Imperial College London

HiPEDS

Modeling Round-off Error in the Fast Gradient Method for Predictive Control

Ian McInerney[a], Eric C. Kerrigan[a,b], George A. Constantinides[a]

[a]Department of Electrical and Electronic Engineering, Imperial College London

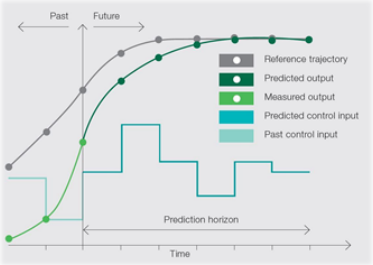
[b]Department of Aeronautics, Imperial College London

HiPEDS

Outline

1. Preliminaries
2. Generic round-off error model
3. Parametric round-off error model
4. Conclusions

Model Predictive Control



$$\min_{u,x} \frac{1}{2} x_N' P x_N + \frac{1}{2} \sum_{k=0}^{N-1} x_k' Q x_k + u_k' R u_k$$

$$\text{s.t. } x_{k+1} = A x_k + B u_k, \; k = 0, \ldots, N-1$$

$$x_0 = \bar{x}_0$$

$$E u_k \leq c_u, \; k = 0, \ldots, N-1$$

$$\min_{u} \frac{1}{2} u' H u + \bar{x}_0' J' u$$

$$\text{s.t. } u \in \mathbb{K} := \{u : Gu \leq g\}$$

Fast Gradient Method

- Accelerated first-order optimization solver
- Project onto the constraint set
- Necessary condition for stability: $\lambda(H) \in (0, 1)$

Algorithm

for $i = 0$ to $I_{max} - 1$ **do**

$f^{(i)} := Hy^{(i)} + J\bar{x}_0$ ▷ Gradient

$t^{(i)} := y^{(i)} - f^{(i)}$ ▷ Update

$u^{(i+1)} := \pi_{\mathbb{K}}(t^{(i)})$ ▷ Projection

$y^{(i+1)} := (1 + \beta)t^{(i+1)} - \beta t^{(i)}$ ▷ Acceleration

end for

Fixed-Point Arithmetic

- Quantize $\mathbb{R}$ into $\mathbb{Z}$ by multiplying by a scale factor and rounding

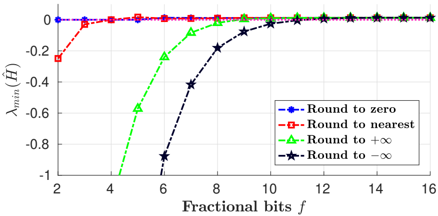
$$\overbrace{18020.017052}^{\mathbb{R}} \xrightarrow{\times 10^5} \overbrace{\underbrace{18020}_{\text{integer}}\underbrace{01705}_{\text{fraction}}}^{\text{Fixed-point}}$$

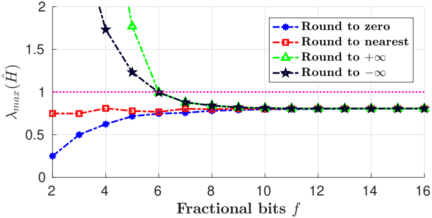
- Only requires integer computations in hardware
- Introduces ϵ_f round-off error

Model rounding action as an additive disturbance matrix:

$$\hat{H} = H + E$$

Rounding Stability Margin





Rounding Stability Margin (η)

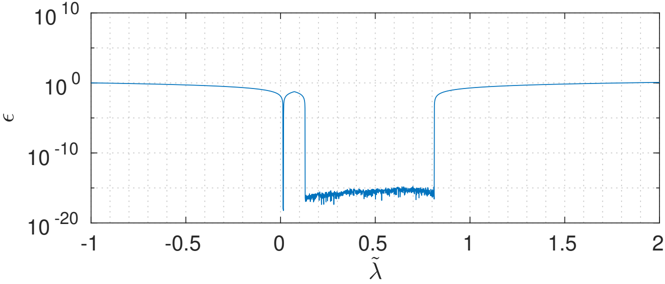
The smallest value of $\|E\|_2$ that causes the eigenvalues of $\hat{H}$ to leave the interval $(0, 1)$.

Rounding Stability Margin

Computation

Assuming $\lambda(H) \in (0, 1)$, use the ϵ-pseudospectrum of H at $\tilde{\lambda} = \{0, 1\}$.

$$\left\|(\tilde{\lambda} I - H)^{-1}\right\| \geq \frac{1}{\epsilon} \xleftrightarrow{\text{Equivalently}} \tilde{\lambda} \in \lambda(H + E) \text{ with } \|E\| \leq \epsilon$$



Generic Rounding Model

Model

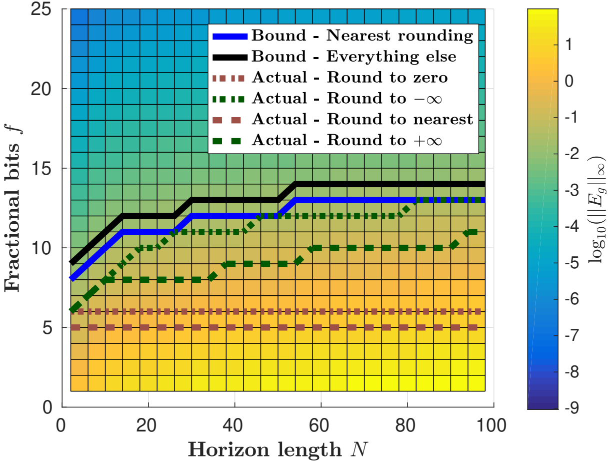
$$E_g := \begin{bmatrix} \pm\epsilon_f & \pm\epsilon_f & \dots \\ \pm\epsilon_f & \pm\epsilon_f & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Fractional Bits Sufficient for $\lambda(\hat{H}) \in (0, 1)$

$$f = \begin{cases} \left\lceil -\log_2\left(\frac{\eta}{mN}\right)\right\rceil - 1 & \text{if using round to nearest,} \\ \left\lceil -\log_2\left(\frac{\eta}{mN}\right)\right\rceil & \text{otherwise.} \end{cases}$$

HiPEDS

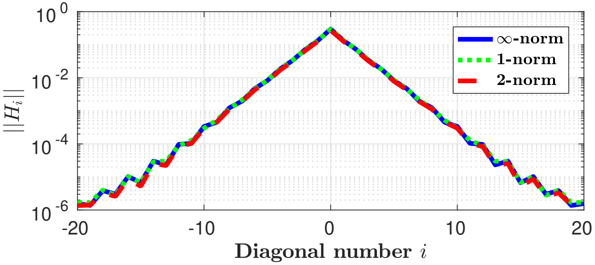
Results



Hessian Structure

Toeplitz Components

$$H_i = \begin{cases} B'(A^i)'PB & \text{if } i > 0, \\ B'PB + R & \text{if } i = 0, \\ B'PA^{|i|}B & \text{if } i < 0. \end{cases}$$



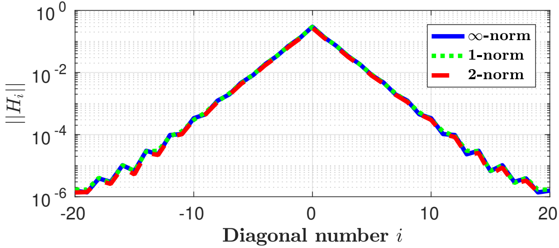
Parametric Rounding Model

Idea

Split round-off error model into two parts: $E_p := E_G + E_T$, with

$$(E_G)_i := \begin{cases} E_g & \text{if } i < k, \\ 0 & \text{otherwise,} \end{cases} \qquad (E_T)_i := \begin{cases} H_i & \text{if } |i| \geq k, \\ 0 & \text{otherwise.} \end{cases}$$

Where k is the diagonal beyond which all blocks in $\hat{H}_i$ are 0



Parametric Rounding Model

Fractional Width Selection

Choose ϵ_f such that

$$|\epsilon_f| m(2k-1) + 2\|\mathcal{P}_{\bar{H}}(k,\cdot)\|_{H_\infty} < \eta,$$

where

$$\mathcal{P}_{\bar{H}}(n,z) := z\mathcal{G}_P(z) - B'P\mathcal{P}_n(z)B \quad \forall z \in \mathbb{T},$$

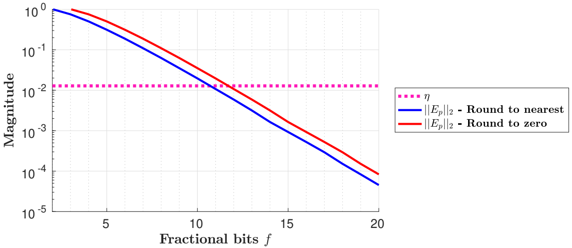
$$\mathcal{G}_P := \begin{cases} x^+ = Ax + Bu \\ y = B'Px \end{cases},$$

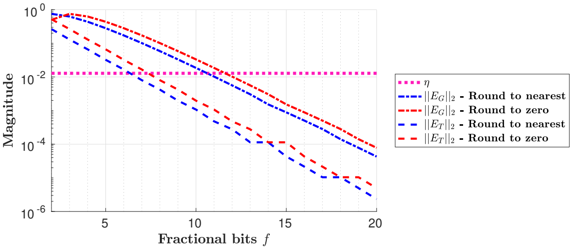
$$\mathcal{P}_n(z) := \sum_{i=0}^{n-1} A^i z^{-i} \quad \forall z \in \mathbb{T}.$$

Imperial College London

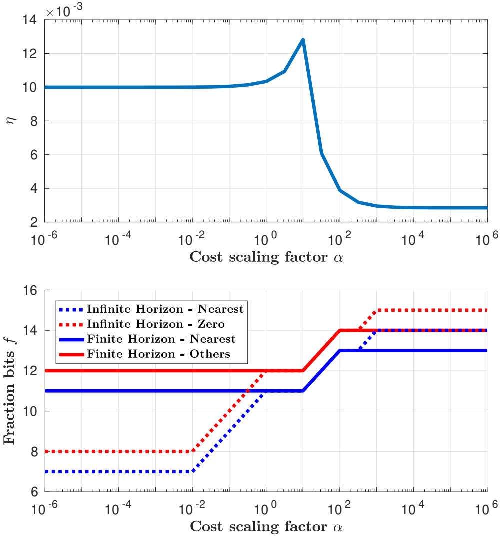
HiPEDS

Results

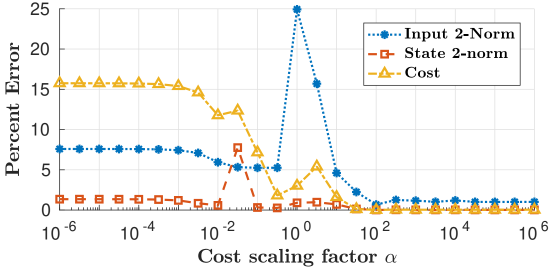




Results


×10^-3
14
12
10
8
6
4
2
η
10^-6
10^-4
10^-2
10^0
10^2
10^4
10^6
Cost scaling factor α
16
14
12
10
8
6
Fraction bits f
Infinite Horizon - Nearest
Infinite Horizon - Zero
Finite Horizon - Nearest
Finite Horizon - Others
10^-6
10^-4
10^-2
10^0
10^2
10^4
10^6
Cost scaling factor α

Results



Fractional Length	Logic Resources LUT	FF	DSP	BRAM	Power (mW)	Solve Time (μs)
f=12	947	768	4	2	20	532.17
f=16	1,136	912	4	2	25	612.17
f=21	887	1,033	8	8	43	701.77
f=26	993	1,237	12	9	48	701.77

Conclusions

Key Contributions

- Compute the Rounding Stability Margin for the Fast Gradient Method using the ϵ-pseudospectrum
- Exploit Toeplitz structure to model exact round-off error
- Reduce the fractional bits needed by 30–45%
- Reduction in hardware usage and solution time by up to 77% and 25% respectively

Future Directions

- Derive sufficient condition for stability of the Fast Gradient Method
- Bound perturbation of optimal vector u^* when problem is quantized