# Homework #1: Computational Accuracy

Professor Deepa Kundur
University of Toronto

## Questions

Please print this out and answer the following questions in the space provided below. Please add additional sheets if necessary or use the backs of sheets.

1. You are using an $n$-bit signed integer fixed-point format to represent a number $x$. Suppose you double the number of bits used to represent $x$, so that it is now $2n$. By what factor (as a function of $n$) does the range of numbers increase? It will help to do this problem using the following steps:

    a. What is the range of numbers represented by an $n$-bit signed integer fixed-point format number? Note the range of $x$ is the number of all possible integers that it can represent, so if $-A \leq x \leq B$ where $A$, $B$ are positive integers then the range of $x$ is $A+B+1$. Another way to think of this is to answer, how many different integers can an $n$-bit signed integer fixed-point format represent?

    b. Compute the above range assuming $2n$-bits are used to represent the same format number.

    c. Determine the ratio if your answer in part b to your answer in part a.

2. You are using an $n$-bit signed fractional fixed-point format to represent a number $x$.

    a. Suppose you double the number of bits used to represent $x$, so that it is now $2n$. By what factor (as a function of $n$) does the precision improve?

    b. Suppose you increase the number of bits used to represent $x$ by $k$, so that it is now $n+k$. By what factor does the precision improve?

    c. Suppose you are using $n = 4$. You would like to improve precision by a factor of at least six by increasing the number of bits. What is the minimum number of bits required to increase precision in this way? Please show all steps and reasoning for full points.

3. You are using an $n$-bit signed integer fixed-point format to represent a number $x$. Suppose you change the number of bits to $n$+1 (i.e., you increase the number of bits used to represent $x$ by 1). By what factor (as a function of $n$) does the range of numbers increase? Please do the problem in the following steps:

    a. How many different integers can an $n$-bit signed integer fixed-point format represent?

    b. Compute the above assuming ($n$+1)-bits are used to represent the same format number.

    c. Determine the ratio of your answer in part b over your answer in part a.

    d. If you wanted to quadruple the range of numbers represented using a signed integer fixed-point format, by how many *additional* bits would you have to increase your number representation?

4. Assume you are using unsigned integer fixed-point format to represent samples of a signal. Show that the dynamic range of this signal increase by approximately 6 dB for each additional bit used to represent its value. It will help to do this problem using the following steps:

    a. What is the dynamic range (ratio of maximum value to minimum non-zero value) of a number represented using $n$-bit unsigned integer fixed-point format? Show your work for full points. Please present your answer as a ratio value (i.e., not in dB).

    b. Repeat part a above for ($n$+1)-bit unsigned integer fixed-point format. That is, what terms can you neglect?

    c. Determine the ratio of the answers in part b to part a.

    d. Using your answer to part c, show how increasing the number of bits to $n$+1 will result in approximately a 6 dB increase in dynamic range. Please explain thoroughly for full points.

5. You are using an $n$-bit floating-point format representation with a $k$-bit mantissa (with fractional representation).

    a. As conducted during the lecture, compute the precision of this format as a percentage resolution. Your answer should be a function of $k$, or $n$, or both.

    b. By what factor does the precision improve (i.e., decrease) for each additional bit used in the mantissa?