**Численное моделирование фотонного тензорного ядра для аппаратного ускорения матрично-векторных вычислений.**

***Колосов Г.А.1, Шорохов А.С2.***

1Студент, 2кандидат физико-математических наук

1,2МГУ им. М.В. Ломоносова, физический факультет, Москва, Россия

E-mail: kolosov.ga20@physics.msu.ru

Темп развития традиционной кремниевой электроники снизился из-за сложностей в фабрикации структур нанометрового размера, а также проявления квантовых эффектов (например, туннелирования электронов через очень малый затвор), принципиально ограничивающих размеры транзисторов. Вместе с этим возник ряд новых задач, для которых архитектура фон Неймана не является оптимальной: создание искусственного интеллекта, моделирование макромолекул и т.д.

Для создания более подходящей вычислительной архитектуры для искусственных нейронных сетей была предложена идея аналогового фотонного вычислителя [1], который может обладать существенно более высокой производительностью и энергоэффективностью относительно электронных аналогов.

В данной работе в программной среде ANSYS Lumerical Interconnect была построена модель аналогового фотонного вычислительного устройства на основе архитектуры кроссбар [2]. В качестве модуляторов входного вектора были использованы модуляторы Маха-Цендера [3] или микрокольцевые модуляторы [4]. При построении модели были учтены потери в каждом элементе оптической цепи (модуляторах [3-4], пересечениях волноводов [5], и т.д.), а также дисперсия направленных ответвителей, так как в системе используются множество длин волн для устранения эффекта деструктивной интерференции. Параметры детекторов на выходе из массива были оптимизированы с точки зрения минимизации количества битовых ошибок вычислений.

Была показана возможность с помощью данной модели производить матрично-векторные вычисления с необходимой для нужд искусственного интеллекта точностью. Так же была исследована эффективность проведения операции свертки в оптическом аналоговом домене, что показано на рис. 1.

Для построенной модели численно были получены оценки максимальной мощности на выходе цепи, мощности минимального полезного сигнала и мощности шумов в системе, что позволило получить оценки предельного размера кроссбар-массива фотонного тензорного ядра для различной требуемой точности вычислений. Исходя из данных о предельном размере массива была рассчитана энергоэффективность и производительность конечного устройства, что отображено на рис. 2.

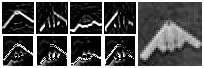
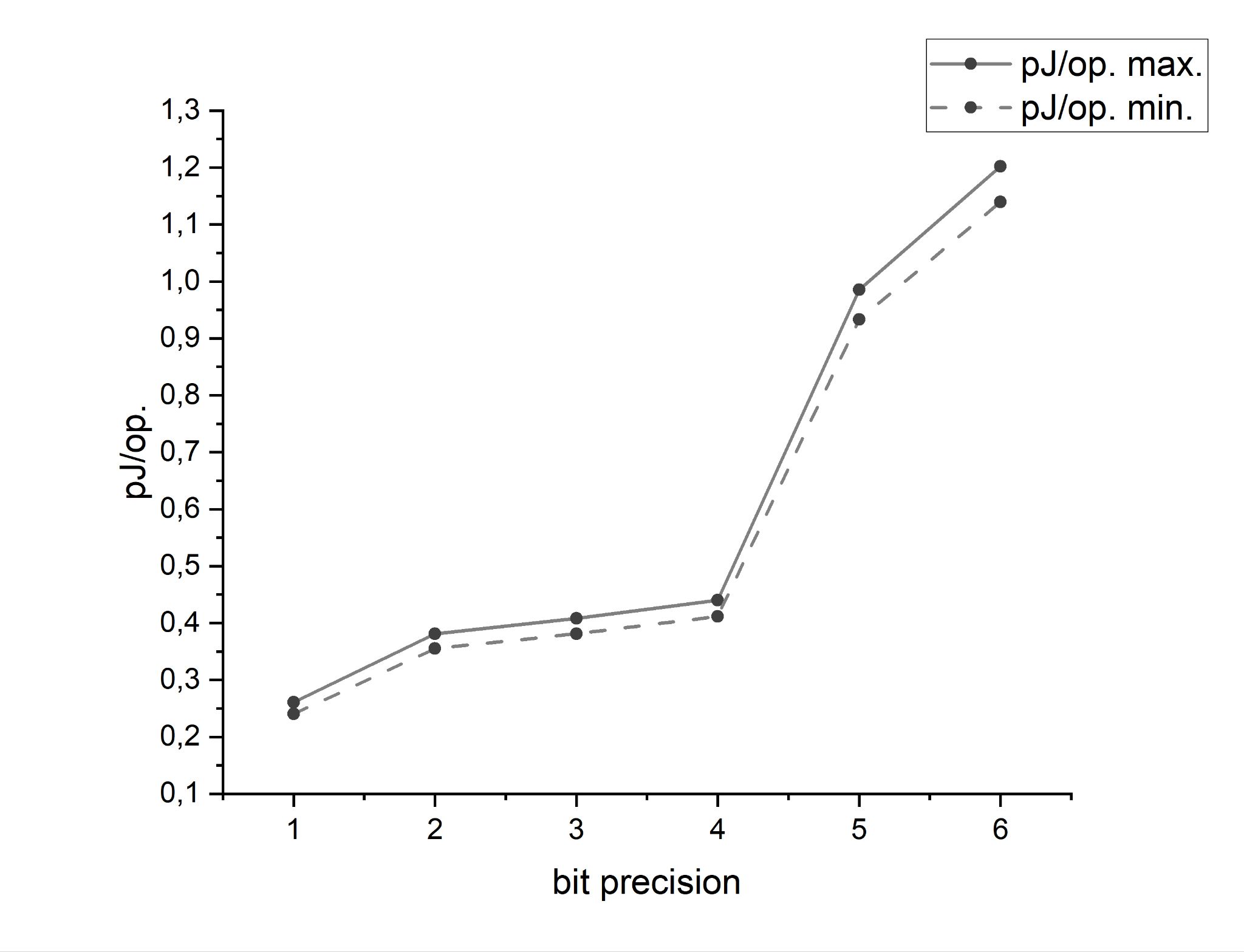
Рис. 1. Пример реализации операции свертки. Верхний ряд — эталонная свертка с помощью пакета numpy, нижний ряд — свертка с помощью модели фотонного тензорного ядра. Справа исходная картинка из датасета CIFAR-10. 

Рис. 2. Слева расчет энергоэффективности конечной системы в пикоджоулях на операцию. Справа расчет производительности в тераоперациях в секунду. На обоих графиках по оси х отложена требуемая битовая точность операций.

**Литература:**

1. Kitayama K. et al. Novel frontier of photonics for data processing—Photonic accelerator //Apl Photonics. – 2019. – Т. 4. – №. 9.

2. Feldmann J. et al. Parallel convolutional processing using an integrated photonic tensor core //Nature. – 2021. – Т. 589. – №. 7840. – С. 52-58.

3. Liao L. et al. High speed silicon Mach-Zehnder modulator //Optics express. – 2005. – Т. 13. – №. 8. – С. 3129-3135.

4. Dong P. et al. Wavelength-tunable silicon microring modulator //Optics express. – 2010. – Т. 18. – №. 11. – С. 10941-10946.

5. Wu S. et al. State-of-the-art and perspectives on silicon waveguide crossings: A review //Micromachines. – 2020. – Т. 11. – №. 3. – С. 326.