1. מימוש ה-dist\_kernel שלנו מקבל כקלט שי מערכים של נתונים A ו-B, ומספר p המייצג את בסיס הנורמה שנחשב, ה-kernel רץ עם 1000 בלוקים של חוטים כאשר כל אחד עם 1000 חוטים, ידוע ש-A ו-B הם בגודל 1000X1000.  
   המימוש שלנו מריץ חוט עם מספר בלוק j ומספר חוט בתוך הבלוק i, מחשב את ההפרש בין A[j][i] ל-B[j][i] בערך מוחלט ומחשב את חזקת p של המספר הזה.  
   כל חוט מוסיף אטומית את הערך שחישב לזיכרון משותף, שמועתק חזרה מה-kernel, הפונקציה שקראה ל-dist\_kernel תחשב את השורש מסדר p של הערך בזיכרון המשותף, הערך הזה הוא סכום ההפרשים בערך מוחלט ובחזקת p של כל התאים בשתי המטריצות, כדי להגיע לערך הרצוי נותר להוציא שורש p מהערך וזה נעשה על ידי הפונקציה הזו.



Gpu speedup = (25.936/0.1423 + 25.683/0.1421)/2 = 182.2

Numba speedup = (25.936/0.094 + 25.683/0.112)/2 = 252.61

הפונקציות המקבילית והפונקציה שרצה על ה-GPU אכן מהירות יותר משמעותית מהפונקציה שרצה על ה-CPU ללא מקבול, וזאת משום שמקבול שקורה בשתי הראשונות מאפשר לבצע את רוב החישובים מתוך חישובים של הפרש והעלאה בחזקת p, הסכימה של האיברים לא בהכרח תהיה יותר מהירה על ה-CPU או על ה-GPU.  
מכיוון שזמן החישוב הוא מאוד קצר, אין יתרון גדול בהרצה על-GPU על פני הרצה מקבילית באמצעות Numba משום שלוקח זמן להעתיק את הזיכרון ל-GPU ולהפעיל אותו מהתוכנית שלנו.