QUANTISATION EFFECTS IN PDMM: A FIRST STUDY FOR SYNCHRONOUS DISTRIBUTED AVERAGING

Daan H. M. Schellekens[a,b], Thomas Sherson[a], and Richard Heusdens[a]

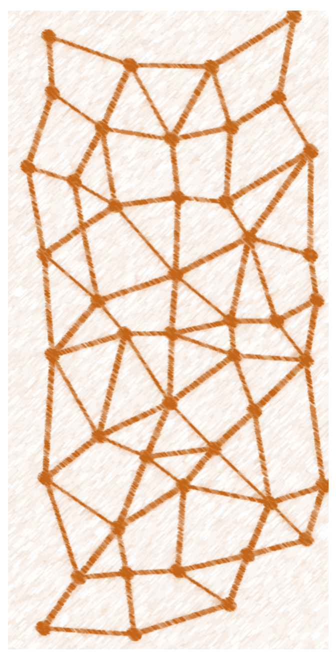
dschellekens@pindrop.com, t.w.sherson@tudelft.nl, r.heusdens@tudelft.nl

INTRODUCTION

Large-scale sensor networks have become commonplace in many applications. To facilitate signal processing in these decentralized networks, iterative distributed algorithms are required.

DISTRIBUTED SIGNAL PROCESSING METHODS

- Distributed consensus [1],
  gossip, path averaging, distributed averaging
- Probabilistic inference [2],
  max-sum and sum-product message-passing, loopy belief
- Convex optimization,
  sub-gradient, primal-dual methods
  - Alternating-direction method of multipliers (ADMM) [3]
  - Primal-dual method of multipliers (PDMM) [4]



PROBLEM STATEMENT

The iterative behaviour of distributed processing algorithms combined with energy, computational power, and bandwidth limitations imposed by these networks, place tight constraints on the transmission capacities of the individual nodes.

The effect of quantisation on the final accuracy and the convergence rate of some distributed (iterative) algorithms has been investigated, including ADMM [5], but no such results are known for PDMM. This paper is a first attempt to investigate the effect of quantisation on synchronous PDMM.

PRIMAL-DUAL METHOD OF MULTIPLIERS

PDMM solves a separable convex optimisation problem over a graph $G = (V, E)$, with $|V| = n$ and $|E| = d$, of the form

$$\min_{\mathbf{x}} \quad \sum_{i \in V} f_i(\mathbf{x}_i)$$
$$\text{subject to} \quad A_{ij}\mathbf{x}_i + A_{ji}\mathbf{x}_j = \mathbf{c}_{ij}, \quad \forall (i,j) \in E,$$

where $\mathbf{x}_i \in \mathbb{R}^{n_i}$, $\mathbf{x} = (\mathbf{x}_1^T, \ldots, \mathbf{x}_n^T)^T$. The update equations for the primal $\mathbf{x} \in \mathbb{R}^{\Sigma n_i}$ and dual $\boldsymbol{\mu} \in \mathbb{R}^d$ domain can be written as a linear system of equations, with vector $\mathbf{y} = (\mathbf{x}^T, \boldsymbol{\mu}^T)^T$ and input $\mathbf{u}$

$$\mathbf{y}^{(k)} = F^k \mathbf{y}^{(0)} + \sum_{i=0}^{k-1} F^i \mathbf{u}.$$

Scalar averaging problem: $f_i(x_i) = \frac{1}{2}(x_i - t_i)^2$, implies $\Sigma n_i = n$.
Assumptions: $\mathbf{y}^{(0)} = \mathbf{0}$ and F is diagonalisable.

With this, the system can be simplified to

$$\mathbf{y}^{(k)} = V_2 \left(\sum_{i=0}^{k-1} \Lambda_2^i \right) V_2^+ \mathbf{u},$$

where Λ_2 contains the complex eigenvalues λ_2 having magnitude strictly less than unity [4]. By defining the error as $\mathbf{e}^{(k)} = \mathbf{y}^{(k)} - \mathbf{y}^*$ we obtain

$$\mathbf{e}^{(k)} = -V_2 \Lambda_2^k \left(I - \Lambda_2 \right)^{-1} V_2^+ \mathbf{u},$$

which shows that the primal squared error $\zeta_x^{(k)}$ converges at a rate $|\lambda_{2,\max}|^{2k}$.

UNIFORM SUBTRACTIVE-DITHER QUANTISATION

Let $Q(\tilde{\mathbf{y}}^{(k)}) = \tilde{\mathbf{y}}^{(k)} + \mathbf{n}_{\mathrm{q}}^{(k)}$ be the uniformly quantised version of $\tilde{\mathbf{y}}^{(k)}$, with reproduction distance $\Delta_{(k)}$. The error then becomes

$$\tilde{\mathbf{e}}^{(k)} = \mathbf{e}^{(k)} + \sum_{i=0}^{k-1} F^{k-i} \mathbf{n}_{\mathrm{q}}^{(i)}.$$

By adding pseudo-random subtractive dither realisations before quantisation, realisations of $\mathbf{n}_{\mathrm{q}}^{(k)}$ can be made

1. i.i.d. uniformly distributed;
2. statistically independent in time of one another.

These two properties of the quantisation noise realisations allow us to write the primal mean squared error (MSE) as

$$\mathbb{E}\left[\tilde{\zeta}_x^{(k)}\right] = \mathbb{E}\left[\zeta_x^{(k)}\right] + \mathbb{E}\left[\zeta_{\mathrm{q},x}^{(k)}\right].$$

If we chose

$$\Delta_{(k)} = |\lambda_{2,\max}|^k \Delta_{(0)},$$

then, again due to dithering, $\mathbb{E}\left[\zeta_{\mathrm{q},x}^{(k)}\right]$ converges at a rate $k \cdot |\lambda_{2,\max}|^{2k}$, while $\mathbb{E}\left[\zeta_x^{(k)}\right]$ converges at a rate $|\lambda_{2,\max}|^{2k}$.

Using a quantiser with decreasing cell width suggests an increase in data rate with increasing iterations. This, however, will *almost* entirely be compensated by the decrease of information that needs to be transmitted with increasing iterations.

SIMULATION RESULTS

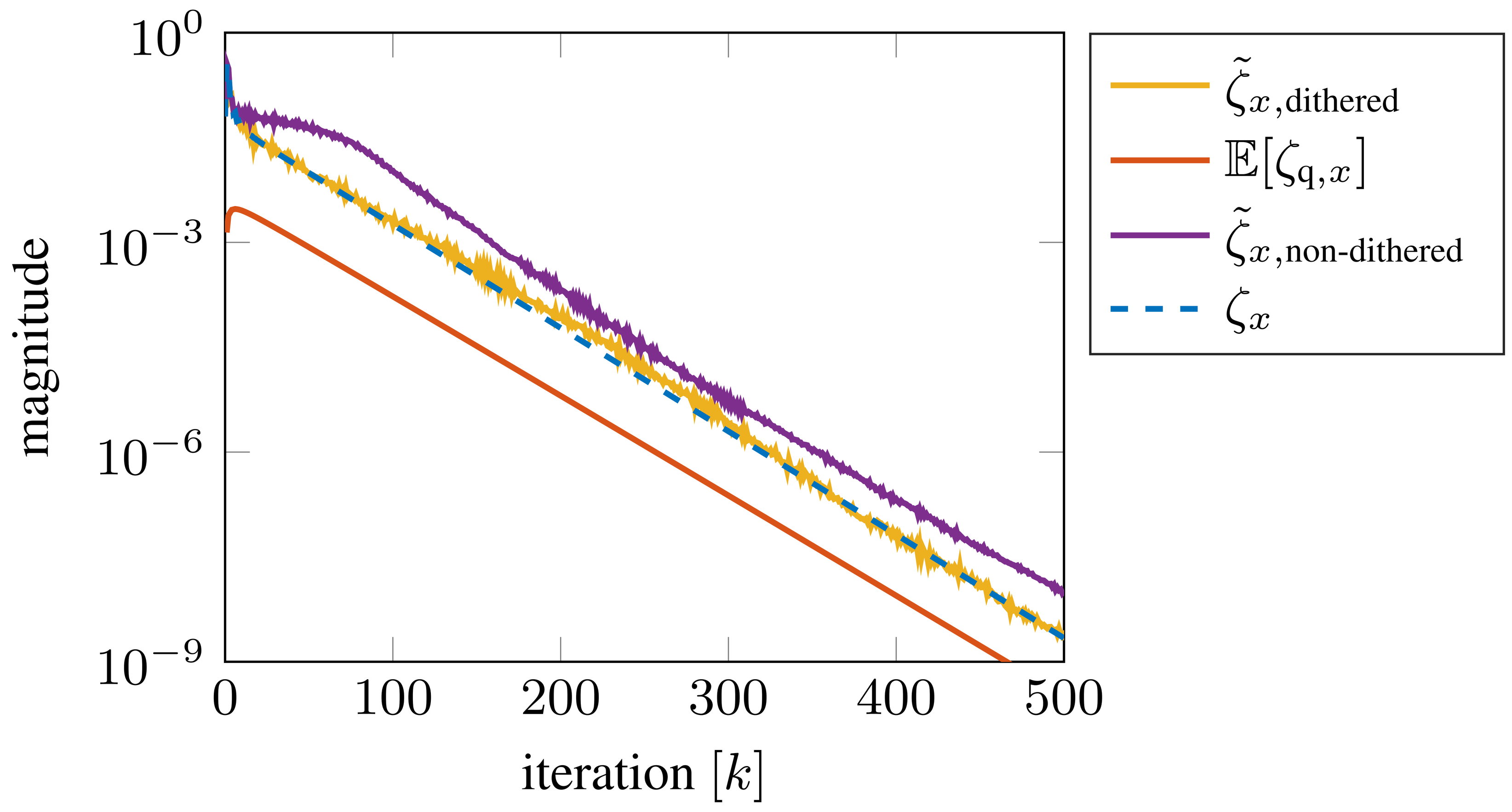


Fig. 1. The primal squared error $\tilde{\zeta}_x$ for quantised PDMM with decreasing cell width $\Delta_{(k)}$ with $\Delta_{(0)} = 10^{-1}$ and $n = 10$.

CONCLUSION

It was found that, for practical applications, quantisation in PDMM can be applied with a fixed-rate quantiser, such that significant data rate reduction can be achieved, without compromising the rate of convergence.

[1] D. Shah, "Gossip algorithms", Now Publishers Inc., 2009.
[2] C. Bishop, "Pattern recognition and machine learning (Information Science and Statistics)", Springer, New York, corr. second printing edition, 2007.
[3] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," in Found. Trends Mach. Learn., vol. 3, no. 1, pp. 1–122, Jan. 2011.
[4] G. Zhang and R. Heusdens, "Distributed optimization using the primal-dual method of multipliers," in IEEE Trans. on Signal and Inform. Processing over Networks, Feb. 2017
[5] S. Zhu and B. Chen, "Quantized consensus by the ADMM: Probabilistic versus deterministic quantizers," in IEEE Trans. on Signal Processing, vol. 64, no. 7, pp. 1700–1713, Apr. 2016.

[a]Circuits and Systems, Delft University of Technology, Delft, The Netherlands

[b]Pindrop
Atlanta, United States
London, United Kingdom