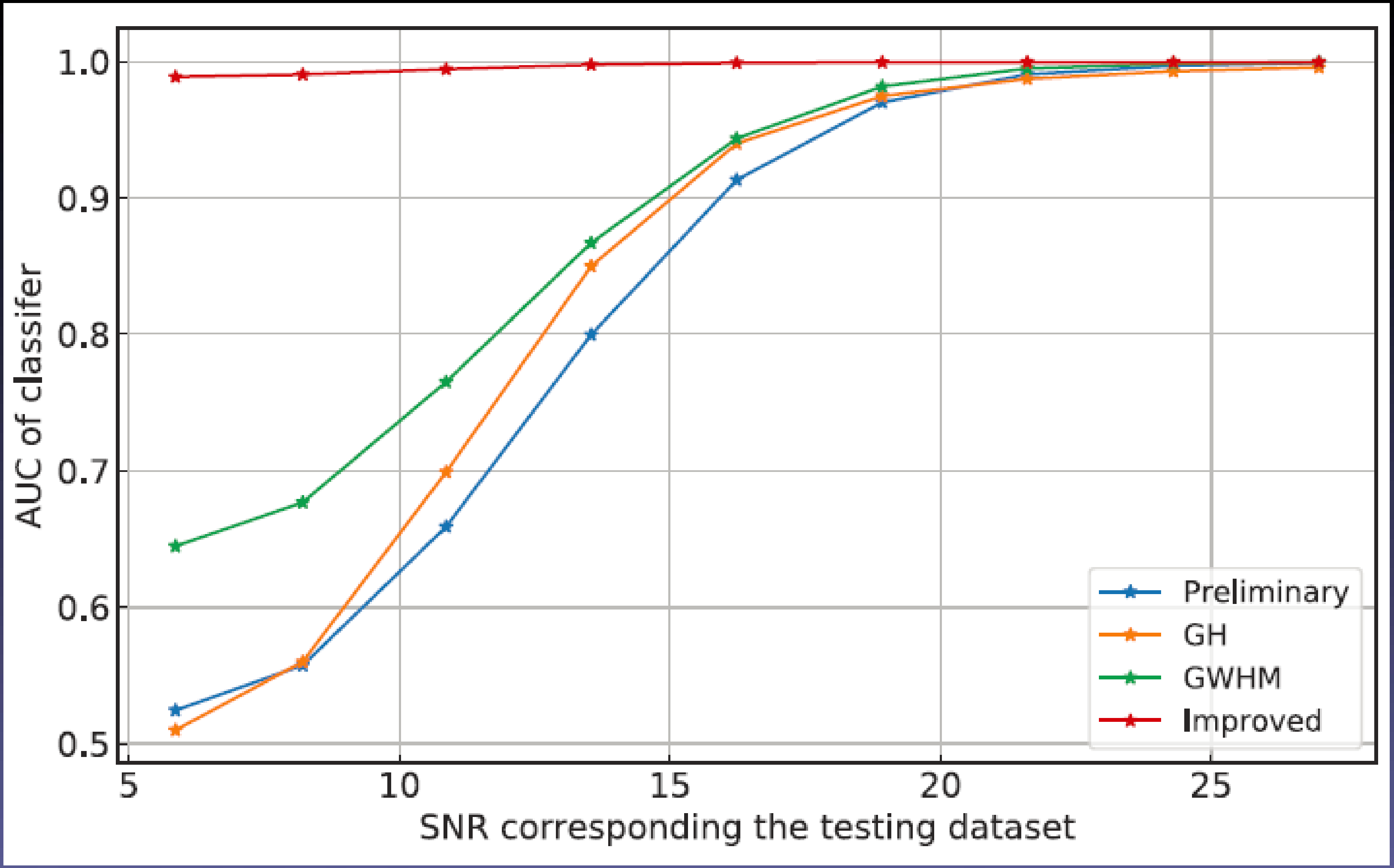


Optimal fully connected layer size: 128

Comparison of networks

	Parameter	Layer								
		1	2	3	4	5	6	7	8	9
Preliminary	Type	CON	CON	CON	MLP	MLP				
	No. of Neurons	16	32	64	64	2				
	Filter size/Dilate	16/1	8/1	8/1	-	-				
	Avg pool size/Stride	16/2	16/2	16/2	-	-				
	Drop out	0	0	0	0	0				
	Activation function	ReLU	ReLU	ReLU	ReLU	SMax				
GH	Type	CON	CON	CON	CON	MLP	MLP	MLP		
	No. of Neurons	64	128	256	512	128	64	2		
	Filter size/Dilate	16/1	16/2	16/2	32/2	-	-	-		
	Max pool size/Stride	4/4	4/4	4/4	4/4	-	-	-		
	Drop out	0	0	0	0	0	0	0		
	Activation function	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	SMax		
GWHM	Type	CON	CON	CON	CON	CON	CON	MLP	MLP	MLP
	No. of Neurons	8	8	16	16	32	32	64	64	2
	Filter size/Dilate	64/1	32/1	32/1	16/1	16/1	16/1	-	-	-
	Max pool size/Stride	-	8/2	-	6/2	-	4/2	-	-	-
	Drop out	0	0	0	0	0	0	0.5	0.5	0
	Activation function	ELU	ELU	ELU	ELU	ELU	ELU	ELU	ELU	SMax
Improved	Type	CON	CON	CON	MLP	MLP				
	No. of Neurons	64	128	256	128	2				
	Filter size/Dilate	16/3	8/3	8/3	-	-				
	Max pool size/Stride	16/2	16/2	16/2	-	-				
	Drop out	0	0	0	0.75	0				
	Activation function	ELu	ELu	ELu	ELu	SMax				

Effect comparison among networks



Summary on optimal network for GW

1. Optimal SNR for training data exist
2. Dilation is not needed in convolution
3. Active function: ELU is better than ReLU
4. Optimal size exits for pooling; maximal pooling or average pooling work equally
5. Optimal numbers for both neurons and network layers exist
6. Optimal dropout probability exist for training

Next step → generalization of DL for GW