Instructions:

The python package that I used in order to generate and graph the data is called openpyxl. This package allows both the manipulation of the data from Excel sheets into python, and then the storage of that data into column in Excel. This is often needed, because if you use an equation in Excel, the “value” of that cell is the equation, not the numerical value. Thus, it is important to understand each of the columns in order to graph or visualize certain subsets of the data.

Each of the columns in the final data sets have the following meanings:

B: Branch index indicates phonon mode (1, 2 transverse acoustic, 3 longitudinal acoustic. 4 longitudinal optical, 5, 6 transverse optical based on similarities but this might be wrong).

C: the x-coordinate in reciprocal space (in Angstroms ^-1)

D: the y coordinate in reciprocal space (in Angstroms ^-1)

E: the z-coordinate in reciprocal space (in Angstroms ^-1)

F: frequency of phonon mode (in THz)

G: F\_0, a normalized measure of how far a phonon mode in q\_j is from equilibrium when **there is no** applied thermal gradient (according to Lukas Lindsay in the word document “Normalized change in phonon number”). Units of J m/K

H: F\_0, a normalized measure of how far a phonon mode in q\_j is from equilibrium when **there is** an applied thermal gradient along the (100) direction (according to Lukas Lindsay in the word document “Normalized change in phonon number”). Units of J m/K

I: F/F\_0, a ratio of F and F\_0

J: Filler column, artifact from when Dr. Lindsay first

K: The phonon population, calculated using Bose-Einstein statistics and the frequency of the phonon mode in column F

L: The change in phonon population due to the applied thermal gradient along (100), as according to the equation:



taken from “Normalized change in phonon number”, although also discussed around equation 10 and 11 in the paper “Ab initio theory of lattice thermal conductivity in diamond” by Ward and Broido. These arise from the linearized Boltzmann transport equations.

M: Total phonon population with an applied gradient: Column K + Column L = Column M

N: The Debye waller factor without an applied thermal gradient:



where a is the generalized coordinate, outlined in section 2.2 of Xu Ruqing’s: X-RAY THERMAL DIFFUSE SCATTERING AND ITS STUDIES OF LATTICE DYNAMICS



O: Same as above, except instead of using eta\_0 (Column K) in the generalized coordinate equation, I used eta (Column M) to determine the total influence on the Debye-Waller factor.

Columns P and Q are vestiges from the initial round of codes, calculating the total Bragg Scattering at each point, which is never relevant to our calculations.

R: The thermal diffuse scattering without the applied thermal gradient, according to the equation:



where both eta\_0 (Column K) and the Debye Waller Factor (Column O) is used.

S: The logarithmic contribution of the TDS

T: The thermal diffuse scattering with an applied thermal gradient along the (100) direction, as per the equation:



except with the eta (Column M) and Debye Waller Factor (Column O) which take into account the thermal gradient.

U and V: the absolute and percent change of the diffuse scattering due to the applied thermal gradient respectively.

W: The amount of counts (proportional to time and incident photons) needed to acquire a signal to noise ratio of 10. This is according to the equation:

which I derived with help from notes provided by Jacob Ruff, and is based on Poisson counting statistics. In this case, F = signal/noise = 10, I\_w is the TDS intensity with an applied thermal current and I\_wo is the TDS intensity without an applied thermal current.

X: The noise of the data, assuming that the signal to noise ratio is 10, as per the equation:

which was, once again, obtained from Poisson statistics with help from Jacob Ruff.

Y: Counts for F = 2

Z: Noise for F = 2

AA – AC: total summed scattering (not relevant for anything)

AD: Total signal for F = 10

AE: Total Signal for F = 2

The reason that so many columns seemingly have no relevance is that the codes that generate and manipulate this data is based upon its location in the spreadsheet. Thus, early misunderstandings and choices persisted throughout the project, without a mind for optimizing the spreadsheets for concision and clarity (and for this I apologize).