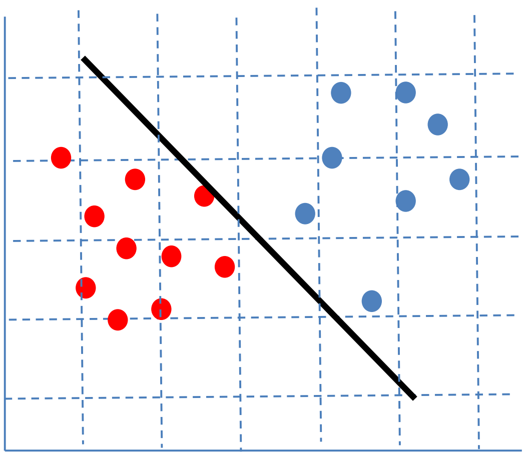
Minimum Precision Requirements for the SVM-SGD Learning Algorithm

Charbel Sakr, Ameya Patil, Sai Zhang, Yongjune Kim, Naresh Shanbhag

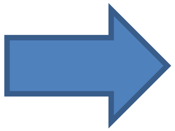
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign

Hyperplane Classification Example

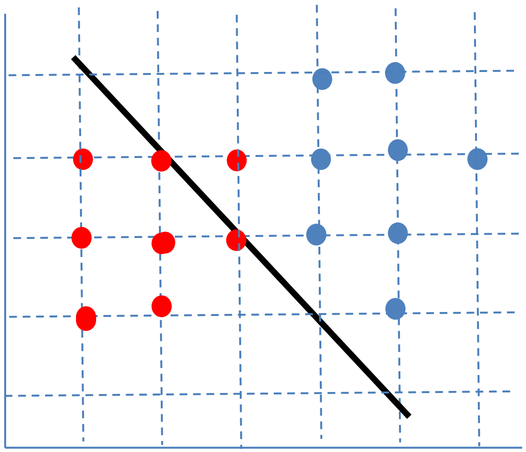
Can we control this behavior?



Geometry of classifier and data



Fixed-point Quantization



Geometry in fixed-point

Prior Work

Fixed-point Training

- Stochastic Rounding for regularization
  - Deep Learning with Limited Numerical Precision (Gupta et al., ICML, 2015)
  - BinaryConnect (Courbariaux et al., NIPS, 2015)
  - BinaryNet (Courbariaux et al., arXiv, 2016)
- Bitwise Operations (Kim & Smaragdis, arXiv, 2016)
- XNOR-net (Rastegari et al., ECCV, 2016)

Training in a discrete space is harder

Fixed-point Quantization After Training

- Exhaustive Search (Hwang & Sung, SiPS, 2014)

Trial-and-error may be hard

- SQNR based precision allocation (Lin et al., ICML, 2016)

No theoretical guarantees on accuracy

The SVM-SGD learning algorithm

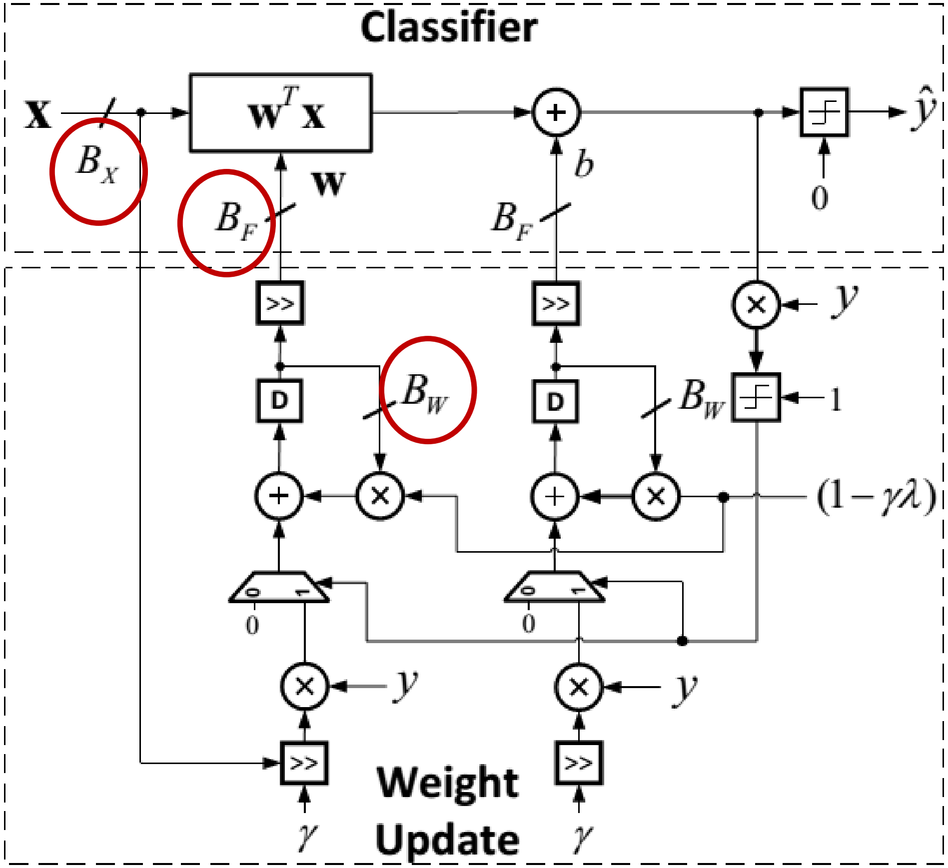
Classification

$$\mathbf{w}^T \mathbf{x}_n + b \underset{\hat{y}_n = -1}{\overset{\hat{y}_n = 1}{\gtrless}} 0$$

Training

$$\mathbf{w}_{n+1} = (1 - \gamma\ \lambda)\mathbf{w}_n + \begin{cases} 0 & \text{if } y_n(\mathbf{w}_n^T \mathbf{x}_n + b) > 1, \\ \gamma\ y_n\ \mathbf{x}_n & \text{otherwise.} \end{cases}$$

The SVM-SGD Architecture



Effects of Quantization

- Decision equation is modified to:

$$(\mathbf{w} + \mathbf{q}_w)^T(\mathbf{x} + \mathbf{q}_x) + b + q_b \gtreqless 0$$

- Decision equation is modified to:

$$\mathbf{q}_x \in \mathcal{R}^N,\ \mathbf{q}_w \in \mathcal{R}^N, \text{ and } q_b \in \mathcal{R}$$

$$\mathbf{q}_x \sim (\mathrm{U}[-\tfrac{\Delta_x}{2}, \tfrac{\Delta_x}{2}])^N \quad ; \Delta_x = 2^{-(B_X - 1)}$$

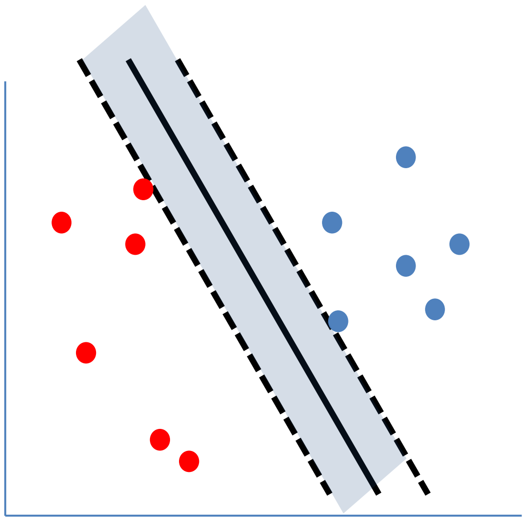
$$\mathbf{q}_w \sim (\mathrm{U}[-\tfrac{\Delta_f}{2}, \tfrac{\Delta_f}{2}])^N \quad ; \Delta_f = 2^{-(B_F - 1)}$$

$$q_b \sim \mathrm{U}[-\tfrac{\Delta_f}{2}, \tfrac{\Delta_f}{2}]$$

- Total quantization noise can be simplified to:

$$\mathbf{q}_w^T\mathbf{x} + \mathbf{w}^T\mathbf{q}_x + q_b$$

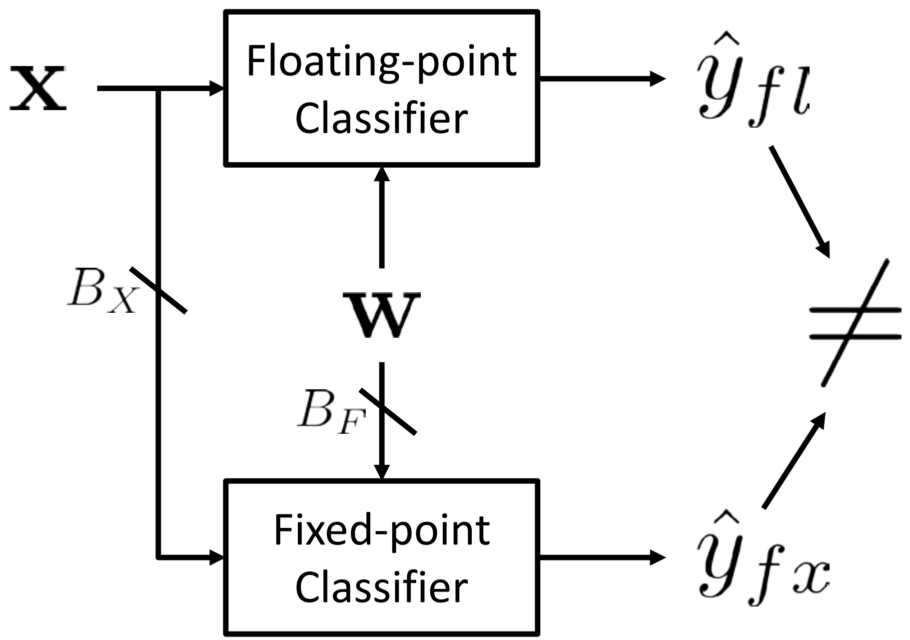
Geometric Bound



$$B_X > \log_2\left(\frac{\sqrt{N}||\mathbf{w}||}{1-(1+\sqrt{N}||\mathbf{x}||)2^{-B_F}}\right)$$

- Dependence on margin
- Trade-off between weight and data precision
- Dependence on dimensionality

Probabilistic Bound

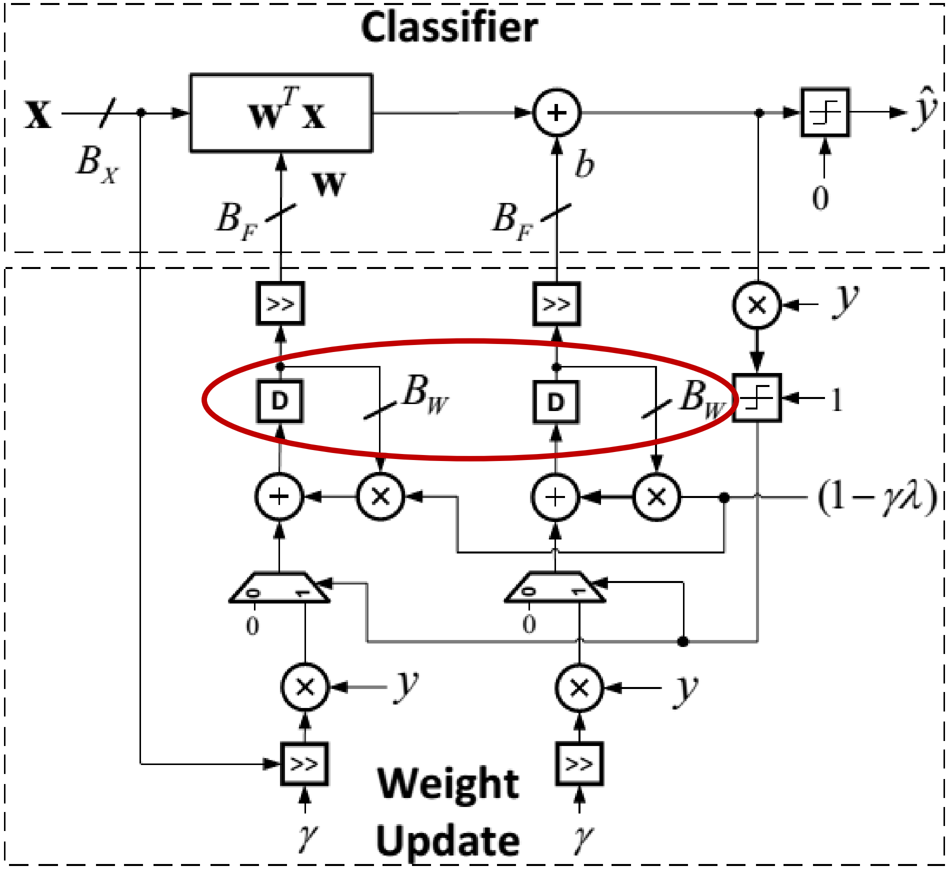


- Data dependence
- Exponential trade-off between precision and accuracy
- Compute once and reuse

$$p_m \leq \frac{1}{24}\left(\Delta_x^2 \mathbb{E}\left[\frac{||\mathbf{w}||^2}{|\mathbf{w}^T\mathbf{X}+b|^2}\right] + \Delta_f^2 \mathbb{E}\left[\frac{||\mathbf{X}||^2+1}{|\mathbf{w}^T\mathbf{X}+b|^2}\right]\right)$$

$$\Delta_x = 2^{-(B_X-1)} \quad \Delta_f = 2^{-(B_F-1)}$$

Precision in training



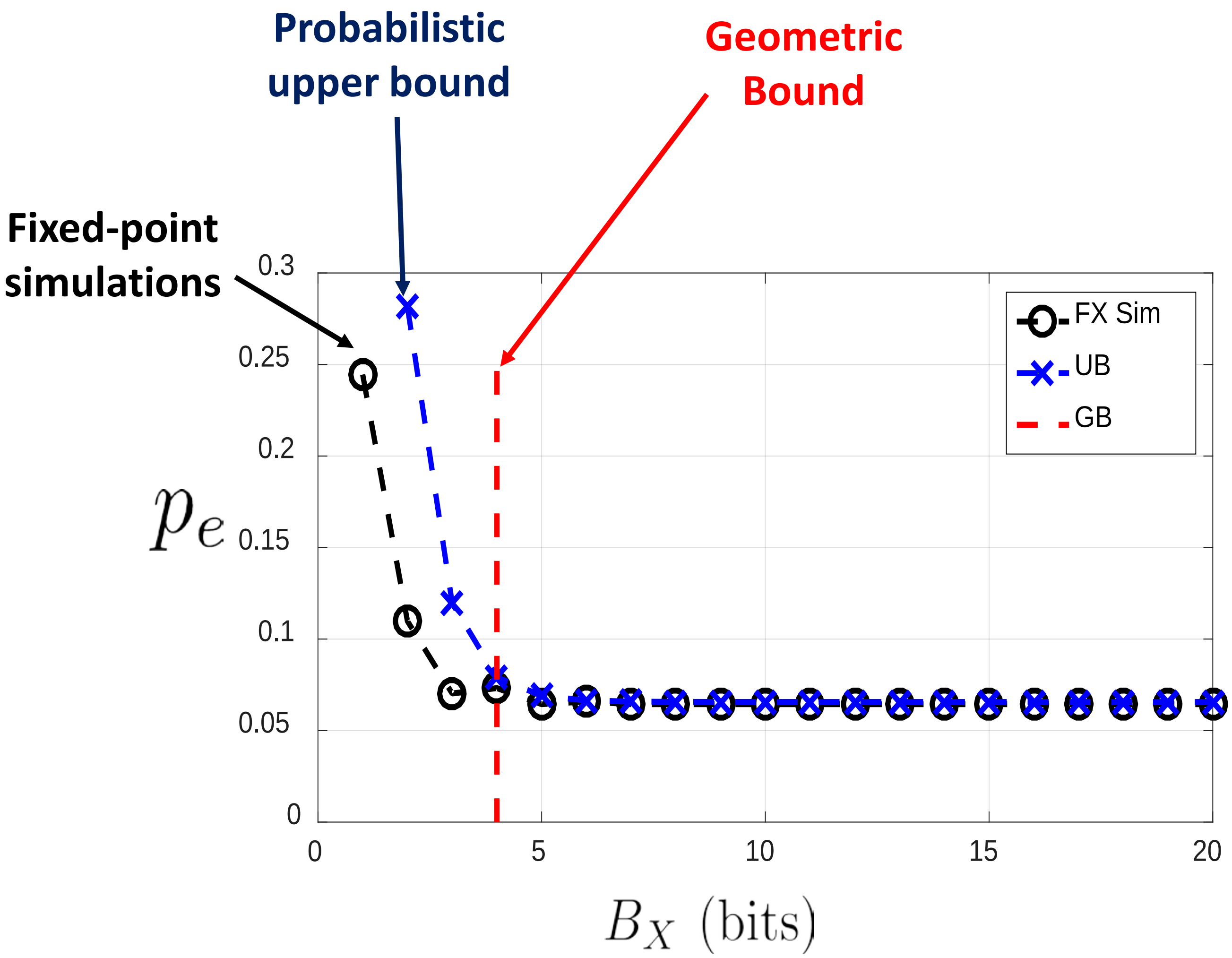
To ensure non-zero updates:

$$B_W \geq B_X - \log_2(\gamma)$$

Simulation Results

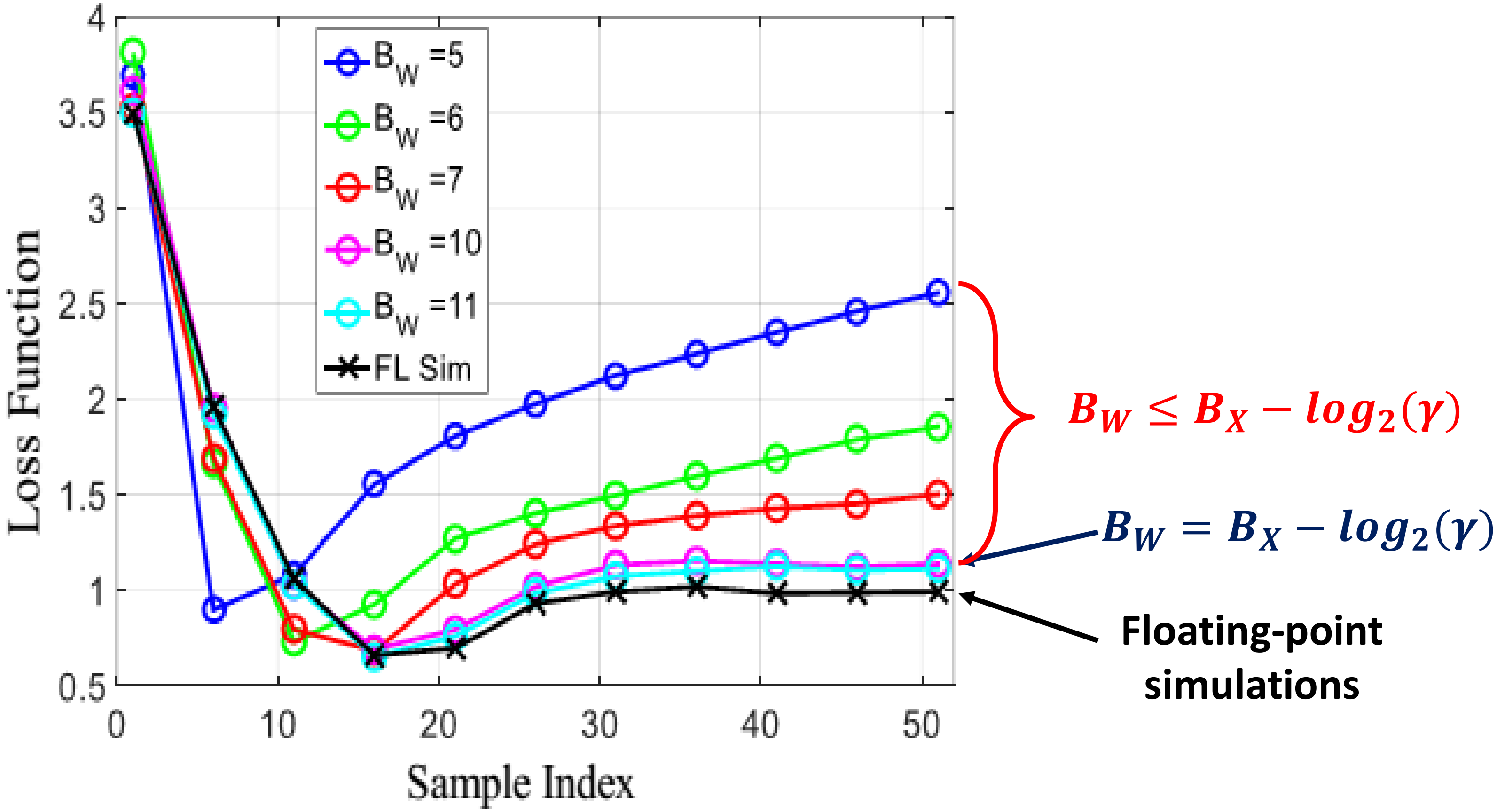
- Dataset: Breast Cancer Dataset from UCI Machine Learning Repository.
- Classification: Fix B_F and sweep B_X. We compare fixed-point simulations to analysis.
- Training: Fix B_X and B_F and sweep B_W. We compare floating-point convergence curves to fixed-point simulations.
- Energy estimation: We use the methodology from (Abdallah & Shanbhag, TVLSI, 2014) on a 45 nm CMOS process to estimate the energy savings of reducing precision.

Classification



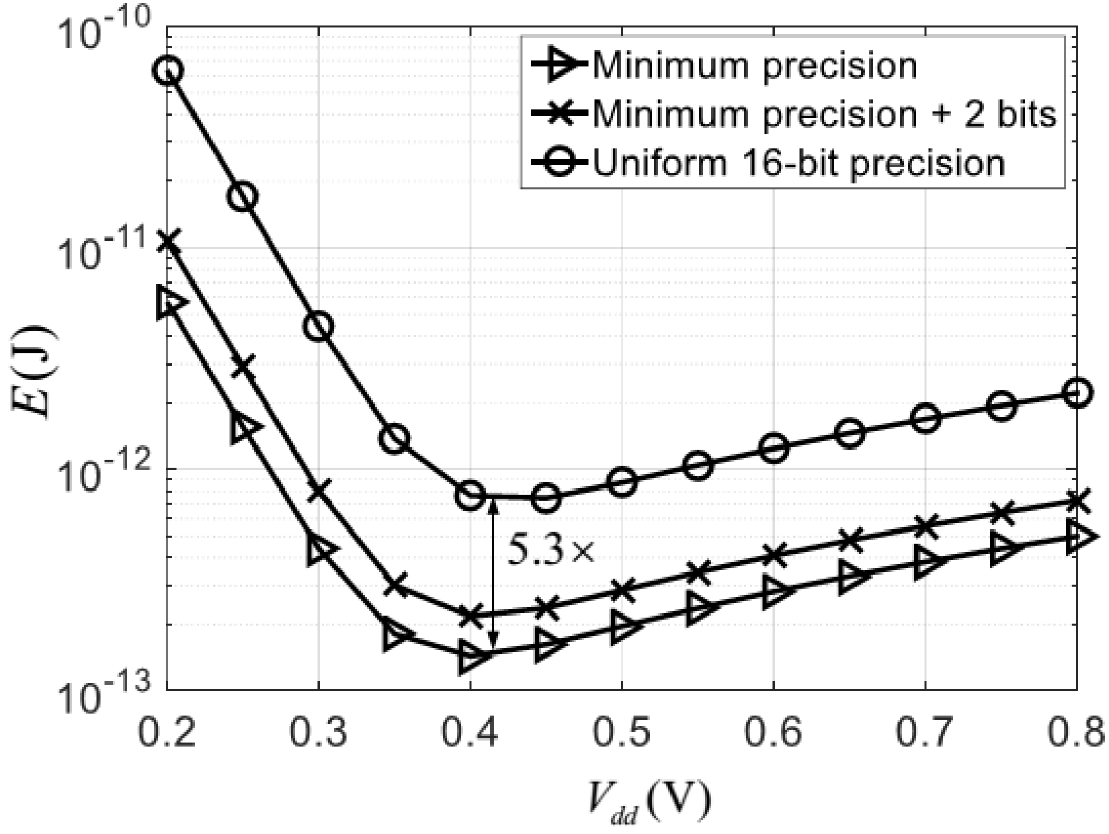
$B_F = 6$

Training



$B_X = 6\,, B_F = 6, \gamma = 2^{-5}$

Energy Savings



Conclusion

- We presented analytical requirements on the fixed-point precision in the context of the SVM-SGD algorithm.
- Ongoing Work: Extension of results for complete deep learning systems.

Acknowledgments

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA.

Thank you!