Problem A High-speed Signal Modeling

Introduction

This contest consists of two problems. This problem is about signal modeling.

High-speed interface, as the key part of communication between chips, plays a crucial role in computing, network, wireless, terminal and other fields. Due to the non-ideal characteristics of analog devices and transmission channels in communication links, the received signals of high-speed interfaces are often accompanied by specific impairments. With the continuous evolution of communication rates, in addition to linear impairments, end-to-end communication signals also contain strong nonlinearity and noise. The increasingly complex damage components make the system analysis and algorithm design more complex.

A typical digital signal processing (DSP) assisted PAM4 communication system is shown in Fig. 1. A digital pattern sequence is coded in PAM4 pattern with the symbols $[-3, -1, 1, 3]$ to represent the 4 levels of a PAM4 signal. After TX DSP, the digital signal is converted to analog waveform using a digital to analog converter (DAC). After transmission by a high speed channel, the analog signal is then received by a receiver. The analog to digital converter (ADC) converts the analog signal to a digital signal. After equalizing and slicing by the RX DSP, the original digital signal from the transmitter is recovered at the receiver.

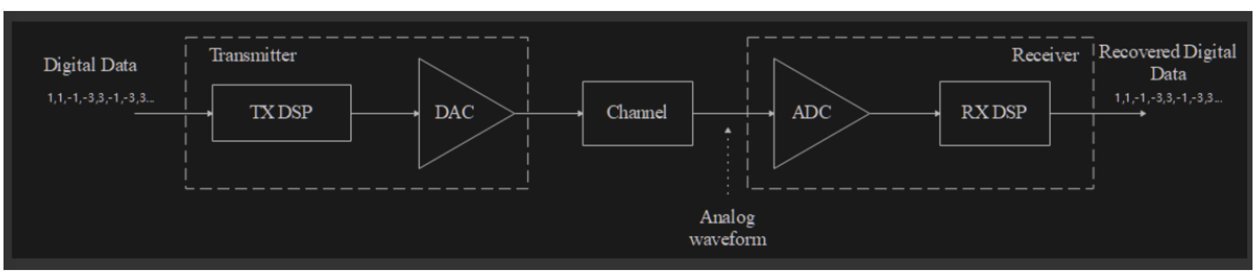

Figure 1: PAM4 DSP-assisted communication system

Challenges

The main challenges of this problem are as follow:

1. The damage in the channel is diverse and coupled, which stems from analog/optical device, passive link and active amplification, including noise, insertion loss, nonlinearity, etc. The nonlinearity and noise are amplified in the band-limited channel when the transmission rate is increased.

2. The energy efficiency per bit is reduced with the rapid growth of transmission speed. The architecture faces a severe challenge of power.

In view of the above background, you should perform damage decomposition on the signal, and accurately model and analyze the signal. Under the given input data, the signal waveform is modeled and fitted, and the original data is restored from the analog waveform through the algorithm. In order to match the requirements of low power consumption, delay, and bit error rate in the actual communication system, this contest must be completed within the specified complexity and accuracy.

Task

As shown in Fig. 2, analog data, $ana\_waveform\_1$, is a captured analog waveform from a real communication system. The corresponding digital pattern is given in dig_data_1. First, you should derive the $signal\_modeling\_f$ used to transfer dig_data_1 to $ana\_waveform\_1$. Then, using the $signal_modeling_f$ and dig_data_2, recover the $ana\_waveform\_2$, shown in Fig.2. You will receive a score based on how accurately your $ana\_waveform\_1$ matches the reference waveform. It is guaranteed that the over sample rate between the digital and analog data is always 16 in this task.

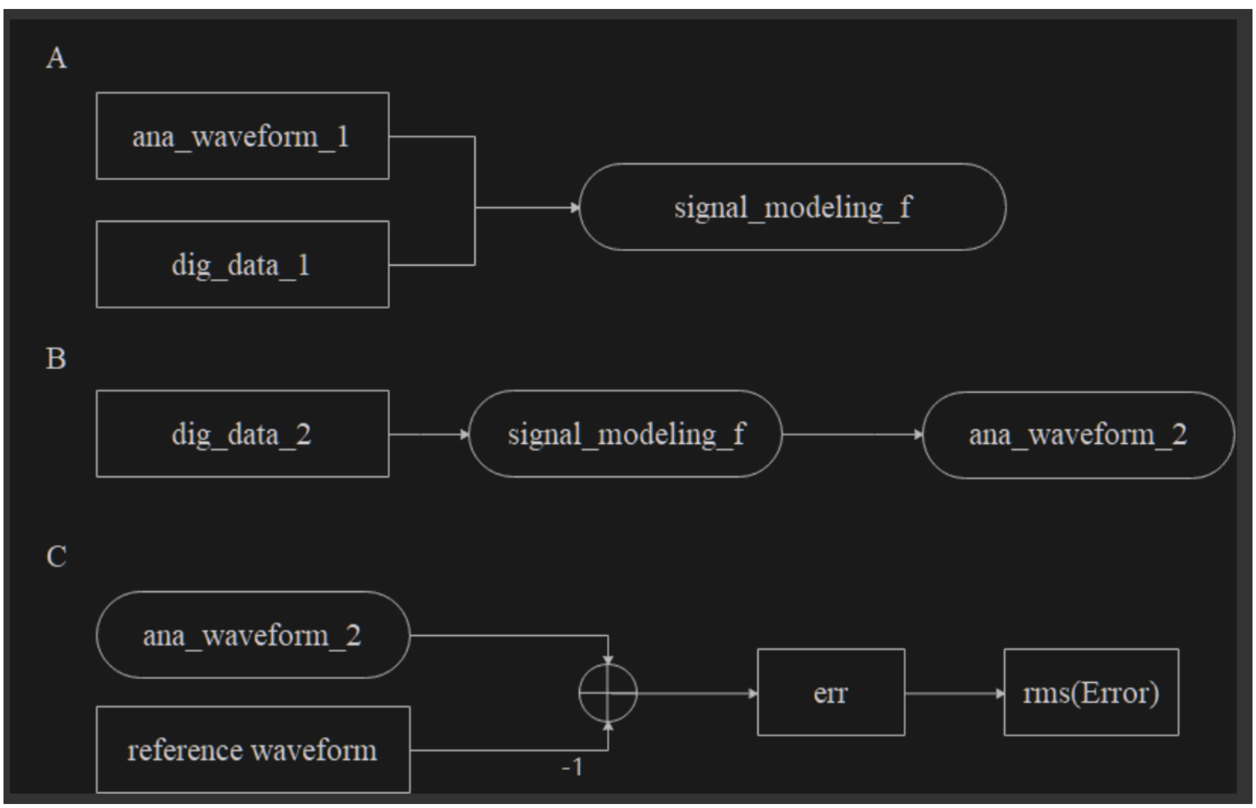

Figure 2: PAM4 DSP-assisted communication system

Input

Your program should read input using standard input (stdin).

- The first line of input contains the integer N, the number of samples in the analog wave.
- The next line contains N space-separated real numbers $a_1, a_2, \ldots, a_N$, the contents of the analog wave. It is guaranteed that N is divisible by 4, and that $-1 \leq a_i \leq 1$ for each $1 \leq i \leq N$.

Note that there is a large amount of input data, so you should use an efficient input method to read the data. Documentation on how to do this in your language of choice can be found at:

`https://open.kattis.com/languages`

Case	N	M	K
Case 1	87360	5460	11000
Case 2	176000	11000	16000
Case 3	176000	11000	16000
Case 4	84800	5300	10982

Figure 3: Challenges and goals

Output

Your program should write output to standard output (stdout).

First, you should write an integer L, the number of samples in your analog signal. This number must be equal to $K * 16$, or you will receive the verdict Wrong Answer.

Then, on the next line, print L space-separated real numbers, your reconstruction of $ana_{w}aveform_{2}$.

Scoring

Your solution will be tested and scored on 4 testcases, which are described above. To prevent overfitting, these testcases are kept hidden.

To receive a score, you must output a reconstruction of $ana\_waveform\_2$ within the time limit. To compute the score of your program, the root mean square error between your $ana\_waveform\_2$ is and the reference waveform is used.

For each testcase, you will receive between 0 and 25 points, with more points awarded for lower error. The total score for the problem is the sum of the score for each testcase. If you make multiple submissions, the one with highest total score is the one counted on the scoreboard.

Sample

A sample input and output is shown below. You will not receive a score on it, and it is only meant to illustrate the input and output format. In particular, no inferences about $signal\_modeling\_f$ should be made from it, as the analog numbers are chosen uniformly at random in the interval $[-1, 1]$. The sample will be run on every submission, and is thus a good way to test that your program works correctly.

Sample Input 1

```
16

0.22061 -0.70137 -0.55763 0.35266 0.96795 0.84704

-0.48585 0.55557 -0.46686 -0.98207 0.13065 0.85421

0.19748 -0.71182 0.38312 0.69907
1
3
2
-1 3
```

Sample Output 1

```
32
-0.52561 0.82934 -0.92622 -0.00621 -0.83611

0.61852 0.23000 -0.51393 -0.83801 -0.50055 0.63153

0.84901 0.50833 -0.71813 0.51604 0.02155 0.93926
```

```
-0.49804 0.43322 -0.28385 0.56230 -0.21097 0.20067

0.12074 -0.62109 -0.90709 -0.73369 -0.01097

0.16630 0.01916 -0.64849 0.63309
```

1 Hints for the competition organizer

1.1 Questions and answers

Question : If my answer is exactly the negative of the standard answer, can I still get points? It seems that if the input is only the *ana_waveform*, we cannot determine the sign of the *digital_data*, right?

Answer : *Dig_data* and *ana_waveform* correspond to the same form and can distinguish the sign of *digital_data*.

Question : My i know what modulation is used to convert pam4 into analog before distortion ?

Answer : The *analog_waveform* is converted from PAM4 digital data by a DAC digital to analog converter. `https://en.wikipedia.org/wiki/Digital-to-analog_converter`

Question : Do we have any info on the treatment of the wave in the Tx phase? Was some filter applied before sending it?

Answer : Yes, a 3tap FFE (pre1,main,post1) digital filter was used before the digital data send to the DAC.

Yes, a 3tap FFE (pre1,main,post1) digital filter was used before the digital data send to the DAC.

Question : can we get at least one real input sample for each problem?

Answer : For problem 1, the *ana_waveform_1* is the answer which means the recovered analog waveform from *dig_data_1* should be similar to *ana_waveform_1* . For problem 2, the problem1 can be used as a reference. The *dig_data_1* is the answer extract from *ana_waveform_1*. A little tips: The ratio between sample input length and sample output length should be attention.