Optimal error bounds for Mann fixed-point iterations via optimal transport

Roberto Cominetti
Universidad Adolfo Ibáñez

Based on joint work with
Mario Bravo, Thierry Champion, Juan Pablo Contreras

"In memory of Professor Ronald E. Bruck"

Workshop on Nonlinear Functional Analysis and Its Applications

Technion, April 4-6, 2022

Mann's sequential averaging

$T : C \to C$ non-expansive / C convex bounded in $(X, \|\cdot\|)$ normed space.
Remark: By rescaling the norm we may and will assume diam$(C) \le 1$.

Starting from x^0 and $y^0 =$ "$T(x^{-1})$", we study the fixed-point iteration

(M) $$x^n = \sum_{i=0}^{n} \pi_i^n \, Tx^{i-1} \qquad (\forall\, n \ge 1)$$

defined by a triangular array of averaging coefficients

$$\pi = \begin{bmatrix} \pi_0^0 & & & & & \\ \pi_0^1 & \pi_1^1 & & & & \\ \pi_0^2 & \pi_1^2 & \pi_2^2 & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ \pi_0^n & \pi_1^n & \pi_2^n & \cdots & \pi_n^n & \\ \vdots & \vdots & \vdots & & \vdots & \ddots \end{bmatrix}$$

$$\pi_i^n \ge 0 \;\; ; \;\; \pi_i^n = 0 \text{ for } i > n \;\; ; \;\; \sum_{i=0}^{n} \pi_i^n = 1.$$

Special cases: $\{\alpha_n\}_{n\geq 1}, \{\beta_n\}_{n\geq 1}$ *sequences in* $[0,1]$

Long history since the early 1950's, dozens of papers mostly in Hilbert settings.

Krasnoselskii-Mann:	$x^n = (1-\alpha_n)x^{n-1} + \alpha_n Tx^{n-1}$	$\Leftrightarrow \ \pi^n = (1-\alpha_n)\pi^{n-1} + \alpha_n \delta^n$
Halpern:	$x^n = (1-\beta_n)y^0 + \beta_n Tx^{n-1}$	$\Leftrightarrow \ \pi^n = (1-\beta_n)\delta^0 + \beta_n \delta^n$
Inertial Halpern:	$x^n = (1-\alpha_n-\beta_n)y^0 + \beta_n Tx^{n-2} + \alpha_n Tx^{n-1}$	
Inertial KM:	$x^n = (1-\alpha_n-\beta_n)x^{n-1} + \beta_n Tx^{n-2} + \alpha_n Tx^{n-1}$	
Extra KM:	$x^n = (1-\alpha_n-\beta_n)x^{n-2} + \beta_n x^{n-1} + \alpha_n Tx^{n-1}.$	
KM-Halpern:	$x^n = (1-\alpha_n-\beta_n)y^0 + \beta_n x^{n-1} + \alpha_n Tx^{n-1}$	
Ishikawa:	$\begin{cases} x^{2n+1} = (1-\beta_n)x^{2n} + \beta_n Tx^{2n} \\ x^{2n+2} = (1-\alpha_n)x^{2n} + \alpha_n Tx^{2n+1} \end{cases}$; $\alpha_n \leq \beta_n$	

- ▷ **fixed points:** algorithms to compute & prove existence of fixed points
- ▷ **convex optimization:** Gradient, Prox, Douglas-Rachford, ADMM, POCS...
- ▷ **stochastic dynamics:** Q-learning in MDPs, lazy random walks, Langevin dynamics
- ▷ **evolution equations:** discretization of $\frac{dx}{dt} + [I-T](x) = 0$

Basic questions:

(M) $$x^n = \sum_{i=0}^n \pi_i^n \, Tx^{i-1} \qquad (\forall\, n \geq 1)$$

1. (*Analysis*) For any given sequence of π^n's determine

a)	$\|x^n - Tx^n\| \to 0$?	(Asymptotic Regularity)
b)	How fast ?	(Rate of Convergence)
c)	Non-asymptotic estimates ?	(Error Bounds)

2. (*Design*) Find π^n's that minimize the worst-case residuals

$$\underset{\pi}{\text{Minimize}} \;\; \Psi_n(\pi) \triangleq \sup_{T,x^0,y^0} \|x^n - Tx^n\|$$

...looks intractable but not really as we shall see.

Previous results for KM: $x^n = (1-\alpha_n)x^{n-1} + \alpha_n Tx^{n-1}$

Conjecture (Baillon-Bruck'1992)

For KM iterates, there exists a universal constant κ such that

$$\|x^n - Tx^n\| \leq \kappa \frac{\text{diam}(C)}{\sqrt{\sum_{i=1}^{n} \alpha_i(1-\alpha_i)}}. \tag{BB}$$

Theorem (Baillon-Bruck'1996)

For constant $\alpha_n \equiv \alpha$, (BB) holds with $\kappa = 1/\sqrt{\pi}$ and rate $\sim O(1/\sqrt{n})$.

Previous results for KM: $x^n = (1-\alpha_n)x^{n-1} + \alpha_n Tx^{n-1}$

Conjecture (Baillon-Bruck'1992)

For KM iterates, there exists a universal constant κ such that

$$\|x^n - Tx^n\| \le \kappa \frac{\mathrm{diam}(C)}{\sqrt{\sum_{i=1}^n \alpha_i(1-\alpha_i)}}. \tag{BB}$$

Theorem (Baillon-Bruck'1996)

For constant $\alpha_n \equiv \alpha$, (BB) holds with $\kappa = 1/\sqrt{\pi}$ and rate $\sim O(1/\sqrt{n})$.

Our contribution (C-Soto-Vaisman'2014, Bravo-C'2016, Bravo-C-Pavez'2017)

1. (BB) holds for general α_n with $\kappa = 1/\sqrt{\pi} \sim 0.5642$
2. Nonlinear maps: the constant $\kappa = 1/\sqrt{\pi}$ is tight
3. Affine maps: tight bound with $\kappa = \max_z \sqrt{z}e^{-z}I_0(z) \sim 0.4688$
4. Error bounds for inexact KM: $x^n = (1-\alpha_n)\,x^{n-1} + \alpha_n(Tx^{n-1} + \varepsilon^n)$

In this talk...

$$\Psi_n(\pi) \triangleq \sup_{T,x^0,y^0} \|x^n - Tx^n\|$$

- Computable tight error bounds for Mann : $\Psi_n(\pi) = R_n(\pi)$
- Lower bound for Krasnoselskii-Mann : $R_n(\pi) \geq \frac{1}{\sqrt{n+1}}$
- Lower bound for general Mann : $R_n(\pi) \geq \frac{1}{n+1}$
- Tight bounds for classical Halpern : $R_n(\pi) \sim O(\frac{1}{n})$
- Optimal β_n for Halpern : $R_n(\pi) \leq \frac{4}{n+4}$

Optimal Mann $\approx$ Optimal Halpern

(up to a small factor)

Tight error bounds via optimal transport

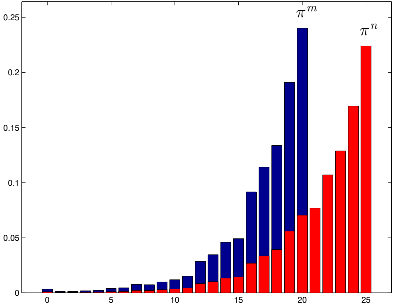
A path towards error bounds...

If some estimates $\|x^m - x^n\| \leq d_{m,n}$ are available then

$$\begin{aligned} \|x^n - Tx^n\| &= \left\| \sum_{i=0}^{n} \pi_i^n (Tx^{i-1} - Tx^n) \right\| \\ &\leq \sum_{i=0}^{n} \pi_i^n \|x^{i-1} - x^n\| \\ &\leq \sum_{i=0}^{n} \pi_i^n d_{i-1,n} \triangleq R_n \end{aligned}$$

How do we find good estimates $d_{m,n}$?

Recursive bounds: $\|x^m - x^n\| \leq d_{m,n}$

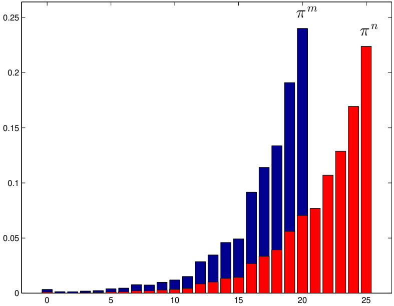


Let $Z(\pi^m, \pi^n)$ be the set of transport plans $z \geq 0$ taking π^m to π^n

$$\pi_i^m = \sum_{j=0}^{n} z_{ij}$$
$$\pi_j^n = \sum_{i=0}^{m} z_{ij}$$

$$x^m - x^n = \sum_{i=0}^{m} \pi_i^m\, Tx^{i-1} - \sum_{j=0}^{n} \pi_j^n\, Tx^{j-1} = \sum_{i=0}^{m} \sum_{j=0}^{n} z_{ij}\,[Tx^{i-1} - Tx^{j-1}]$$

Recursive bounds: $\|x^m - x^n\| \leq d_{m,n}$

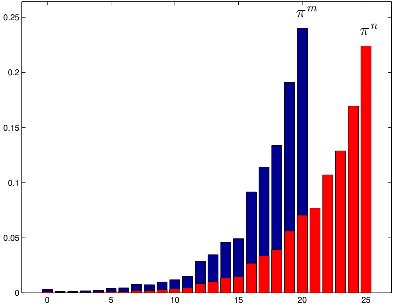


Let $Z(\pi^m, \pi^n)$ be the set of transport plans $z \geq 0$ taking π^m to π^n

$$\pi_i^m = \sum_{j=0}^n z_{ij}$$
$$\pi_j^n = \sum_{i=0}^m z_{ij}$$

$$\|x^m - x^n\| \leq \sum_{i=0}^m \sum_{j=0}^n z_{ij} \|x^{i-1} - x^{j-1}\| \leq \sum_{i=0}^m \sum_{j=0}^n z_{ij}\, d_{i-1,j-1}$$

Recursive bounds: $\|x^m - x^n\| \leq d_{m,n}$



Let $Z(\pi^m, \pi^n)$ be the set of transport plans $z \geq 0$ taking π^m to π^n

$$\pi_i^m = \sum_{j=0}^n z_{ij}$$
$$\pi_j^n = \sum_{i=0}^m z_{ij}$$

$$\|x^m - x^n\| \leq \sum_{i=0}^m \sum_{j=0}^n z_{ij} \|x^{i-1} - x^{j-1}\| \leq \sum_{i=0}^m \sum_{j=0}^n z_{ij}\, d_{i-1,j-1}$$

An optimal transport z minimizing the rhs yields a recursive bound $d_{m,n}$

Optimal transport error bounds

Set $d_{-1,-1} = 0$, $d_{-1,n} = 1$ for $n \geq 0$, and define inductively

$$\begin{aligned} d_{m,n} &= \min_{z \in Z(\pi^m,\pi^n)} \sum_{i=0}^{m} \sum_{j=0}^{n} z_{ij}\, d_{i-1,j-1} \\ R_n &= \sum_{i=0}^{n} \pi_i^n\, d_{i-1,n} \end{aligned}$$

Proposition (C-Soto-Vaisman'2014, Bravo-Champion-C'2021)

Mann iterates satisfy $\|x^m - x^n\| \leq d_{m,n}$ and $\|x^n - Tx^n\| \leq R_n$, for every nonexpansive $T : C \to C$ with C convex and *diam*$(C) \leq 1$.

Optimal transport error bounds

Set $d_{-1,-1} = 0$, $d_{-1,n} = 1$ for $n \geq 0$, and define inductively

$$\begin{aligned} d_{m,n} &= \min_{z \in Z(\pi^m,\pi^n)} \sum_{i=0}^{m} \sum_{j=0}^{n} z_{ij}\, d_{i-1,j-1} \\ R_n &= \sum_{i=0}^{n} \pi_i^n\, d_{i-1,n} \end{aligned}$$

Proposition (C-Soto-Vaisman'2014, Bravo-Champion-C'2021)

Mann iterates satisfy $\|x^m - x^n\| \leq d_{m,n}$ and $\|x^n - Tx^n\| \leq R_n$, for every nonexpansive $T : C \to C$ with C convex and *diam*$(C) \leq 1$.

These bounds depend only on $\pi = (\pi^n)_{n\in\mathbb{N}}$ and hold for every T, so that

$$\Psi_n(\pi) \triangleq \sup_{T,x^0,y^0} \|x^n - Tx^n\| \leq R_n(\pi).$$

In fact, these are the best possible bounds...

Optimal transport error bounds are tight...

Theorem (Bravo-C.'2016, Bravo-Champion-C'2021, C-Contreras'2021)

For any given sequence of π^n's, there exists a nonexpansive $T : C \to C$ on the unit cube $C = [0,1]^{\mathbb{N}} \subseteq \ell^\infty(\mathbb{N})$ and a corresponding Mann sequence satisfying $\|x^m - x^n\|_\infty = d_{m,n}$ and $\|x^n - Tx^n\|_\infty = R_n$ for all m, n.

Proof: T built from dual solutions of the optimal transports, which are then glued together using Aronszajn-Panitchpakdi's extension theorem. $\square$

$\Rightarrow \quad \Psi_n(\pi) \equiv R_n(\pi)$ provides tight error bounds !

Example – Krasnoselskii-Mann with $\alpha_n = \frac{1}{n+1}$

This gives a uniform distribution $\pi_i^n = \frac{1}{n+1}$ for $0 \leq i \leq n$

$$d_{m,n} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{2}{3} & \frac{3}{4} & \frac{4}{5} & \frac{5}{6} & \frac{6}{7} & \cdots \\ \frac{1}{2} & 0 & \frac{1}{4} & \frac{3}{8} & \frac{7}{15} & \frac{19}{36} & \frac{97}{168} & \cdots \\ \frac{2}{3} & \frac{1}{4} & 0 & \frac{23}{144} & \frac{47}{180} & \frac{1}{3} & \frac{47}{120} & \cdots \\ \frac{3}{4} & \frac{3}{8} & \frac{23}{144} & 0 & \frac{329}{2880} & \frac{1681}{8640} & \frac{1733}{6720} & \cdots \\ \frac{4}{5} & \frac{7}{15} & \frac{47}{180} & \frac{329}{2880} & 0 & \frac{7609}{86400} & \frac{7793}{50400} & \cdots \\ \frac{5}{6} & \frac{19}{36} & \frac{1}{3} & \frac{1681}{8640} & \frac{7609}{86400} & 0 & \frac{257219}{3628800} & \cdots \\ \frac{6}{7} & \frac{97}{168} & \frac{47}{120} & \frac{1733}{6720} & \frac{7793}{50400} & \frac{257219}{3628800} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$R_n = \left(1, \frac{3}{4}, \frac{23}{36}, \frac{329}{576}, \frac{7609}{14400}, \frac{257219}{518400}, \frac{2401963}{5080320}, \frac{245771507}{541900800}, \ldots\right)$$

No closed-form formula, but numerics suggest $R_n = \Omega(\frac{1}{\sqrt{\ln n}})$

Example – Krasnoselskii-Mann with $\alpha_n \equiv \alpha \geq \frac{1}{2}$

Here we have $\pi^n = (1-\alpha)\pi^{n-1} + \alpha\delta^n$, and we get

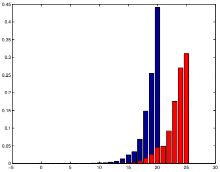
$$R_n(\alpha) = \frac{1}{\alpha} d_{n,n+1}(\alpha) \leq {}_2F_1\left(-n, \tfrac{1}{2}; 2; 4\alpha(1-\alpha)\right)$$

where ${}_2F_1$ is Gauss' hypergeometric function.

$$\begin{aligned} d_{6,10}(\alpha) = &\alpha(4 - 36\,\alpha + 328\,\alpha^2 - 2671\,\alpha^3 + 19853\,\alpha^4 - 132880\,\alpha^5 + 785003\,\alpha^6 \\ &- 4016624\,\alpha^7 + 17541102\,\alpha^8 - 64796454\,\alpha^9 + 201809157\,\alpha^{10} \\ &- 530670200\,\alpha^{11} + 1183318617\,\alpha^{12} - 2250818306\,\alpha^{13} + 3675506816\,\alpha^{14} \\ &- 5184593492\,\alpha^{15} + 6352439437\,\alpha^{16} - 6792441644\,\alpha^{17} + 6361687020\,\alpha^{18} \\ &- 5232669869\,\alpha^{19} + 3785701567\,\alpha^{20} - 2409974375\,\alpha^{21} + 1348858198\,\alpha^{22} \\ &- 662337623\,\alpha^{23} + 284299971\,\alpha^{24} - 106102624\,\alpha^{25} + 34171973\,\alpha^{26} \\ &- 9400913\,\alpha^{27} + 2178730\,\alpha^{28} - 417352\,\alpha^{29} + 64328\,\alpha^{30} \\ &- 7667\,\alpha^{31} + 663\,\alpha^{32} - 37\,\alpha^{33} + \alpha^{34}) \end{aligned}$$

...?????????

Metric properties of recursive optimal transports

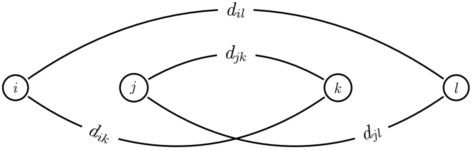


$$d_{m,n} = \min_{z \in Z(\pi^m,\pi^n)} \sum_{i=0}^{m} \sum_{j=0}^{n} z_{ij}\, d_{i-1,j-1}$$

Theorem (Aygen-Satik'1994, Bravo-C'2016, Bravo-Champion-C'2021)

The $d_{m,n}$'s define a distance in $\mathbb{N}$. For monotone π^n's ($\pi_i^{n+1} \leq \pi_i^n$ for $0 \leq i \leq n$), we have the stronger convex quadrangle inequality

$$d_{il} + d_{jk} \leq d_{ik} + d_{jl} \qquad \text{for all } i \leq j \leq k \leq l.$$



No flow-crossing in optimal transport

⇓

Greedy algorithm

⇓

Simpler expressions for $R_n(\pi)$

Optimal Mann iterations

Optimal Mann iteration – Fixed Horizon

Finding optimal π's for a fixed horizon n

$$\underset{\pi}{\text{Minimize}}\ \ \Psi_n(\pi) \equiv R_n(\pi)$$

can be formulated as a non-convex program

$$
\begin{array}{rcll}
\mathsf{FH}_n & \triangleq & \displaystyle\min_{\pi,z,d} & \sum_{i=0}^{n} \pi_i^n \, d_{i-1,n} & \\
& & \text{s.t.} & \pi^k \in \Delta^k & \forall k = 0,\dots,n \\
& & & z^{k,m} \in Z(\pi^k,\pi^m) & \forall 0 \le k < m \le n \\
& & & d_{-1,k} = 1 & \forall k = 0,\dots,n \\
& & & d_{k,k} = 0 & \forall k = -1,\dots,n \\
& & & d_{k,m} = \sum_{i=0}^{k}\sum_{j=0}^{m} z_{i,j}^{k,m}\, d_{i-1,j-1} & \forall 0 \le k < m \le n \\
& & & d_{m,k} = d_{k,m} & \forall 0 \le k < m \le n
\end{array}
$$

however state-of-the-art solvers have trouble solving beyond $n = 7$.

Sequential optimal Mann iteration

A simpler incremental approach minimizes $R_n(\pi)$ only with respect to π^n, by fixing the previous optima $\pi^1, \ldots, \pi^{n-1}$ and its corresponding $d_{i-1,j-1}$'s.

The n-th stage problem

$$\min_{\pi^n \in \Delta^n} R_n(\pi^n) \triangleq \sum_{k=0}^{n} \pi_k^n \, d_{k-1,n}(\pi^n)$$

can be formulated as

$$\begin{aligned} \mathsf{S}_n \;\triangleq\; \min_{\pi^n, z} \quad & \pi_0^n + \sum_{k=1}^{n} \sum_{i=0}^{k-1} \sum_{j=0}^{n} \pi_k^n \, z_{i,j}^{k-1,n} \, d_{i-1,j-1} \\ \text{s.t.} \quad & \pi^n \in \Delta^n \\ & z^{k,n} \in Z(\pi^k, \pi^n), \quad \forall k = 0, ..., n-1 \end{aligned}$$

Monotone optimal Mann iteration

Restricting to monotone π^n's with $\pi_i^{n+1} \leq \pi_i^n$ for $0 \leq i \leq n$ and $\pi_n^n \geq \frac{1}{2}$, the convex quadrangle inequality gives an explicit formula for the optimal transport

$$D_{m,n}(\pi^n) \triangleq \sum_{i=0}^{m}(\pi_i^m - \pi_i^n)\, d_{i-1,n-1} + \sum_{j=m+1}^{n} \pi_j^n\, (d_{m-1,j-1} - d_{m-1,n-1})$$

and the n-th stage sequential problem simplifies to

$$\begin{aligned} \mathsf{MS}_n \;\triangleq\; \min_{\pi^n} \quad & \pi_0^n + \sum_{i=1}^{n} \pi_i^n \, D_{i-1,n}(\pi^n) \\ \text{s.t.} \quad & \pi^n \in \Delta^n \\ & \pi_i^n \leq \pi_i^{n-1}, \quad \forall i = 0, ..., n-1 \\ & \pi_n^n \geq \tfrac{1}{2} \end{aligned}$$

A similar monotone version of FH_n yields an easier-to-solve problem MFH_n.

Numerical optimization: FH_n, MFH_n, S_n, MS_n

Property : $\mathsf{FH}_n \leq \mathsf{S}_n \leq \mathsf{MS}_n$

Conjecture : $\mathsf{FH}_n = \mathsf{MFH}_n$ and $\mathsf{S}_n = \mathsf{MS}_n$

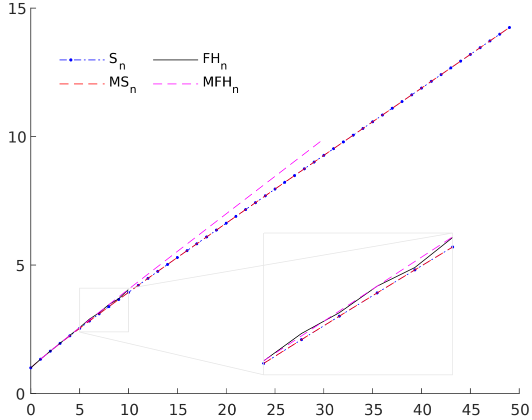
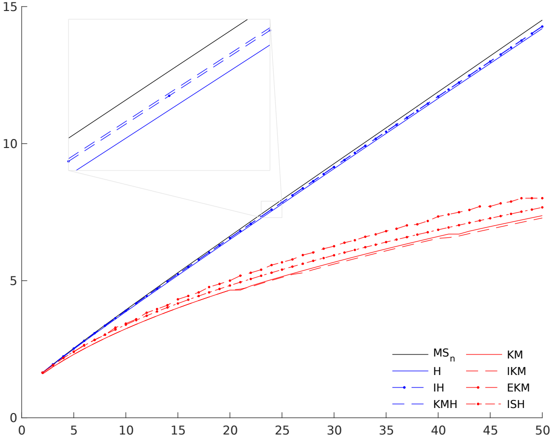


Figure: $\frac{1}{R_n}$ versus n

Optimal monotone-sequential vs Halpern and KM



MS minimum &
Halpern variants
$R_n \sim O(1/n)$

KM variants
$R_n \sim O(1/\sqrt{n})$

Figure: $\frac{1}{MS_n}$ versus n

Lower bounds & optimal Halpern

Lower bounds

The right-shift operator $T(x_0, x_1, x_2, \dots) = (0, x_0, x_1, x_2, \dots)$ is an isometry in every $\ell^p = \ell^p(\mathbb{N})$ for $1 \le p \le \infty$, with unique fixed point $x^* = (0, 0, \dots)$

Theorem (Contreras-C'2021)

a) *For every sequence* π^n *the Mann iterates for the right-shift* $T : \ell^\infty \to \ell^\infty$*, started from* $x^0 = y^0 = (1, 1, 1, \dots)$*, satisfy*

$$\|x^n - Tx^n\|_\infty \ge \tfrac{1}{n+1}.$$

b) *For every sequence* α_n *the Krasnoselskii-Mann iterates for the right-shift* $T : \ell^1 \to \ell^1$*, started from* $x^0 = y^0 = (1, 0, 0, ...)$*, satisfy*

$$\|x^n - Tx^n\|_1 \ge \tfrac{1}{\sqrt{n+1}}.$$

Tight bounds for Halpern: $x^n = (1-\beta_n)y^0 + \beta_n Tx^{n-1}$

Previous upper bounds for Halpern's iteration:

- (Sabach-Shtern'2017): for $\beta_n = \frac{n}{n+2}$ we have $R_n \leq \frac{4}{n+1}$
- (Leider'2021): for $\beta_n = \frac{n}{n+1}$ in Hilbert spaces $R_n \leq \frac{1}{n+1}$

The tight bounds for these β_n's in normed spaces involve the harmonic numbers

$$H_n = \sum_{k=1}^n \frac{1}{k} \sim \ln(n).$$

Theorem (Bravo-Champion-C'2021)

a) For Halpern with $\beta_n = \frac{n}{n+2}$ we have $R_n = \frac{4}{n+1}(1 - \frac{H_{n+2}}{n+2})$.

b) For Halpern with $\beta_n = \frac{n}{n+1}$ we have $R_n = \frac{H_{n+1}}{n+1}$.

Similar bounds can be obtained for various specific choices of β_n's.

Optimal Halpern: $x^n = (1-\beta_n)y^0 + \beta_n Tx^{n-1}$

Theorem (C-Contreras'2021)

- The optimal stepsizes for Halpern are $\beta_{n+1} = \frac{1}{2}(1+\beta_n^2)$ with $\beta_0 = 0$.
- They yield $\|x_n - Tx_n\| \leq R_n \leq \frac{4}{n+4}$ where the tight optimal bounds R_n satisfy the recursion $R_{n+1} = R_n - \frac{1}{4}R_n^2$ with $R_0 = 1$.

Theorem (C-Contreras'2021)

For affine nonexpansive maps the optimal Halpern is $\beta_n = \frac{n}{n+1}$ with $R_n = \frac{1}{n+1}$.

Remark: The latter coincides with Leider's bound for nonlinear maps in Hilbert spaces.

Thanks !

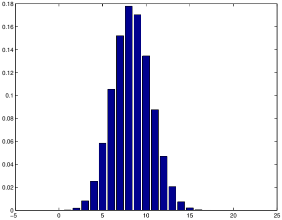
Papers available at
https://sites.google.com/site/cominettiroberto/

Example – Right-shift in $C=\{x \in \ell^1_+(\mathbb{N}) : \sum_{i=0}^{\infty} x_i = 1\}$

$T(x_0, x_1, x_2, \ldots) = (0, x_0, x_1, x_2, \ldots)$ is an ℓ^1-isometry

The KM iteration for this map yields

$$
\begin{aligned}
x^0 &= (1,0,0,0,\ldots)\\
x^1 &= (1-\alpha_1, \alpha_1, 0, 0, \ldots)\\
x^2 &= ((1-\alpha_2)(1-\alpha_1), (1-\alpha_2)\alpha_1 + \alpha_2(1-\alpha_1), \alpha_2\alpha_1, 0, \ldots)\\
x^3 &= \ldots
\end{aligned}
$$



$x_k^n = \mathbb{P}(X_1 + \cdots + X_n = k)$

$X_i \sim \text{Bernoulli}(\alpha_i)$

$\|x^n - Tx^n\|_1 = 2\max_k x_k^n$

Example – Lazy random walks

A random walk on a metric space (M, d) is described by $\mu^{n+1} = T\mu^n$, with the evolution operator T induced by a Markov transition kernel $(m_x)_{x\in M}$

$$T\mu(A) \triangleq \int_M m_x(A)\, d\mu(x) \qquad \forall A \in \mathcal{B}(M), \forall \mu \in \mathcal{P}_1(M).$$

- ▷ Fix (T) is the set of invariant measures
- ▷ T is nonexpansive in total variation $\|\cdot\|_{\text{tv}}$
- ▷ residuals $\|\mu^n - T\mu^n\|_{\text{tv}}$

For *lazy random walks* with kernel

$$m_x^\alpha = (1-\alpha)\delta_x + \alpha\, m_x$$

the evolution is a (KM) iteration.

Example – Optimal transport distances

For $\pi^n = (1-\alpha)\delta^{n-1} + \alpha\delta^n$ with $\alpha \geq \frac{1}{2}$

$$
\begin{gathered}
1\\
1\\
1\\
4\alpha - 6\alpha^2 + 4\alpha^3 - \alpha^4\\
-3 + 30\alpha - 90\alpha^2 + 130\alpha^3 - 90\alpha^4 + 24\alpha^5\\
4 - 10\alpha - 35\alpha^2 + 200\alpha^3 - 350\alpha^4 + 272\alpha^5 - 80\alpha^6\\
4 - 43\alpha + 178\alpha^2 - 340\alpha^3 + 280\alpha^4 + 2\alpha^5 - 144\alpha^6 + 64\alpha^7\\
0\\
37\alpha - 289\alpha^2 + 931\alpha^3 - 1510\alpha^4 + 1180\alpha^5 - 204\alpha^6 - 272\alpha^7 + 128\alpha^8\\
18 - 194\alpha + 821\alpha^2 - 1636\alpha^3 + 1255\alpha^4 + 814\alpha^5 - 2357\alpha^6 + 1728\alpha^7 - 448\alpha^8\\
-20 + 147\alpha - 252\alpha^2 - 686\alpha^3 + 3675\alpha^4 - 6825\alpha^5 + 6566\alpha^6 - 3276\alpha^7 + 672\alpha^8\\
-1 + 112\alpha - 952\alpha^2 + 3640\alpha^3 - 7770\alpha^4 + 9912\alpha^5 - 7532\alpha^6 + 3152\alpha^7 - 560\alpha^8\\
20 - 245\alpha + 1295\alpha^2 - 3745\alpha^3 + 6545\alpha^4 - 7119\alpha^5 + 4725\alpha^6 - 1755\alpha^7 + 280\alpha^8\\
-15 + 168\alpha - 756\alpha^2 + 1904\alpha^3 - 2940\alpha^4 + 2856\alpha^5 - 1708\alpha^6 + 576\alpha^7 - 84\alpha^8\\
7 - 56\alpha + 224\alpha^2 - 504\alpha^3 + 700\alpha^4 - 616\alpha^5 + 336\alpha^6 - 104\alpha^7 + 14\alpha^8\\
8\alpha - 28\alpha^2 + 56\alpha^3 - 70\alpha^4 + 56\alpha^5 - 28\alpha^6 + 8\alpha^7 - \alpha^8\\
1\\
1\\
1
\end{gathered}
$$

Figure: $d_{7,n}(\alpha)$ for $n = 0, \ldots, 18$

Example – Optimal transport bounds for KM

For $\pi^n = (1-\alpha)\pi^{n-1} + \alpha\delta^n \quad \Leftrightarrow \quad$ KM with constant $\alpha_n \equiv \alpha \geq \frac{1}{2}$

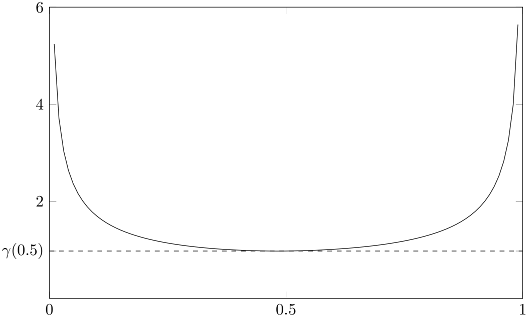
$$
\begin{aligned}
R_1 &= 1 - \alpha + \alpha^2 \\
R_2 &= 1 - 2\alpha + 4\alpha^2 - 4\alpha^3 + 2\alpha^4 \\
R_3 &= 1 - 3\alpha + 9\alpha^2 - 18\alpha^3 + 25\alpha^4 - 21\alpha^5 + 9\alpha^6 - \alpha^7 \\
R_4 &= 1 - 4\alpha + 16\alpha^2 - 48\alpha^3 + 112\alpha^4 - 192\alpha^5 + 230\alpha^6 - 180\alpha^7 \\
&\quad +84\alpha^8 - 20\alpha^9 + 2\alpha^{10} \\
R_5 &= 1 - 5\alpha + 25\alpha^2 - 100\alpha^3 + 331\alpha^4 - 876\alpha^5 + 1795\alpha^6 - 2762\alpha^7 \\
&\quad + 3106\alpha^8 - 2482\alpha^9 + 1366\alpha^{10} - 500\alpha^{11} + 117\alpha^{12} - 16\alpha^{13} + \alpha^{14} \\
R_6 &= 1 - 6\alpha + 36\alpha^2 - 180\alpha^3 + 775\alpha^4 - 2806\alpha^5 + 8324\alpha^6 - 19778\alpha^7 \\
&\quad + 37023\alpha^8 - 53948\alpha^9 + 60623\alpha^{10} - 52122\alpha^{11} + 34044\alpha^{12} - 16770\alpha^{13} \\
&\quad + 6163\alpha^{14} - 1652\alpha^{15} + 308\alpha^{16} - 36\alpha^{17} + 2\alpha^{18} \\
R_7 &= 1 - 7\alpha + 49\alpha^2 - 294\alpha^3 + 1562\alpha^4 - 7222\alpha^5 + 28408\alpha^6 - 93187\alpha^7 \\
&\quad + 251365\alpha^8 - 552678\alpha^9 + 985643\alpha^{10} - 1422448\alpha^{11} + 1660135\alpha^{12} \\
&\quad - 1567511\alpha^{13} + 1198337\alpha^{14} - 741914\alpha^{15} + 371352\alpha^{16} - 149443\alpha^{17} \\
&\quad + 47802\alpha^{18} - 11909\alpha^{19} + 2233\alpha^{20} - 297\alpha^{21} + 25\alpha^{22} - \alpha^{23}
\end{aligned}
$$

Degrees: $\{2, 4, 7, 10, 14, 18, 23, 28, 34, 40, 47, 54, 62, 70, 79, 88, 98, 108, 119, \ldots\}$
Coincide with $\lfloor (n^2+6n+1)/4 \rfloor$ up to $n = 40$, the solution of the *postage stamp problem* with n stamps and 2 denominations.

Best constant stepsize for KM: $\alpha_n \equiv \alpha$?

$$\| Tx^n - x^n \| \leq \gamma(\alpha)/\sqrt{n}$$

$$\gamma(\alpha) = \sup_{n\in\mathbb{N}} \sqrt{n}\, R_n(\alpha)$$



The minimum of $\gamma(\alpha)$ is attained near $\alpha \approx 0.4623$

Which is the best α if n is fixed, say $n = 150$?

Krasnoselskii's original iteration $\alpha_n \equiv \frac{1}{2}$

$$\gamma(\alpha) = \sup_{n\in\mathbb{N}} \sqrt{n}\, R_n(\alpha)$$

For $\alpha = \frac{1}{2}$ the sup seems to be attained at $n = 8$

$$\begin{aligned} R_8(\alpha) \;=\; & 1 - 8\,\alpha + 64\,\alpha^2 - 448\,\alpha^3 + 2835\,\alpha^4 - 16008\,\alpha^5 + 79034\,\alpha^6 \\ & - 334908\,\alpha^7 + 1201873\,\alpha^8 - 3622324\,\alpha^9 + 9129380\,\alpha^{10} \\ & - 19214722\,\alpha^{11} + 33796129\,\alpha^{12} - 49776610\,\alpha^{13} + 61566687\,\alpha^{14} \\ & - 64152608\,\alpha^{15} + 56488500\,\alpha^{16} - 42133404\,\alpha^{17} + 26651679\,\alpha^{18} \\ & - 14288252\,\alpha^{19} + 6472429\,\alpha^{20} - 2462126\,\alpha^{21} + 778478\,\alpha^{22} \\ & - 201354\,\alpha^{23} + 41584\,\alpha^{24} - 6604\,\alpha^{25} + 758\,\alpha^{26} - 56\,\alpha^{27} + 2\,\alpha^{28} \end{aligned}$$

$\Rightarrow$ the sharp rate in Krasnoselskii's iteration would be

$$\gamma(\tfrac{1}{2}) = \frac{46302245}{67108864}\sqrt{2} \sim 0.9757 \qquad \text{(smaller than } \tfrac{2}{\sqrt{\pi}} \sim 1.1284\text{)}$$

Sharp rates for KM in Hilbert spaces?

- For Krasnoselskii-Mann in Hilbert spaces with $\alpha_n \equiv \alpha$ we have

$$\sum_{n=0}^{\infty} \|x^{n+1} - x^n\|^2 < \infty \qquad \text{(Browder-Petryshin'1966)}$$

.

- Since $\|x^{n+1} - x^n\|$ is decreasing this gives a faster rate

$$\|x^n - Tx^n\| = o(1/\sqrt{n}) \qquad \text{(Baillon-Bruck'1992)}$$

- Which is the exact rate? $\Omega(1/\sqrt{n \log n})$?

Remark: in Hilbert spaces (BB) holds with $\kappa = 1$

The residuals $r^k \triangleq x^k - Tx^k$ decrease in norm:

$$\begin{aligned} \|r^k\| &= \|(1-\alpha_k)r^{k-1} + Tx^{k-1} - Tx^k\| \\ &\leq (1-\alpha_k)\|r^{k-1}\| + \|x^{k-1} - x^k\| \;=\; \|r^{k-1}\|. \end{aligned}$$

Let $a^k \triangleq x^k - x^*$ and $b^k \triangleq Tx^k - Tx^*$ with $x^* \in \mathrm{Fix}\,(T)$, so that $r^k = a^k - b^k$ and $a^{k+1} = (1-\alpha_{k+1})a^k + \alpha_{k+1}b^k$. Using the parallelogram identity

$$\|(1-\alpha)a + \alpha b\|^2 = (1-\alpha)\|a\|^2 + \alpha\|b\|^2 - \alpha(1-\alpha)\|a - b\|^2$$

and noting that $\|b^k\| \leq \|a^k\|$, we get

$$\|a^{k+1}\|^2 \leq \|a^k\|^2 - \alpha_{k+1}(1-\alpha_{k+1})\|r^k\|^2.$$

After telescoping we obtain

$$\textstyle\sum_{k=0}^n \alpha_{k+1}(1-\alpha_{k+1})\|r^k\|^2 \leq \|a^0\|^2 \leq \mathrm{diam}(C)^2$$

and the conclusion follows from the monotonicity of $\|r^k\|$. □