Some Sparse Optimization Problems and How to Solve Them

Stephen Wright

University of Wisconsin-Madison

May 2013

Two Topics

I. Identification of low-dimension subspace from incomplete data. (+ Laura Balzano — Michigan)

II. Packing ellipsoids with overlap. (+ Caroline Uhler — IST Austria)

I. Identifying Subspaces from Partial Observations

Often we observe a certain phenomenon on a high-dimensional ambient space, but the phenomenon lies on a low-dimension subspace. Moreover, our observations may not be complete: missing data

Can we recover the subspace of interest?

- Matrix completion, e.g. Netflix. Observe partial rows of an $m \times n$ matrix; each row lies (roughly) in a low-d subspace of $\mathbb{R}^n$.
- Background/foreground separation in video data.
- Mining of spatal sensor data (traffic, temperature) with high correlation between locations.
- Linear system identification in control, with streaming data?
- Structure from Motion: Observe a 3-d object from different camera angles, noting the location of reference points. Some points are occluded from some angles.

Structure from Motion

(Kennedy, Balzano, Taylor, Wright, 2013)

Subspace Identification: Formalities

- Seek subspace $S \subset \mathbb{R}^n$ of known dimension $d \ll n$.
- Know certain components $\Omega_t \subset \{1, 2, \ldots, n\}$ of vectors $v_t \in \mathcal{S}$, $t = 1, 2, \ldots$ — the subvector $[v_t]_{\Omega_t}$.
- Assume that $\mathcal{S}$ is incoherent w.r.t. the coordinate directions.

Assume that

- $v_t = \bar{U} s_t$, where range$(\bar{U}) = \mathcal{S}$, and $\bar{U}$ is $n \times d$ orthonormal, and the components of $s_t \in \mathbb{R}^d$ are i.i.d. normal with mean 0.
- Sample set Ω_t is independent for each t with $|\Omega_t| \geq q$, for some q between d and n.
- Observation subvectors $[v_t]_{\Omega_t}$ contain no noise.

Full-data case $\Omega_t \equiv \{1, 2, \ldots, n\}$ gives the solution after d steps — but the algorithm still yields an interesting result.

Sampled Data: An Online / Incremental Algorithm

Balzano (2012)

GROUSE (Grassmannian Rank-One Update Subspace Estimation).

- Process the v_t sequentially.
- Maintain an estimate U_t (orthonormal $n \times d$) of subspace basis $\bar{U}$.
- Simple update formula $U_t \to U_{t+1}$, based on $(v_t)_{\Omega_t}$.

Note:

- Setup is similar to incremental and stochastic gradient methods.
- Rank-one update formula for U_t is akin to updates in quasi-Newton Hessian and Jacobian approximations in optimization.
- Projection, so that all iterates U_t are $n \times d$ orthonormal.

One GROUSE Step

Given current estimate U_t and partial data vector $[v_t]_{\Omega_t}$, where $v_t = \bar{U} s_t$:

$$
\begin{aligned}
w_t &:= \arg\min_w \|[U_t w - v_t]_{\Omega_t}\|_2^2; \\
p_t &:= U_t w_t; \\
[r_t]_{\Omega_t} &:= [v_t - U_t w_t]_{\Omega_t}; \quad [r_t]_{\Omega_t^c} := 0; \\
\sigma_t &:= \|r_t\| \|p_t\|; \\
\text{Choose } \eta_t &> 0;
\end{aligned}
$$

$$
U_{t+1} := U_t + \left[(\cos \sigma_t \eta_t - 1) \frac{p_t}{\|p_t\|} + \sin \sigma_t \eta_t \frac{r_t}{\|r_t\|} \right] \frac{w_t^T}{\|w_t\|};
$$

We focus on the (locally acceptable) choice

$$
\eta_t = \frac{1}{\sigma_t} \arcsin \frac{\|r_t\|}{\|p_t\|}, \quad \text{which yields } \sigma_t \eta_t = \arcsin \frac{\|r_t\|}{\|p_t\|} \approx \frac{\|r_t\|}{\|p_t\|}.
$$

GROUSE Comments

With the particular step above, and assuming $\|r_t\| \ll \|p_t\|$, have

$$[U_{t+1}w_t]_{\Omega_t} \approx [p_t + r_t]_{\Omega_t} = [v_t]_{\Omega_t},$$
$$[U_{t+1}w_t]_{\Omega_t^c} \approx [p_t + r_t]_{\Omega_t^c} = [U_t w_t]_{\Omega_t^c}.$$

Thus

- On sample set Ω_t, $U_{t+1}w_t$ matches obervations in v_t;
- On other elements, the components of $U_{t+1}w_t$ and $U_t w_t$ are similar.
- $U_{t+1}z = U_t z$ for any z with $w_t^T z = 0$.

The GROUSE update is essentially a projection of a step along the search direction $r_t w_t^T$, which is a negative gradient of the inconsistency measure

$$\mathcal{E}(U_t) := \min_{w_t} \|[U_t]_{\Omega_t} w_t - [v_t]_{\Omega_t}\|_2^2.$$

The GROUSE update makes the **minimal adjustment required to match the latest observations**, while retaining a certain desired structure — orthonormality, in this case.

GROUSE Local Convergence Questions

- How to measure discrepancy between current estimate $R(U_t)$ and $\mathcal{S}$?
- Convergence behavior is obviously random, but what can we say about expected rate? Linear? If so, how fast?
- How many components q of v_t are needed at each step?

For the first question, can use *angles between subspaces* $\phi_{t,i}$, $i = 1, 2, \ldots, d$.

$$\cos \phi_{t,i} = \sigma_i(U_t^T \bar{U}),$$

where $\sigma_i(\cdot)$ denotes the ith singular value. Define

$$\epsilon_t := \sum_{i=1}^{d} \sin^2 \phi_{t,i} = d - \sum_{i=1}^{d} \sigma_i(U_t^T \bar{U})^2 = d - \|U_t^T \bar{U}\|_F^2.$$

We seek a bound for $E[\epsilon_{t+1}|\epsilon_t]$, where the expectation is taken over the random vector s_t for which $v_t = \bar{U} s_t$.

Full-Data Case ($q = n$)

Full-data case vastly simpler to analyze than the general case. Define

- $\theta_t := \arccos(\|p_t\|/\|v_t\|)$ is the angle between range(U_t) and $\mathcal{S}$ that is revealed by the update vector v_t;
- Define $A_t := U_t^T \bar{U}$, $d \times d$. We have $\epsilon_t = d - \|A_t\|_F^2$.

Lemma

$$\epsilon_t - \epsilon_{t+1} = \frac{\sin(\sigma_t \eta_t)\sin(2\theta_t - \sigma_t \eta_t)}{\sin^2 \theta_t} \left(1 - \frac{s_t^T A_t^T A_t A_t^T A_t s_t}{s_t^T A_t^T A_t s_t}\right),$$

The right-hand side is nonnegative for $\sigma_t \eta_t \in (0, 2\theta_t)$, and zero if $v_t \in R(U_t) = \mathcal{S}_t$ or $v_t \perp \mathcal{S}_t$.

The favored choice of η_t (defined above) yields $\sigma_t \eta_t = \theta_t$, thus:

$$\epsilon_t - \epsilon_{t+1} = 1 - \frac{s_t^T A_t^T A_t A_t^T A_t s_t}{s_t^T A_t^T A_t s_t}.$$

Full-Data Result

Need to calculate an expected value of the bound, over the random vector s_t. Needs some work, but we end up with:

Theorem

Suppose that $\epsilon_t \leq \bar{\epsilon}$ for some $\bar{\epsilon} \in (0, 1/3)$. Then

$$E\left[\epsilon_{t+1} \,|\, \epsilon_t\right] \leq \left(1 - \left(\frac{1 - 3\bar{\epsilon}}{1 - \bar{\epsilon}}\right)\frac{1}{d}\right)\epsilon_t.$$

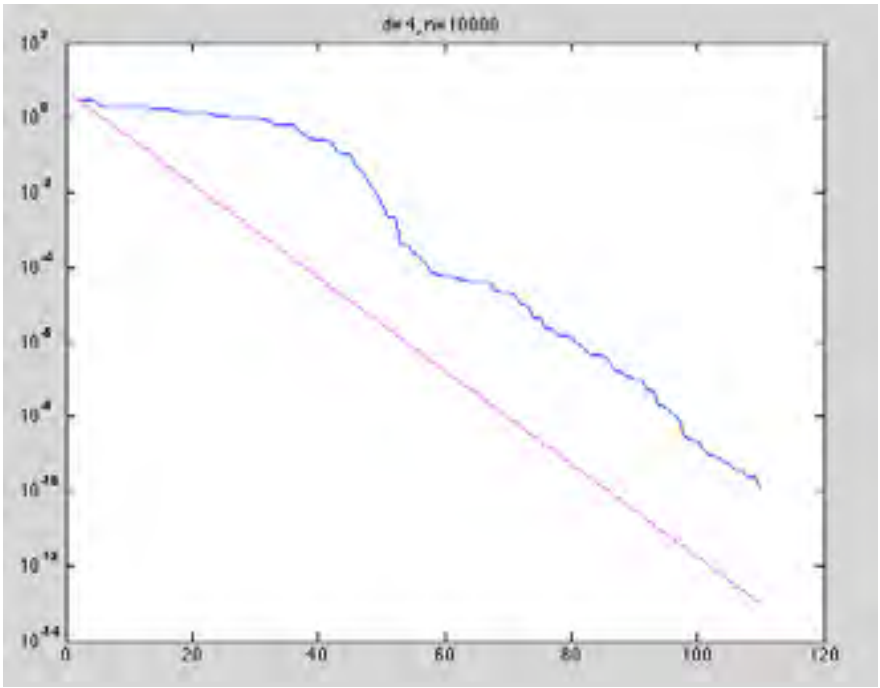
Linear convergence rate is asymptotically $1 - 1/d$.

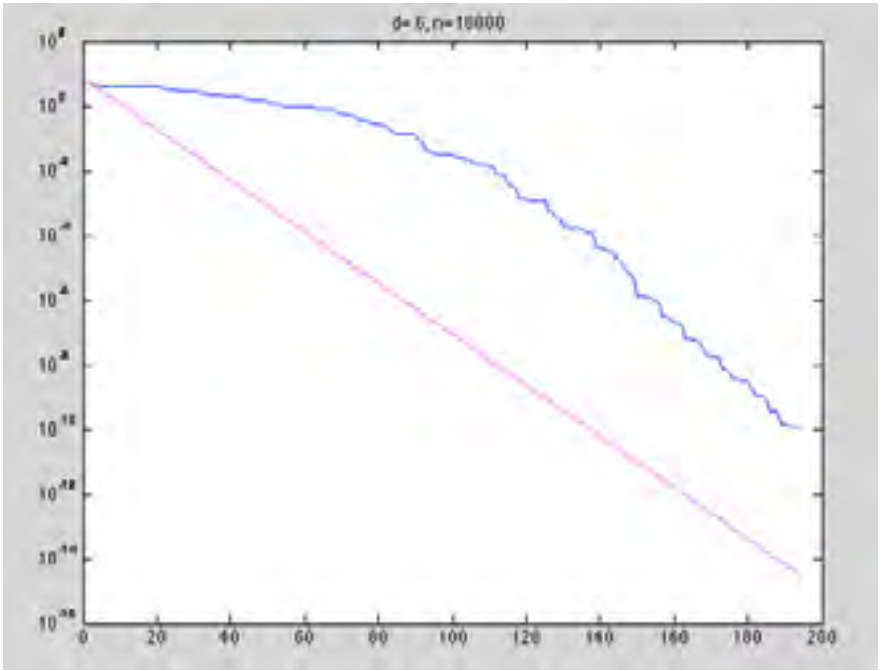
- For $d = 1$, get near-convergence in one step (thankfully!)
- Generally, in d steps (which is number of steps to get the exact solution using SVD), improvement factor is

$$(1 - 1/d)^d < \frac{1}{e}.$$

Computations confirm: Slow early, then linear with rate $(1 - 1/d)$.

ϵ_t vs expected $(1 - 1/d)$ rate (for various d)





General Case: Preliminaries

Assume a regime in which ϵ_t is small.

Define coherence of $\mathcal{S}$ (w.r.t. coordinate directions) by

$$\bar{\mu} := \frac{n}{d} \max_{i=1,2,\ldots,n} \|P_{\mathcal{S}} e_i\|_2^2.$$

It's in range $[1, n/d]$, nearer the bottom if "incoherent."

Add a safeguard to GROUSE: Take the step only if

$$\sigma_i([U_t]_{\Omega_t}^T [U_t]_{\Omega_t}) \in \left[.5\frac{|\Omega_t|}{n}, 1.5\frac{|\Omega_t|}{n}\right], \quad i = 1, 2, \ldots, d,$$

i.e. the sample is big enough to capture accurately the expression of v_t in terms of the columns of U_t. Can show that this will happen w.p. $\geq .9$ if

$$|\Omega_t| \geq q \geq C_1 (\log n)^2 d\bar{\mu} \log(20d), \quad C_1 \geq \frac{64}{3}.$$

The Result

Require conditions on q and the fudge factor C_1:

$$q \geq C_1 (\log n)^2 d\bar{\mu} \log(20d), \quad C_1 \geq \frac{64}{3};$$

Also need C_1 large enough that the coherence in the residual between v_t and current subspace estimate U_t satisfies a certain (reasonable) bound w.p. $1 - \bar{\delta}$, for some $\bar{\delta} \in (0, .6)$. Then for

$$\epsilon_t \leq (8 \times 10^{-6})(.6 - \bar{\delta})^2 \frac{q^3}{n^3 d^2},$$
$$\epsilon_t \leq \frac{1}{16} \frac{d}{n} \bar{\mu},$$

we have

$$E[\epsilon_{t+1} \,|\, \epsilon_t] \leq \left(1 - (.16)(.6 - \bar{\delta}) \frac{q}{nd}\right) \epsilon_t.$$

The Result: Comments and Steps

The decrease constant it not too far from that observed in practice; we see a factor of about

$$1 - X\frac{q}{nd}$$

where X is not too much less than 1.

The threshold condition on ϵ_t is quite pessimistic, however. Linear convergence behavior is seen at much higher values of ϵ_t.

18 pages (SIAM format) of highly technical analysis, involving:

- High-probability estimates of residual $\|r_t\|$;
- Noncommutative Bernstein inequality;
- Deterministic bound on ϵ_{t+1} in terms of ϵ_t and $\|r_t\|^2/\|p_t\|^2$;
- Incoherence assumption on the *error* identified by most samples;
- An expectation argument like the ones used in the full-data case.

Computations for GROUSE with Sampling

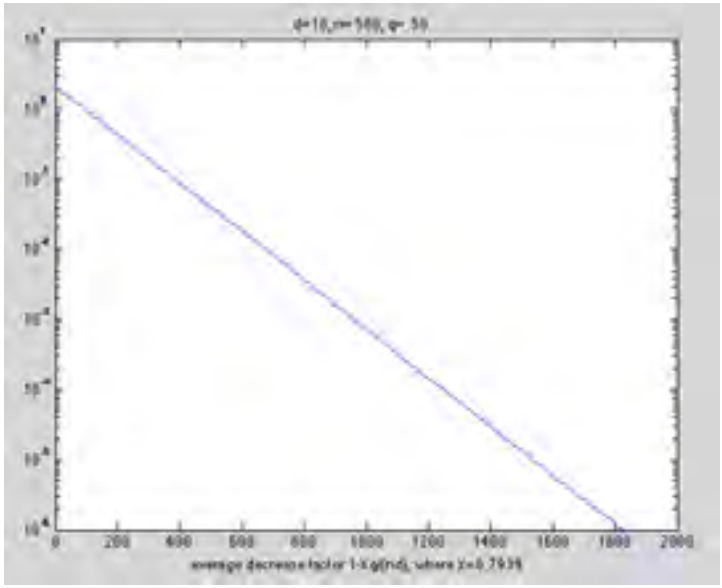
- Choose U_0 so that ϵ_0 is between 1 and 4.
- Stop when $\epsilon_t \leq 10^{-6}$.
- Calculate average convergence rate: value X such that

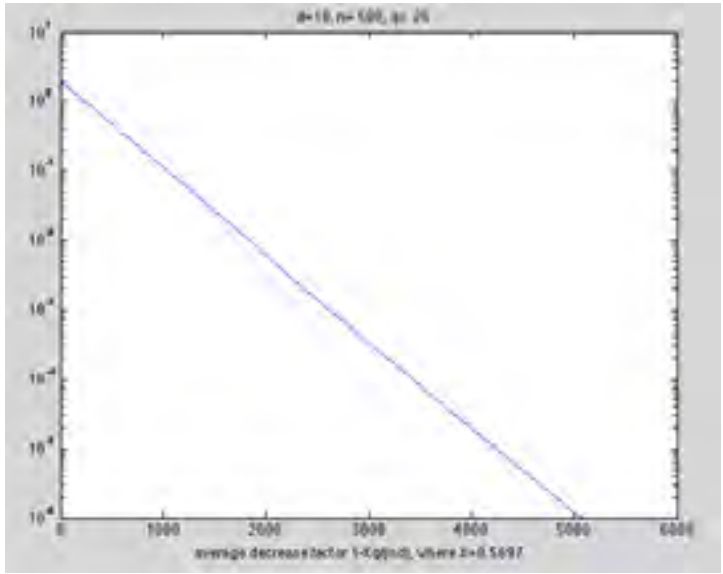
$$\epsilon_N \approx \epsilon_0 \left(1 - X\frac{q}{nd}\right)^N.$$

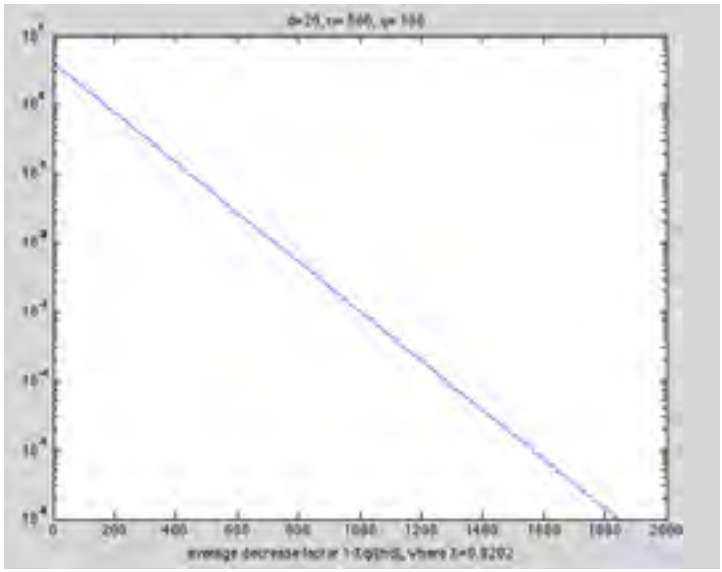
We find that X is not too much less than 1!

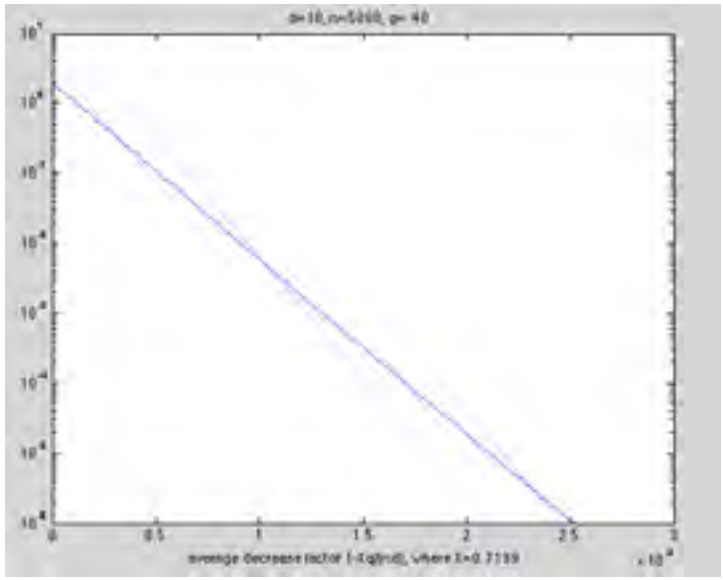
n	d	q	X
500	10	50	.79
500	10	25	.57
500	20	100	.82
5000	10	40	.72

Computations: Straight Downhill









iSVD (Incremental SVD)

GROUSE is closely related to the following incremental SVD approach.

Given U_t and $[v_t]_{\Omega_t}$:

- Compute w_t as in GROUSE:

$$w_t := \arg\min_w \|[U_t w - v_t]_{\Omega_t}\|_2^2.$$

- Use w_t to impute the unknown elements $(v_t)_{\Omega_t^c}$, and fill out v_t with these estimates:

$$\tilde{v}_t := \begin{bmatrix} [v_t]_{\Omega_t} \\ [U_t]_{\Omega_t^c} w_t \end{bmatrix}.$$

- Append $\tilde{v}_t$ to U_t and take the SVD of the resulting $n \times (d+1)$ matrix $[U_t \ : \ \tilde{v}_t]$;
- Define U_{t+1} to be the leading d singular vectors. *(Discard the singular vector that corresponds to the smallest singular value of the augmented matrix.)*

Relating iSVD and GROUSE

Theorem

Suppose we have the same U_t and $[v_t]_{\Omega_t}$ at the t-th iterations of iSVD and GROUSE. Then there exists $\eta_t > 0$ in GROUSE such that the next iterates U_{t+1} of both algorithms are identical, to within an orthogonal transformation.

The choice of η_t (details below) is *not* the same as the "optimal" choice in GROUSE, but it works fairly well in practice.

$$\lambda = \frac{1}{2}\left[(\|w_t\|^2 + \|r_t\|^2 + 1) + \sqrt{(\|w_t\|^2 + \|r_t\|^2 + 1)^2 - 4\|r_t\|^2}\right]$$

$$\beta = \frac{\|r_t\|^2\|w_t\|^2}{\|r_t\|^2\|w_t\|^2 + (\lambda - \|r_t\|^2)^2}$$

$$\eta_t = \frac{1}{\sigma_t}\arcsin\beta.$$

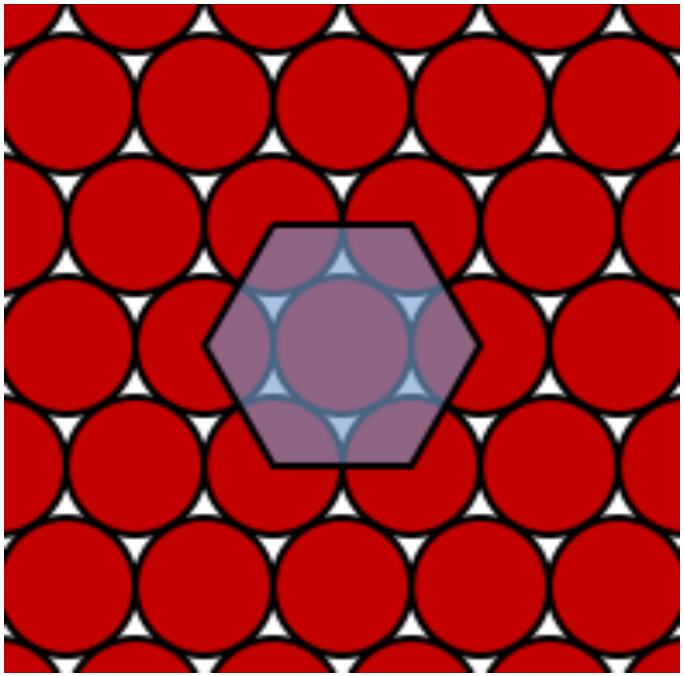
II. Packing Circles and Ellipses (and Chromosomes)

- Classical Results in Circle Packing
- Packing Circles with Minimal Overlap
  - Formulation
  - Algorithm
  - Results
- Packing Ellipsoids with Minimal Overlap
  - Formulation
  - Algorithm
  - Results
- Chromosome Arrangement.
  - Background
  - Investigate: Can geometry explain arrangements?

Circle Packing: Classical Questions

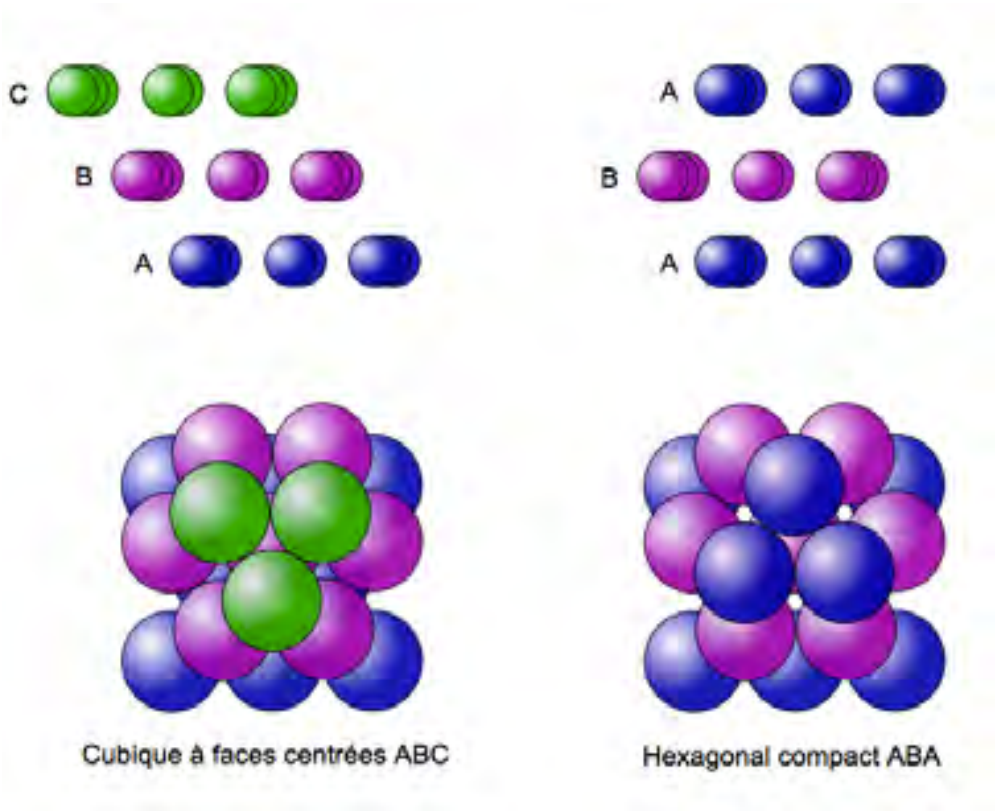
1. "Pack identical circles as densely as possible in the infinite plane."
2. "Pack identical spheres (3d and higher) as densely as possible in infinite space."
3. "Pack N identical circles in enclosing circle of minimal radius."

For Q.1, the hexagonal packing has optimal density, which is $\pi/\sqrt{12}$. (Thue, 1910; Toth, 1940). Each circle has six neighbors.



Sphere Packing

For Q.2, there are "close-packed structures" with dense layers, each layer arranged "hexagonally." Each sphere has 12 neighbors. All achieve densities of $\pi/\sqrt{18}$. *Face-centered cubic (FCC)* is one such structure.



- Gauss (1831) proved that the structures described above have the highest density ($\pi/\sqrt{18}$) among regular packings.
- Kepler (1611) conjectured that this density is the highest achievable among all packings, regular or irregular.
- Hales (1998, 2005) following Toth (1953) proved Kepler's conjecture.

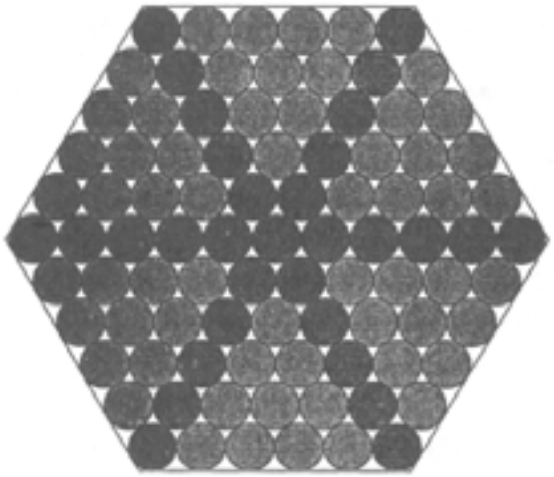
Hales' proof required computational solution of 100,000 LPs.

Requires checking of many irregular packings, some of which have higher local density than the best regular packings, but which cannot be extended infinitely.

Hales is working on a version of the proof that can be formally verified (Flyspeck project).

Q.3: Circle Packings in a Circle

Consider 91 identical circles arranged hexagonally:



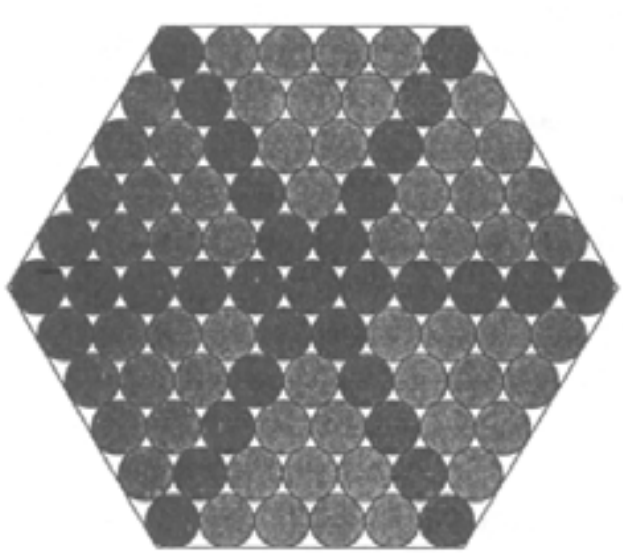
91 disks
density = 0.88434871149353
D/d = 11.154700538379

Can we rearrange these to fit them in a smaller circle, without overlap?

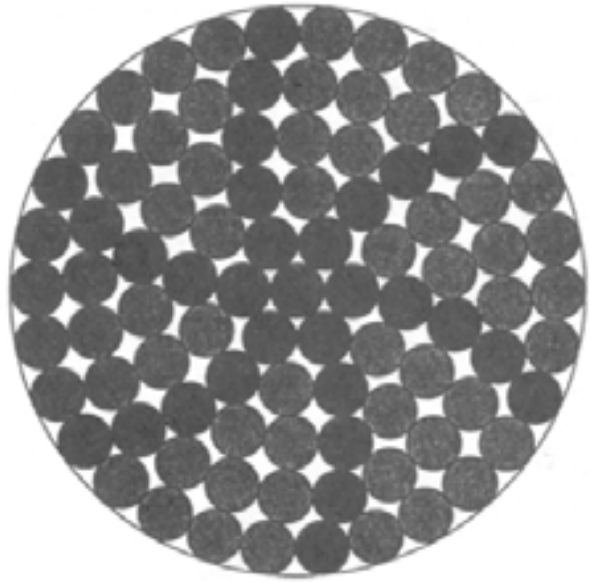
R. L. Graham, B. D. Lubachevsky et al, *Discrete Mathematics* 181 (1998), pp. 139–154.

Curved Hexagonal Packings

YES! A slight twisting of the hexagonal arrangement reduces by about 5% the radius of the enclosing circle.



91 disks
density = 0.88434871149353
D/d = 11.154700538379



91 disks
density = 0.81499829406214
D/d = 10.566772233506

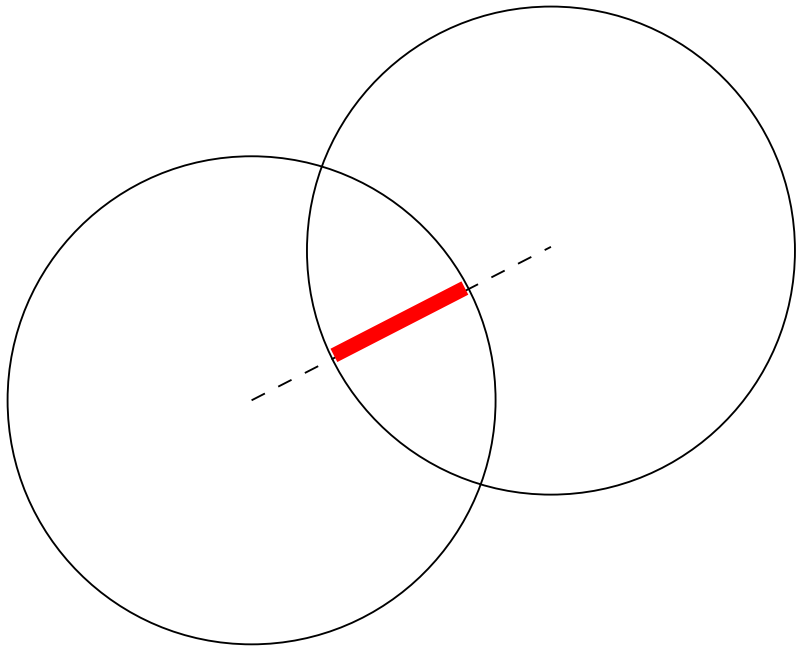
Google "circle packing Magdeburg" for best known packings up to about $N = 1100$. (Only $N = 1, 2, \ldots, 13$ and $N = 19$ are proved optimal.)

Packing Spheres with Overlap

(Pose and solve in $\mathbb{R}^n$; not restricted to $n = 2$ or $n = 3$.)

> *Given N spheres of prescribed radius r_i, $i = 1, 2, \ldots, N$, and a convex set Ω, choose centers $c_i \in \mathbb{R}^n$ so that the cirles lie within Ω and some measure of total overlap is minimized.*

Measure overlap between two spheres by *diameter of largest sphere inscribed in their intersection:* $r_i + r_j - \|c_i - c_j\|_2$.



Formulation

Given convex enclosing set Ω, can define a convex set Ω_i of allowable values for the center c_i of circle i.

Capture overlap between circles i and j by ξ_{ij}:

$$\xi_{ij} := \max(0, (r_i + r_j) - \|c_i - c_j\|_2), \qquad \xi := (\xi_{ij})_{1 \le i < j \le N}.$$

Aggregate the pairwise overlaps ξ_{ij} into a single objective H (for example, sum of squares or $\max_{i,j} \xi_{ij}$).

Optimization formulation, with unknowns c_i, $i = 1, 2, \ldots, N$ and ξ:

$$\begin{aligned} \min_{c,\xi} \quad & H(\xi) \\ \text{subject to} \quad & (r_i + r_j) - \|c_i - c_j\|_2 \le \xi_{ij} \qquad \text{for } 1 \le i < j \le N \\ & 0 \le \xi, \\ & c_i \in \Omega_i, \qquad \text{for } i = 1, 2, \ldots, N. \end{aligned}$$

Optimality Conditions

Highly nonconvex problem. Conditions for a *Clarke stationary point* are that there exist $\lambda_{ij} \in \mathbb{R}$ such that

$$0 \le g_{ij} - \lambda_{ij} \perp \xi_{ij} \ge 0 \;\text{ for some } g_{ij} \in \partial_{\xi_{ij}} H(\xi), \qquad 1 \le i < j \le N,$$

$$\sum_{j=i+1}^{N} \lambda_{ij} w_{ij} - \sum_{j=1}^{i-1} \lambda_{ji} w_{ji} \in N_{\Omega_i}(c_i), \qquad i = 1, 2, \ldots, N,$$

$$0 \le \xi_{ij} + \|c_i - c_j\| - (r_i + r_j) \perp \lambda_{ij} \ge 0, \qquad 1 \le i < j \le N,$$

where $\|w_{ij}\|_2 \le 1$, with $w_{ij} = \dfrac{c_i - c_j}{\|c_i - c_j\|_2}$ when $c_i \ne c_j$, $1 \le i < j \le N$

Here

- $N_{\Omega_i}(c_i)$ is the normal cone to Ω_i at c_i;
- ∂ denotes subdifferential.

Algorithm: Key Subproblem

Linearize the constraint defining ξ_{ij} about the current point c^-, to define the subproblem to be solved at each iteration:

$$
\begin{aligned}
P(c^-) := \min_{c,\bar{\xi}} \; & H(\bar{\xi}) \\
\text{subject to} \quad & (r_i + r_j) - z_{ij}^T (c_i - c_j) \leq \bar{\xi}_{ij}, \quad \text{for } 1 \leq i < j \leq N, \\
& 0 \leq \bar{\xi}, \\
& c_i \in \Omega_i, \quad \text{for } i = 1, \ldots, N,
\end{aligned}
$$

$$
\text{where} \quad z_{ij} := \begin{cases} (c_i^- - c_j^-)^T / \|c_i^- - c_j^-\| & \text{when } c_i^- \neq c_j^- \\ 0 & \text{otherwise.} \end{cases}
$$

Use the original objective H — no need to approximate since it's simple.

Depending on the form of H and Ω_i, $P(c^-)$ could be a linear program, quadratic program, or more general conic program.

Algorithm

Given $r_i > 0$ and constraint sets Ω_i, $i = 1, 2, \ldots, N$;
Choose $c^0 \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_N$;
for $k = 0, 1, 2, \ldots$ **do**
  Generate z_{ij} for $1 \leq i < j \leq N$;
  Solve subproblem $P(c^k)$ to obtain $(c^{k+1}, \bar{\xi}^{k+1})$;
  if $H(\bar{\xi}^{k+1}) = H(\xi^k)$ **then**
    stop and return c^k;
  end if
  Set $\xi_{ij}^{k+1} = \max(0, (r_i + r_j) - \|c_i^{k+1} - c_j^{k+1}\|)$ for $1 \leq i < j \leq N$;
end for

Convergence

- If (c^k, ξ^k) solves the subproblem $P(c^k)$ (i.e. algorithm doesn't move), then it is stationary for the main problem.
- If the current (c^k, ξ^k) is stationary for the main problem, with $c_i^k \neq c_j^k$ for $i \neq j$, then it also solves the subproblem $P(c^k)$.
- If (c^k, ξ^k) is *not* stationary for the main problem, then the subproblem predicts a strict reduction in objective: $H(\bar{\xi}^{k+1}) < H(\xi^k)$.
- Objective H improves even more than forecast: $H(\xi^{k+1}) < H(\bar{\xi}^{k+1})$. Linearization overestimates the true overlap. Thus, no need for trust region or line search.

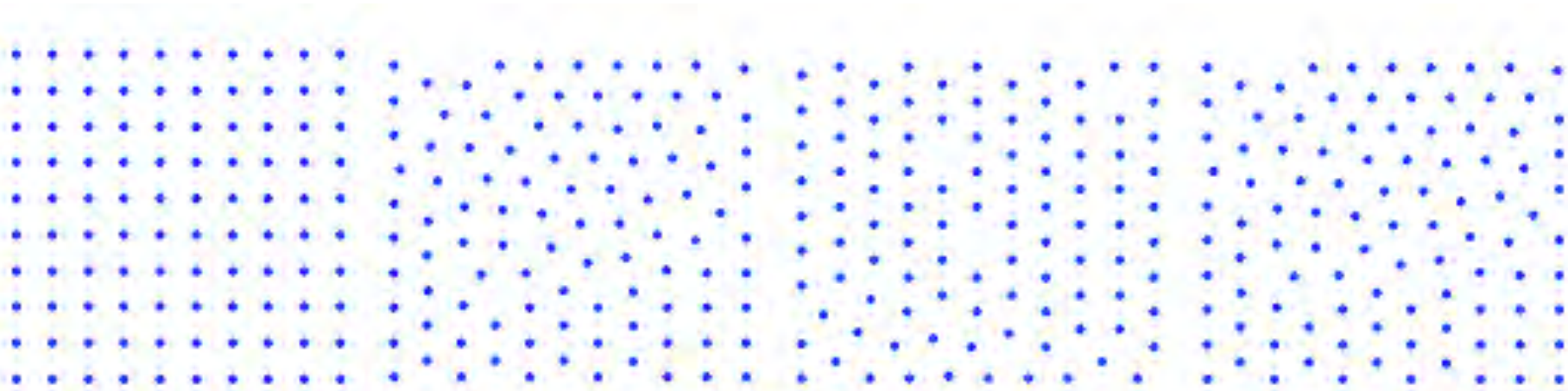
Result: **All accumulation points $\hat{c}$ of the sequence $\{c^k\}$ are either stationary or degenerate (i.e. $\hat{c}_i = \hat{c}_j$ for $i \neq j$).**

There are typically many local minima, or families of minima. Computed solution depends on starting point.

Results: Emergence of Hexagons

Packing 100 circles into a square with min-max overlap.

Many different local minima obtained. The square grid is one such, but there are better solutions in which hexagonal structure emerges in large parts of the domain.



(a) $o = .1192514295$ (b) $o = .1188906843$ (c) $o = .1181440939$ (d) $o = .1179656050$

(Hexagonal packing overlap is $\approx .1149$.)

Packing Ellipsoids: M&Ms

Improving the Density of Jammed Disordered Packings Using Ellipsoids

Aleksandar Donev,[1,4] Ibrahim Cisse,[2,5] David Sachs,[2] Evan A. Variano,[2,6] Frank H. Stillinger,[3] Robert Connelly,[7] Salvatore Torquato,[1,3,4*] P. M. Chaikin[2,4]

Packing problems, such as how densely objects can fill a volume, are among the most ancient and persistent problems in mathematics and science. For equal spheres, it has only recently been proved that the face-centered cubic lattice has the highest possible packing fraction $\varphi = \pi/\sqrt{18} \approx 0.74$. It is also well known that certain random (amorphous) jammed packings have $\varphi \approx 0.64$. Here, we show experimentally and with a new simulation algorithm that ellipsoids can randomly pack more densely—up to $\varphi = 0.68$ to 0.71 for spheroids with an aspect ratio close to that of M&M's Candies—and even approach $\varphi \approx 0.74$ for ellipsoids with other aspect ratios. We suggest that the higher density is directly related to the higher number of degrees of freedom per particle and thus the larger number of particle contacts required to mechanically stabilize the packing. We measured the number of contacts per particle $Z \approx 10$ for our spheroids, as compared to $Z \approx 6$ for spheres. Our results have implications for a broad range of scientific disciplines, including the properties of granular media and ceramics, glass formation, and discrete geometry.

The structure of liquids, crystals, and glasses is intimately related to volume fractions of ordered and disordered (random) hard-sphere packings, as are the transitions between these phases (*1*). Packing problems (*2*) are of current interest in dimensions higher than three

13 FEBRUARY 2004 VOL 303 SCIENCE www.sciencemag.org

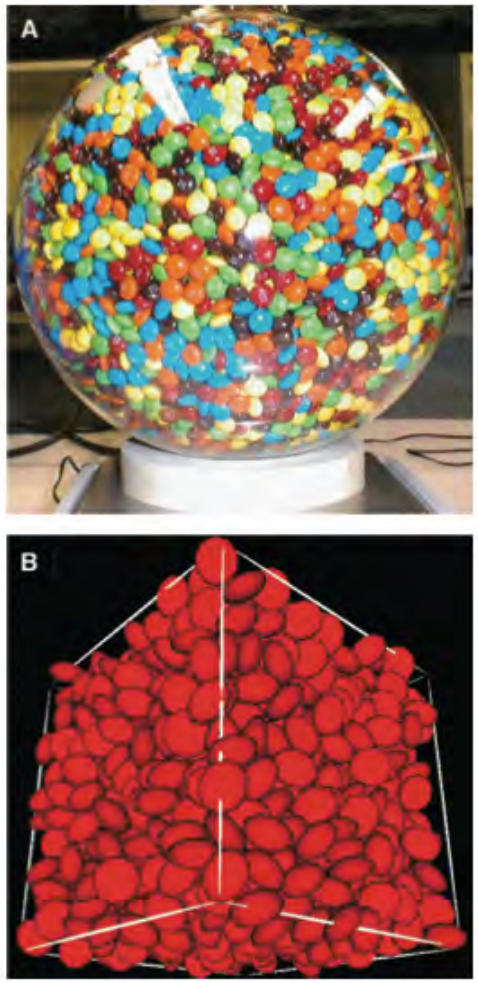


Fig. 1. **(A)** An experimental packing of the regular candies. **(B)** Computer-generated packing of 1000 oblate ellipsoids with $\alpha = 1.9^{-1}$.

Packing 3D Ellipsoids: Results (from Science, 2004)

- Optimal ordered packing for spheres has density $\pi/\sqrt{18} \approx .74$.
- "Random" spherical packings have density $\approx .64$, with each sphere touching about 6 of its neighbors (on average).
- Densities of random packings increase when spheres become ellipsoids. More contacts with neighbors are required for a "jammed" configuration.
- Among prolate and oblate ellipsoids, best packing is attained by ellipsoids with aspect ratio similar to M&Ms. Density $\approx .685$ with about 10 contacts per ellipsoid.
- Donev et al verified by measurements with actual M&Ms and a molecular-dynamics simulation. Other authors did experiments on sphere packing with ball bearings.

Our algorithm applied to uniform spheres gave packings with an average of 11.5 neighbors per sphere — close to the FCC count of 12, much higher than random packing count of 6.

Ellipsoids: S-Lemma

Given two ellipses:

$$\mathcal{E} = \{x \in \mathbb{R}^3 \mid (x-c)^T S^{-2}(x-c) \leq 1\} = \{c + Su \mid \|u\|_2 \leq 1\},$$
$$\bar{\mathcal{E}} = \{x \in \mathbb{R}^3 \mid (x-\bar{c})^T \bar{S}^{-2}(x-\bar{c}) \leq 1\} = \{\bar{c} + \bar{S}u \mid \|u\|_2 \leq 1\}.$$

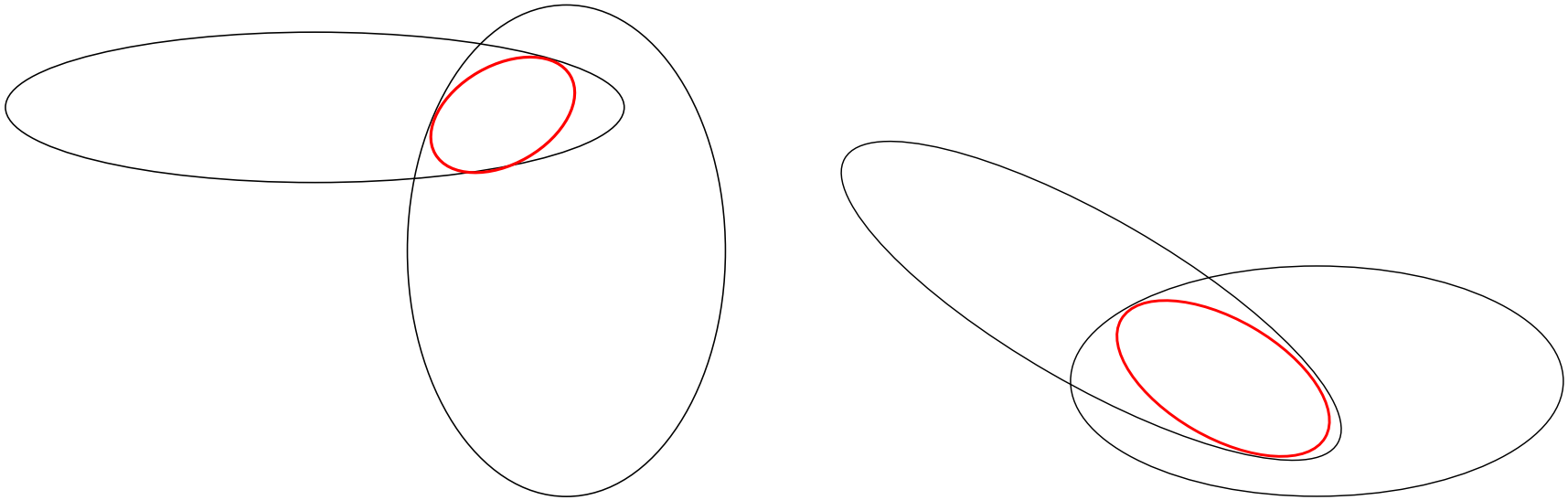
The containment condition $\bar{\mathcal{E}} \subset \mathcal{E}$ can be represented as the following linear matrix inequality (LMI) in parameters $\bar{c}$, $\bar{S}$, c, and S^2: There exists $\lambda \in \mathbb{R}$ such that

$$\begin{pmatrix} -\lambda I & 0 & \bar{S} \\ 0 & \lambda - 1 & (\bar{c}-c)^T \\ \bar{S} & \bar{c}-c & -S^2 \end{pmatrix} \preceq 0.$$

For two ellipsoids $\mathcal{E}_i$ and $\mathcal{E}_j$, with $1 \leq i < j \leq N$, denote their parameters by (c_i, S_i) and (c_j, S_j).

It's also useful to define $\Sigma_i := S_i^2$ and $\Sigma_j := S_j^2$.

Ellipsoid Overlaps



Measure the overlap between two ellipsoids as the maximal sum of principal axes of any ellipsoid inscribed in the intersection.

Ellipsoids: Measuring Overlap

Use the S-Lemma containment result above to formulate a subproblem to measure overlap, denoted by $\hat{O}(c_i, c_j, \Sigma_i, \Sigma_j)$

$$\max_{S_{ij} \succeq 0, c_{ij}, \lambda_{ij1}, \lambda_{ij2}} \quad \operatorname{trace}(S_{ij})$$

$$\text{subject to} \quad \begin{pmatrix} -\lambda_{ij1} I & 0 & S_{ij} \\ 0 & \lambda_{ij1} - 1 & (c_{ij} - c_i)^T \\ S_{ij} & c_{ij} - c_i & -\Sigma_i \end{pmatrix} \preceq 0,$$

$$\begin{pmatrix} -\lambda_{ij2} I & 0 & S_{ij} \\ 0 & \lambda_{ij2} - 1 & (c_{ij} - c_j)^T \\ S_{ij} & c_{ij} - c_j & -\Sigma_j \end{pmatrix} \preceq 0.$$

Dual Formulation of Overlap

Introduce matrices M_{ij1} and M_{ij2} defined by

$$M_{ij1} := \begin{pmatrix} R_{ij1} & r_{ij1} & P_{ij1} \\ r_{ij1}^T & p_{ij1} & q_{ij1}^T \\ P_{ij1} & q_{ij1} & Q_{ij1} \end{pmatrix}, \qquad M_{ij2} := \begin{pmatrix} R_{ij2} & r_{ij2} & P_{ij2} \\ r_{ij2}^T & p_{ij2} & q_{ij2}^T \\ P_{ij2} & q_{ij2} & Q_{ij2} \end{pmatrix},$$

Now can write the dual explicitly as follows:

$$\min_{M_{ij1} \succeq 0, M_{ij2} \succeq 0, T_{ij} \succeq 0} \quad p_{ij1} + p_{ij2} + 2q_{ij1}^T c_i + 2q_{ij2}^T c_j + \langle Q_{ij1}, \Sigma_i \rangle + \langle Q_{ij2}, \Sigma_j \rangle$$

$$\begin{aligned} \text{subject to} \quad & 0 = I + T_{ij} - 2P_{ij1} - 2P_{ij2} \\ & 0 = \operatorname{trace}(R_{ij1}) - p_{ij1} \\ & 0 = \operatorname{trace}(R_{ij2}) - p_{ij2} \\ & 0 = q_{ij1} + q_{ij2}. \end{aligned}$$

Overlap Problem: Sensitivity of Objective

- Since the dual always has a strictly feasible point, strong duality holds: the optimal primal and dual objectives are the same.
- In the dual formulation of $\hat{O}(c_i, c_j, \Sigma_i, \Sigma_j)$, parameters defining the two ellipses $c_i, c_j, \Sigma_i, \Sigma_j$ enter only into the objective, not the constraints.
- Sensitivity of $\hat{O}(c_i, c_j, \Sigma_i, \Sigma_j)$ to the parameters can be obtained from the dual optimal values, in particular q_{ij1}, q_{ij2}, Q_{ij1}, and Q_{ij2}.

Hence, when the dual solution exists, we can use it to construct a linearized model of the overlap, as a function of the positions and orientations of $\mathcal{E}_i$ and $\mathcal{E}_j$.

Packing Ellipses in an Ellipse

Using the overlap notation, formulation the problem of packing ellipses in an ellipse with min-max overlap as follows:

$$\begin{aligned}
\min_{\xi,(c_i,S_i,\Sigma_i),i=1,2,\ldots,N} \quad & \xi \\
\text{subject to} \quad & \xi \geq \hat{O}(c_i, c_j, \Sigma_i, \Sigma_j), && 1 \leq i < j \leq N, \\
& \mathcal{E}_i \subset \mathcal{E}, && i = 1, 2, \ldots, N, \\
& \Sigma_i = S_i^2, && i = 1, 2, \ldots, N, \\
& \text{semi-axes of } \mathcal{E}_i \text{ have lengths } r_{i1}, r_{i2}, r_{i3}, && i = 1, 2, \ldots, N.
\end{aligned}$$

- The scalar ξ captures the maximum overlap;
- $\mathcal{E}_i$, $i = 1, 2, \ldots, N$ denote the ellipses with specified axes (r_{i1}, r_{i2}, r_{i3}),
- $\mathcal{E}$ denotes the circumscribing ellipse.

Nonconvexity

The problem is highly nonconvex.

- Each $\hat{O}(c_i, c_j, \Sigma_i, \Sigma_j)$ is a nonconvex function of its arguments - this is intrinsic.
- Constraint $\Sigma_i = S_i^2$ is nonconvex. Can easily replace it by the following convex pair of constraints:

$$\begin{bmatrix} \Sigma_i & S_i \\ S_i & I \end{bmatrix} \succeq 0, \qquad S_i \succeq 0.$$

- Constraints on the eigenvalues of S_i are nonconvex. We replace these by convex relaxations:

$$S_i - r_{i1} I \preceq 0, \quad S_i - r_{i3} I \succeq 0, \quad \mathrm{trace}(S_i) = r_{i1} + r_{i2} + r_{i3}.$$

We formulate the inclusion condition $\mathcal{E}_i \subset \mathcal{E}$ in a convex fashion, using the S-Lemma, as above.

Successive Linearization Strategy

ELL: The min-max-overlap problem, after relaxations.

Propose a trust-region bilevel successive linearization strategy for finding local solutions of ELL. Each iteration solves one "big" top-level conic program, and many small conic programs corresponding to the pairwise dual overlaps.

- Solve the dual overlaps for each pair of nearby ellipses (i, j);
- Use dual optimal values to linearize $\xi \geq \hat{O}(c_i, c_j, \Sigma_i, \Sigma_j)$ for the pairs (i, j) with significant overlap;
- Incorporate the other formulation elements described above, to get a conic programming subproblem;
- Add a trust-region constraint on the steps;
- If the step gives a sufficient improvement in the max overlap, accept it. Otherwise, shrink the trust-region radius and try again.

Trust-Region Subproblem

$$\min_{\xi,(\lambda_i,c_i,S_i,\Sigma_i),i=1,2,\ldots,N} \quad \xi$$

$$\text{subject to} \quad \xi \geq p_{ij1} + p_{ij2} + 2q_{ij1}^T c_i + 2q_{ij2}^T c_j + \langle Q_{ij1}, \Sigma_i \rangle + \langle Q_{ij2}, \Sigma_j \rangle, \qquad \text{for } (i,j) \in \mathcal{I},$$

$$\begin{bmatrix} -\lambda_i I & 0 & S_i \\ 0 & \lambda_i - 1 & (c_i - c)^T \\ S_i & c_i - c & -\Sigma \end{bmatrix} \preceq 0, \quad i = 1, 2, \ldots, N,$$

$$\begin{bmatrix} \Sigma_i & S_i \\ S_i & I \end{bmatrix} \succeq 0, \qquad i = 1, 2, \ldots, N,$$

$$S_i - r_{i1} I \preceq 0, \quad S_i - r_{i3} I \succeq 0, \qquad i = 1, 2, \ldots, N,$$

$$\text{trace}(S_i) = r_{i1} + r_{i2} + r_{i3}, \qquad i = 1, 2, \ldots, N,$$

$$\|c_i - c_i^-\|_2^2 \leq \Delta_c^2, \qquad i = 1, 2, \ldots, N,$$

$$\|S_i - S_i^-\| \leq \Delta_S, \qquad i = 1, 2, \ldots, N,$$

$$|\lambda_i - \lambda_i^-| \leq \Delta_\lambda, \qquad i = 1, 2, \ldots, N.$$

Framework for Analysis

We simplify and generalize the problem for purposes of convergence analysis. Each pairwise overlap problem is stated as an objective-parametrized SDP:

$$P(I, C): \qquad t_I^*(C) := \min_{M_I} \langle C, M_I \rangle$$
$$\text{s.t. } \langle A_{I,i}, M_I \rangle = b_{I,i}, \quad i = 1, 2, \ldots, p_I, \quad M_I \succeq 0,$$

which is assumed to satisfy a Slater condition. Each index I represents a single pair of ellipsoids.

The top-level problem is

$$\min_{C \in \Omega} t^*(C) := \max_{I=1,2,\ldots,m} t_I^*(C),$$

where Ω is a closed convex set. (Actually, Ω is the intersection of a closed convex set with nonempty interior and a hyperplane.)

Convergence

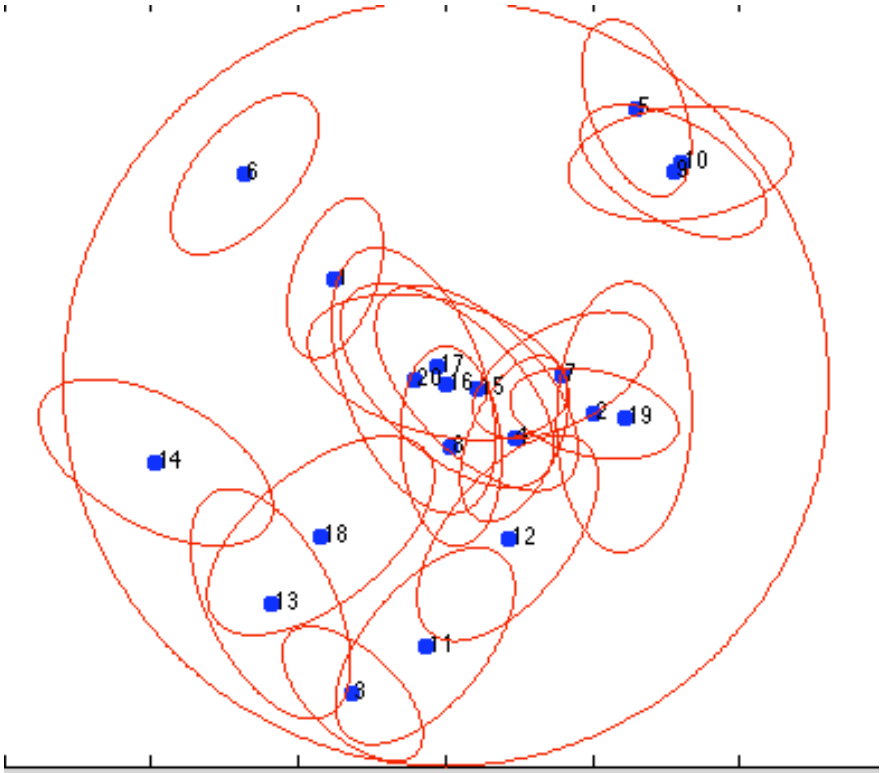
Convergence analysis uses

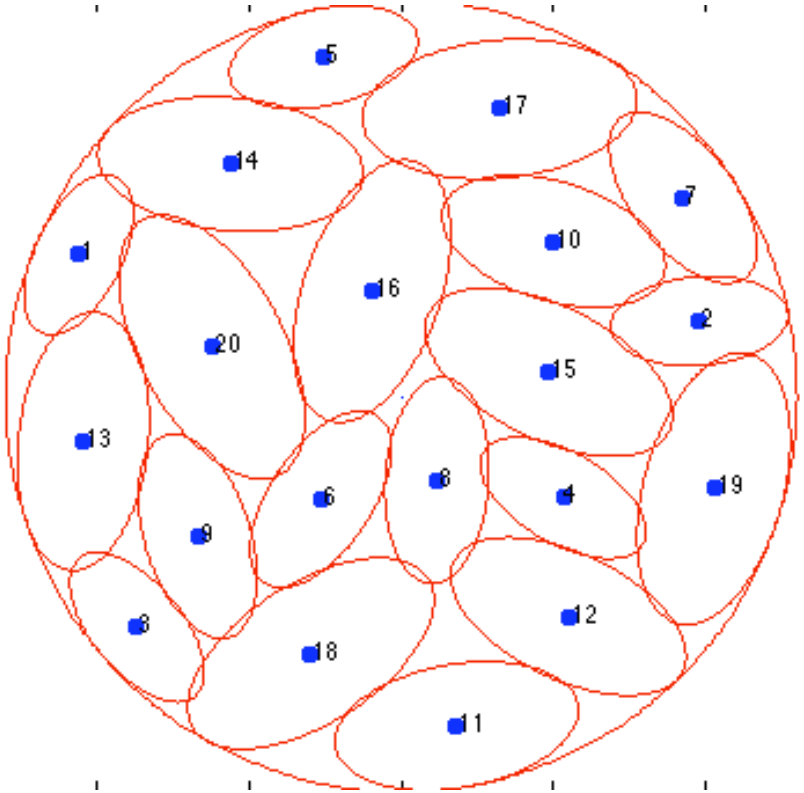
- both convex and nonconvex analysis, particularly Clarke's (1983) concepts of generalized gradients and stationary;
- SDP duality and optimality conditions;
- trust-region machinery.

Result: Except for finitely-terminating degenerate cases, **convergent subsequences of the algorithm are Clarke-stationary or no-overlap points of ELL.**

Implemented in Matlab and CVX (Grant and Boyd, `cvxr.com`)

Results: 20 Ellipses (40 iterations)

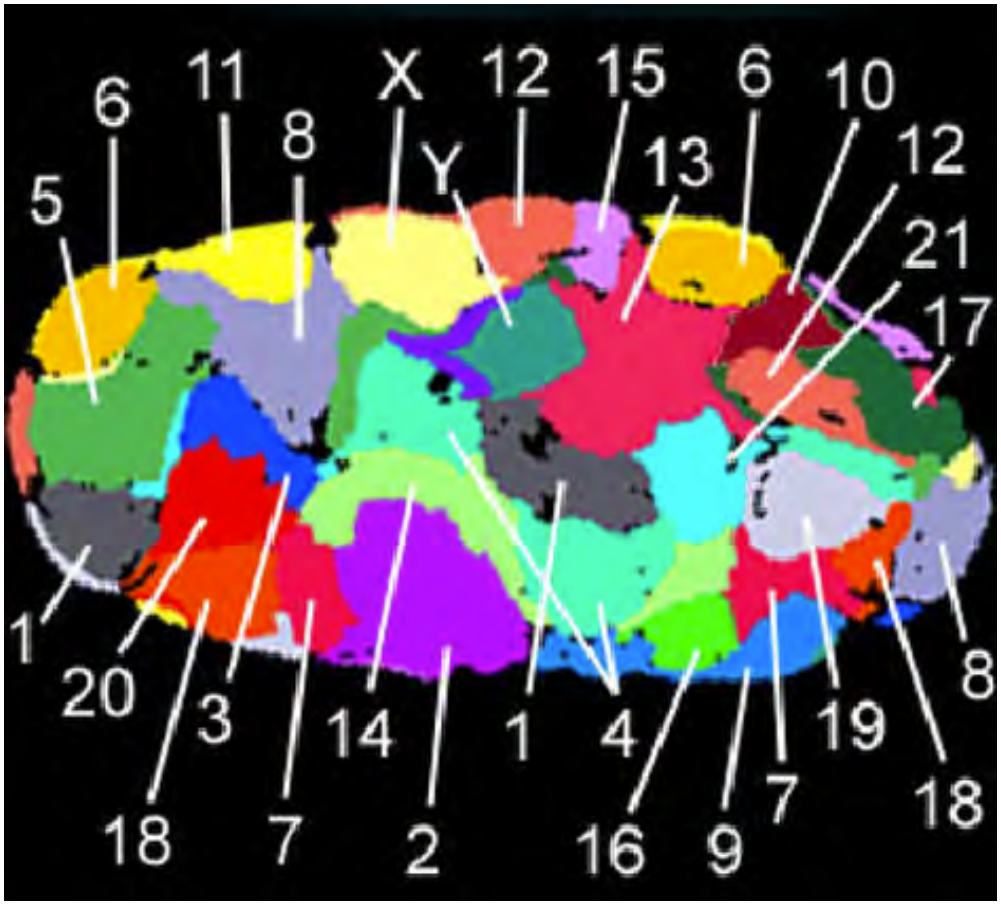




Chromosome Packing

Study interphase arrangement of chromosome territories in cell nuclei.

Chromatine fibers in DNA are not jumbled together randomly. Rather, the fibers corresponding to a single chromosome tend to associate in a particular region of the nucleus, forming a chromosome territory (CT).



Packing of CT affects cell biology:

- Chromosome locked in the interior may not be expressed.
- Overlap of CTs allows for co-regulation of genes.
- Interchromatime compartments (internal DNA-free channels) allow access to CTs in the interior.

Different cell nuclei have different sizes and shapes, forcing different packings of CTs.

Arrangement of CTs is believed to change during cell division and differentiation.

Cremer and Cremer (2010): *"The search for nonrandom chromatin assemblies, the mechanisms responsible for their formation, and their functional implications is one of the major goals of nuclear architecture research. This search is still in its beginning."*

Locational preferences for CT have been noted experimentally:

- Radial Preference;
  - In spherical nuclei (e.g. lymphocytes), gene-dense chromosomes tend to be in the interior
  - In ellipsoidal nuclei (e.g. fibroblasts), small chromosomes tend to be in the interior.
- Neighbor Preference:
  - proximity to co-regulated genes.
- Separated homologs:
  - Homologous chromosomes tend to be separated further than heterologs.

"Tethering" effects, and adhesion to nuclear walls, also may help determine CT arrangement.

Goal and Method

Goal: Determine whether the locational preference can be explained by purely geometrical, packing considerations.

Method: Use our algorithm to identify packings that are "locally optimal" in minimizing maximum overlap. Plot CT size vs distance from center of nucleus.

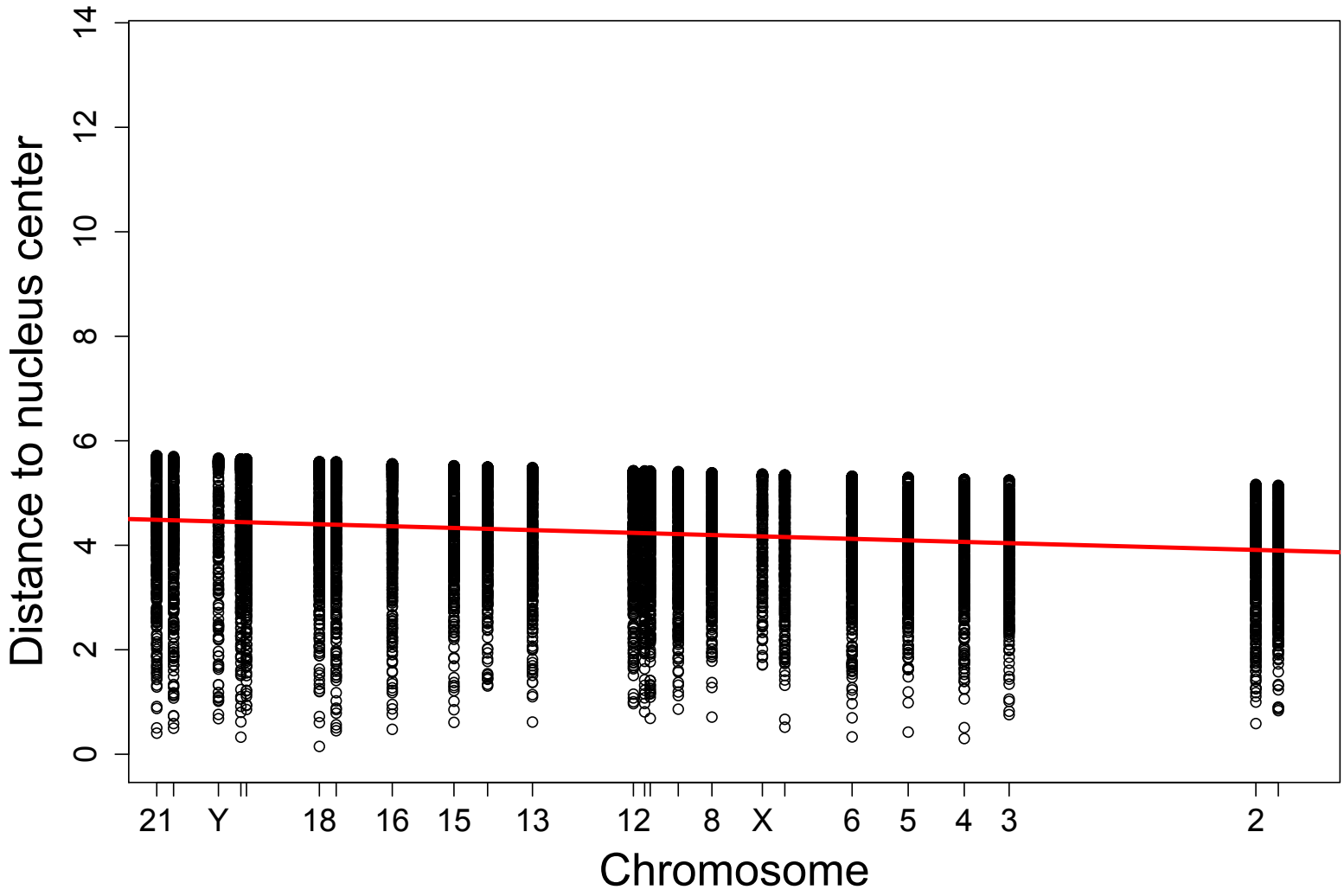
Use ellipsoids to model CTs. An approximation, but much more realistic than the circles used in an earlier phase of the study.

Setup

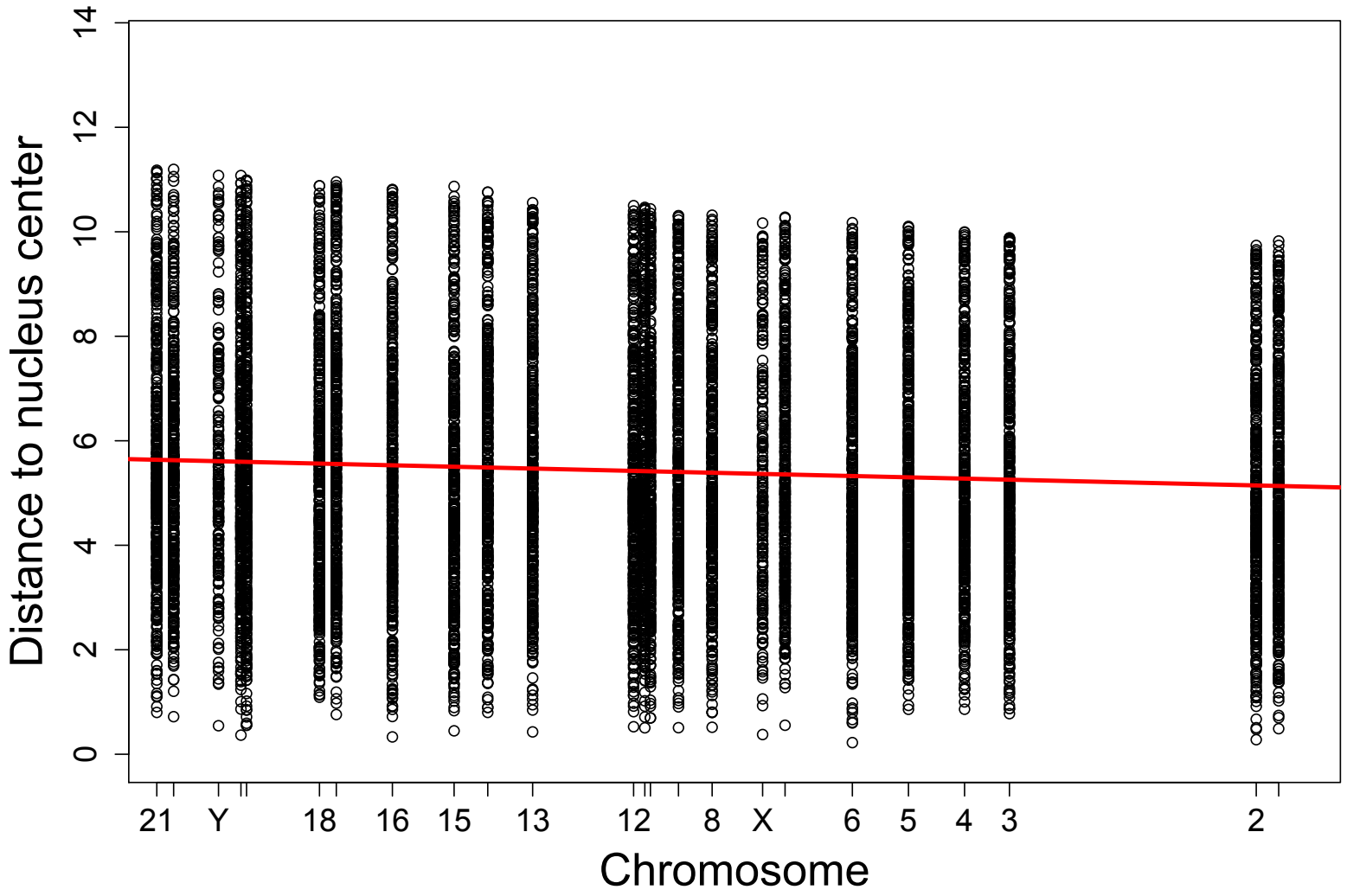
- Three different nucleus sizes: 500, 1000, 1600 $\mu\mathrm{m}^3$.
- Two nucleus shapes: spherical, and ellipsoidal with axis ratio approximately $1:2:4$
- Volumes of CTs based on known number of base pairs in each, and average density.
- Shapes of CTs based on observations of mouse chromosomes, approximate axis ratios $1:2.9:4.4$.
- Generated 50 problems for each parameter combination, by tweaking axis ratios and CT volumes.
- Plot distance of CT to nucleus center vs volume of CT.

CT	1	2	3	4	5	6	7	8
volume	37.05	36.45	29.85	28.65	27.15	25.65	23.85	21.90
CT	9	10	11	12	13	14	15	16
volume	21.00	20.25	20.10	19.80	17.10	15.90	15.00	13.35
CT	17	18	19	20	21	22	X	Y
volume	11.85	11.40	9.45	9.30	7.05	7.50	23.25	8.70

Medium Spherical Nucleus (No Homolog Separation)



Medium Ellipsoidal Nucleus (No Homolog Separation)



Observations, Modified Formulation

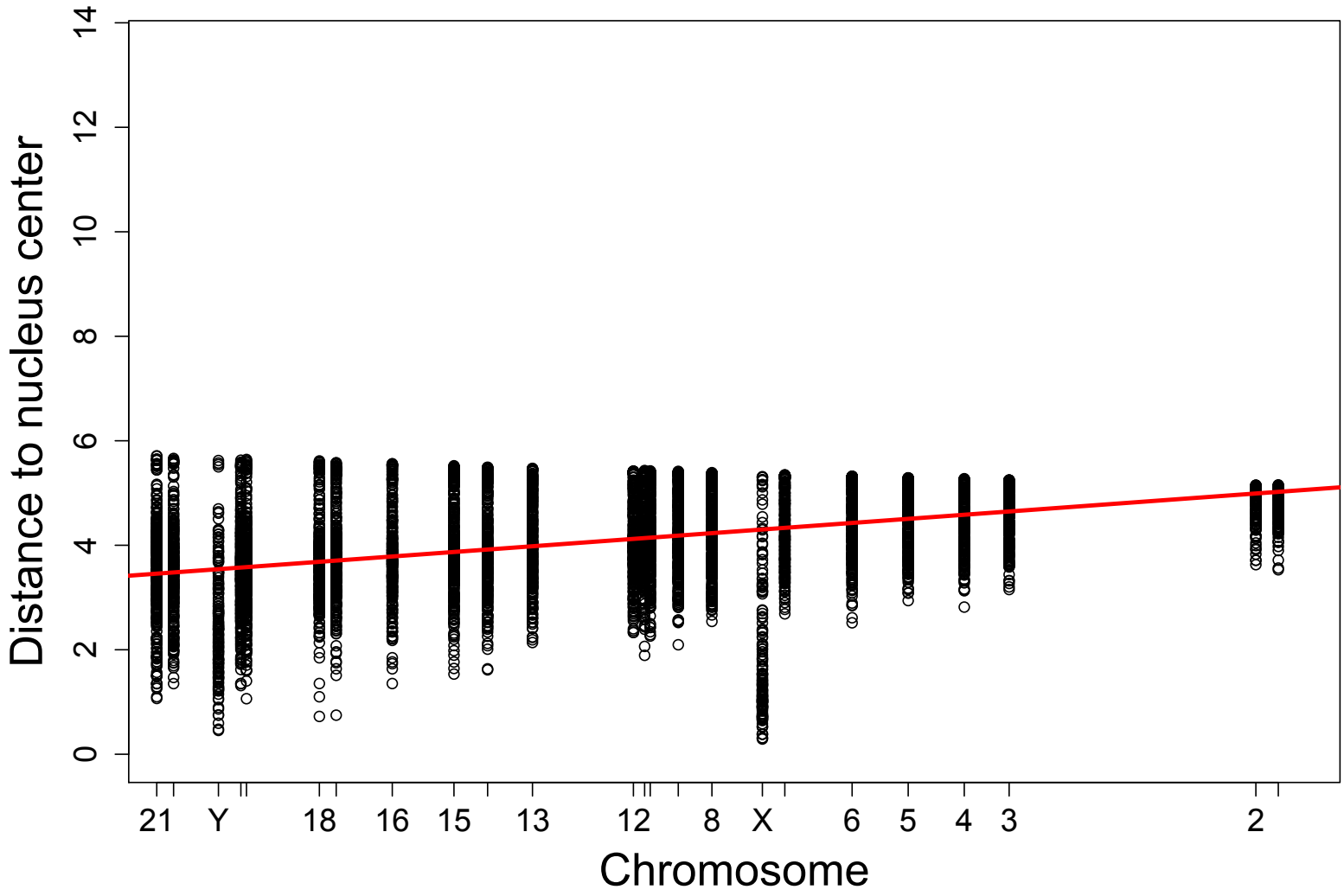
We see a slight radial preference for packing larger CTs toward the center — opposite to biological observations so far. Suggests that the observed radial preference cannot be explained simply by min-overlap packing.

Change formulation by adding a **penalty for overlap of homologs**. (Affects 22 constraints, one for each of the homologous pairs in human DNA.)

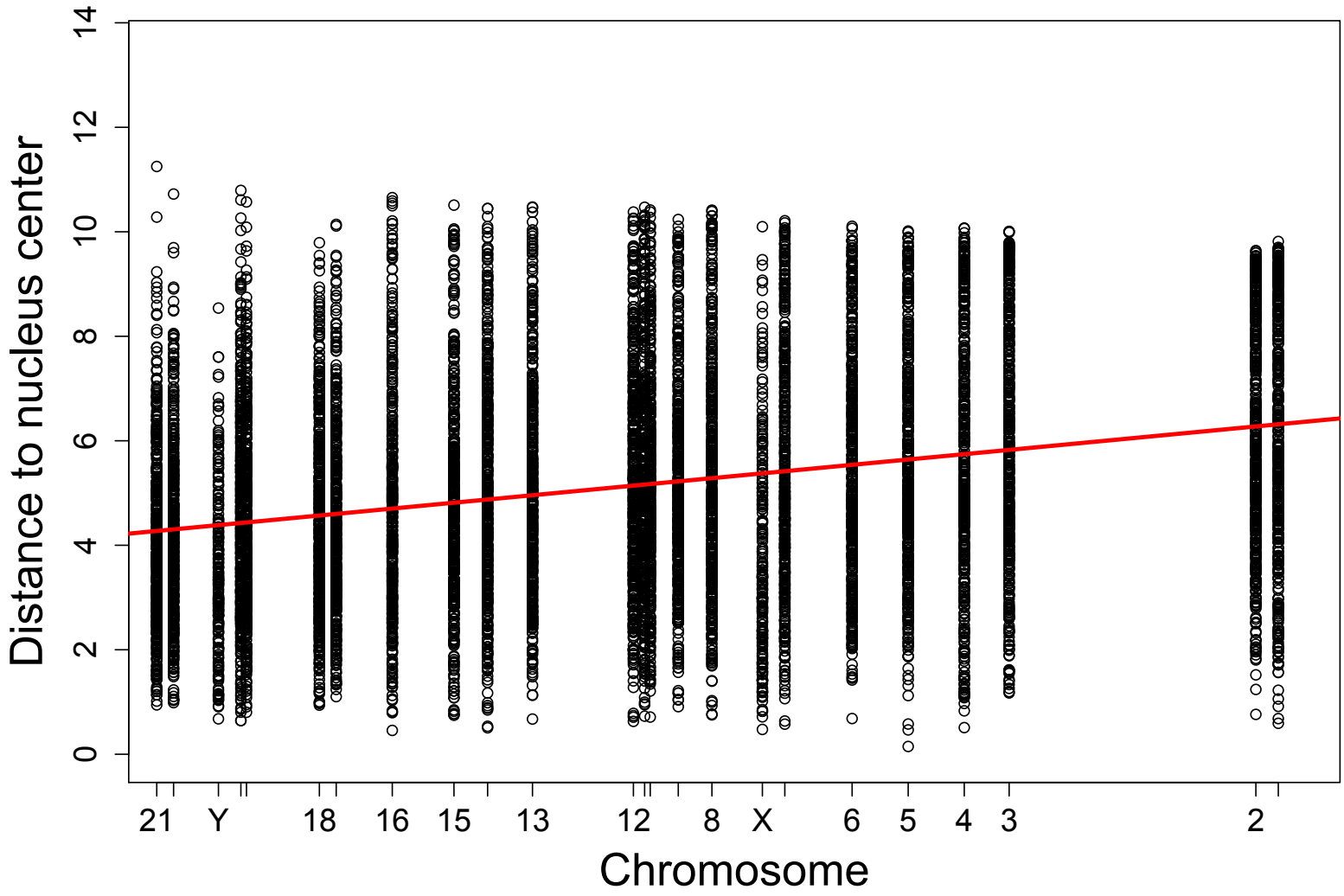
Possible bio explanations for separated homologs:

- avoiding DNA recombination between homologs;
- avoiding co-regulation of genes.

Medium Spherical Nucleus, Penalized Homolog Overlaps



Medium Ellipsoidal Nucleus, Penalized Homolog Overlaps



Conclusions

Geometrical considerations (minimizing overlap) plus the tendency for homolog separation are enough to explain the observed tendency for larger CTs to be further from the center.

Results are preliminary — there's much more to learn from the bio side, and much more to try from the formulation and algorithmic side.

(Teaming with C. Lanctôt (Prague) for experiments with C. elegans.)

We believe that algorithms and experiments like ours will help in understanding CT arrangement, in particular, its dependence on basic biological and geometrical principles.

Paper: C. Uhler and S. J. Wright, "Packing Ellipsoids with Overlap," *SIAM Review*, to appear.
`www.optimization-online.org/DB_HTML/2012/04/3418.html`