# Image Processing, Internet-of-Things, and Inverse Problems: Blind Deconvolution Meets Blind Demixing

Thomas Strohmer
Department of Mathematics
University of California, Davis

SIAM Conference on Imaging Science
Albuquerque May 24, 2016

# Acknowledgements

Research in collaboration with:

Xiaodong Li (UC Davis)
Shuyang Ling (UC Davis)
Ke Wei (UC Davis)

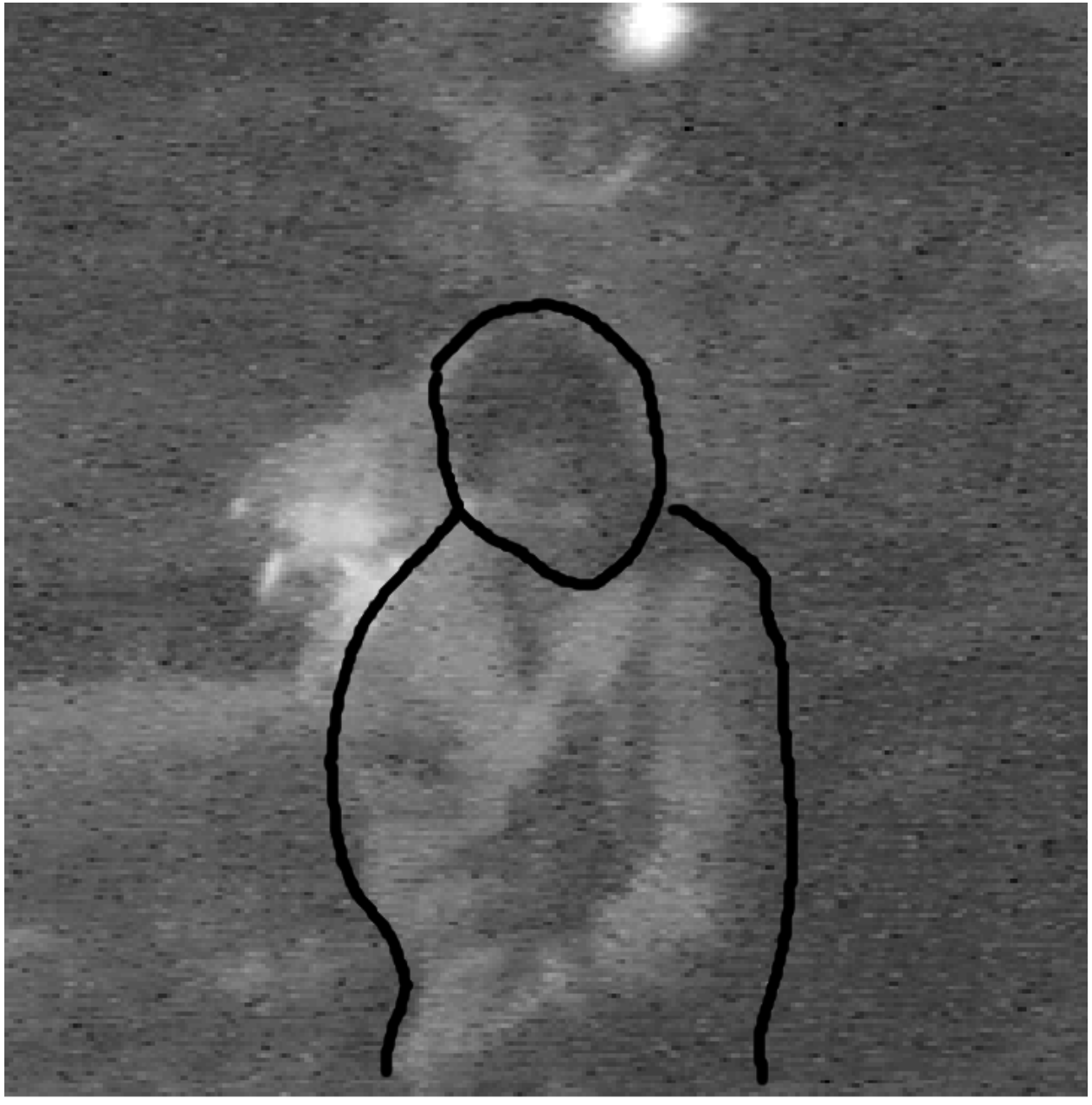
This work is sponsored by the Defense Threat Reduction Agency and NSF-DMS.


# The Graveyard Murderer

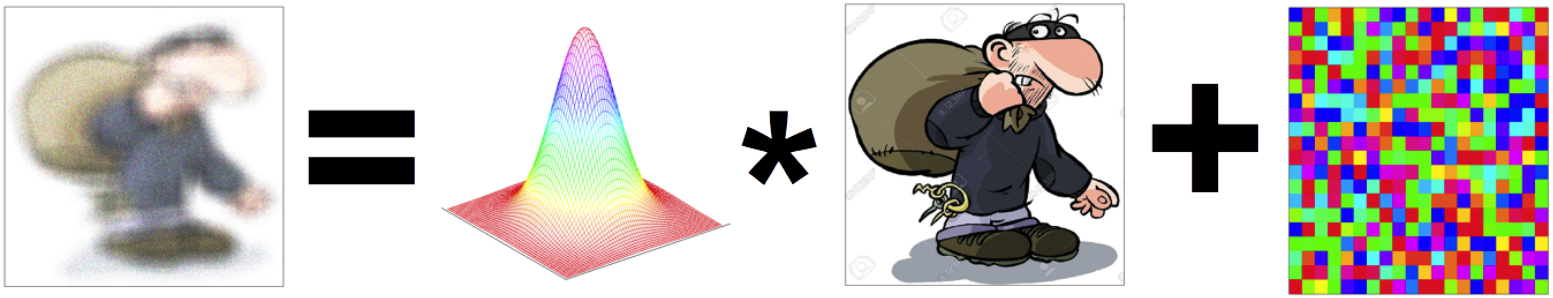
# The Graveyard Murderer

# The Graveyard Murderer

# The Graveyard Murderer



# Blind Deconvolution



$$y = g * z + w$$

We observe a function $y$ which consists of the convolution of two unknown functions, the blurring function $g$ and the signal of interest $z$, plus noise $w$

# Blind deconvolution

Given: $y = g * z$, where "$*$" denotes convolution.

"Ordinary" deconvolution: $y$ and $g$ are known, we compute $z$ from $y$ and $g$ via "deconvolution".

Blind deconvolution: Both $g$ and $z$ are unknown. Need to recover $g$ *and* $z$ from $y$.

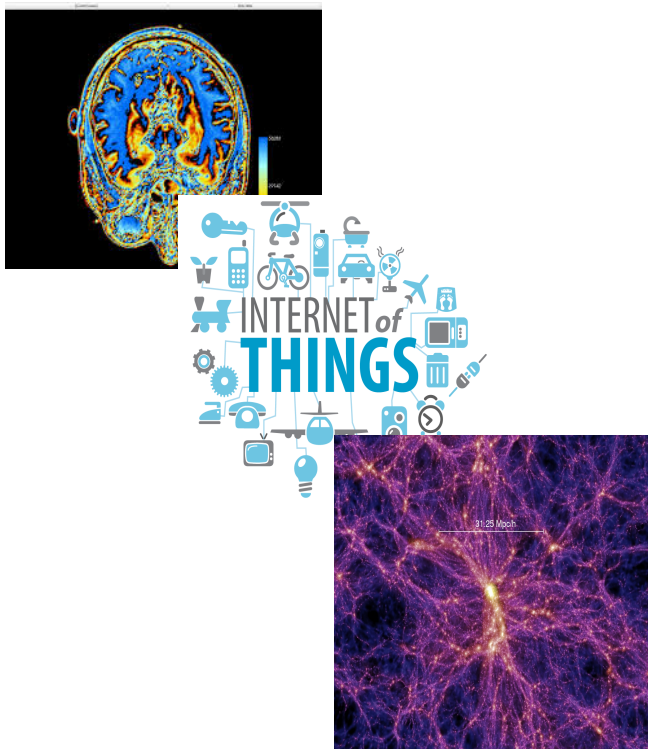
Solvability: The blind deconvolution problem is not solvable without further assumptions on $g$ and $z$.

Note: If $z = \overline{g}$, then this is the famous phase retrieval problem

# Blind deconvolution is ubiquitous

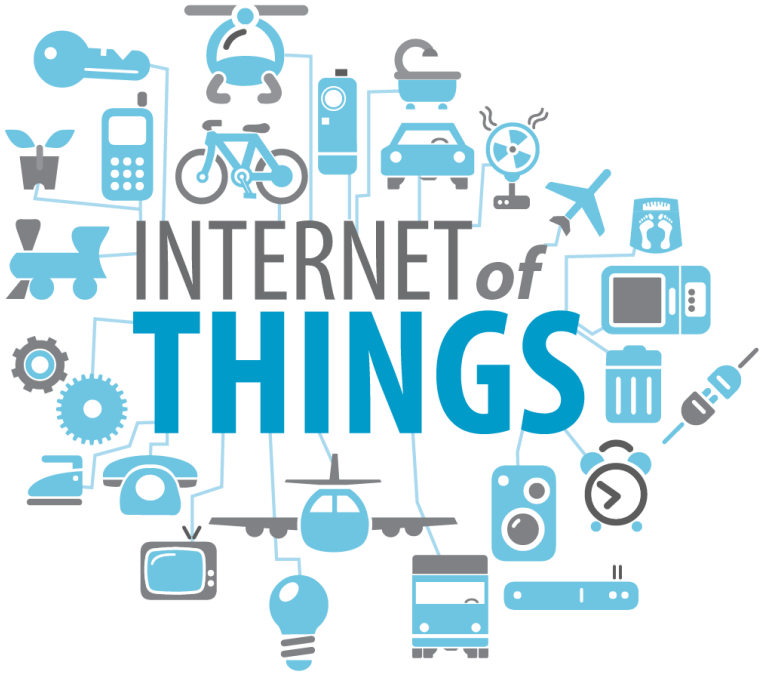
Blind deconvolution problems appear in numerous scientific disciplines and applications:

- Astronomy
- Neuroscience
- Spectroscopy
- Audio processing
- Image processing
- Wireless communications



# The Internet of Things and 5G

Internet of Things (IoT):
Billions of everyday objects have network connectivity, allowing them to send and receive data.





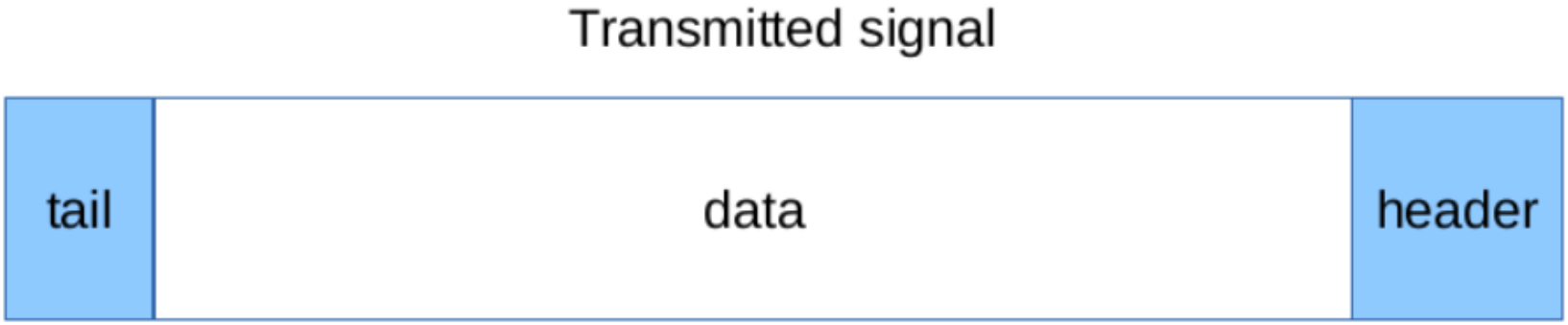
The current wireless system (4G) will not be able to handle the IoT.

The next generation system, 5G, will be designed to run the IoT.

5G needs radical departure from current communication procedures

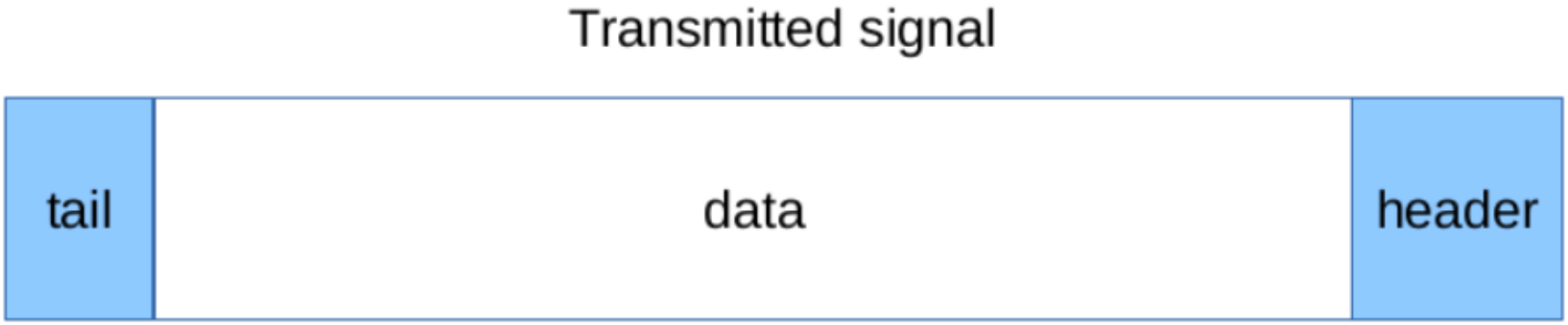
# The headache with the overhead

For each real-time connected device a significant control signaling overhead is necessary to allow for swift channel estimation, equalization and demodulation.

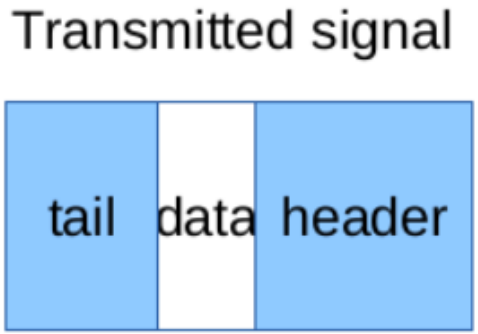


# The headache with the overhead

For each real-time connected device a significant control signaling overhead is necessary to allow for swift channel estimation, equalization and demodulation.




In the IoT most devices will transmit very few data and will do so only sporadically at random time instances



# When the overhead takes over

# When the overhead takes over

# IoT, sporadic traffic and random access

In the IoT we can no longer afford to send signals with the usual control signaling overhead.

Is it possible to design communication systems where the transmitters essentially just send overhead-free signals and the receiver can still extract the information?

Note: Need to do this in a multi-user environment.

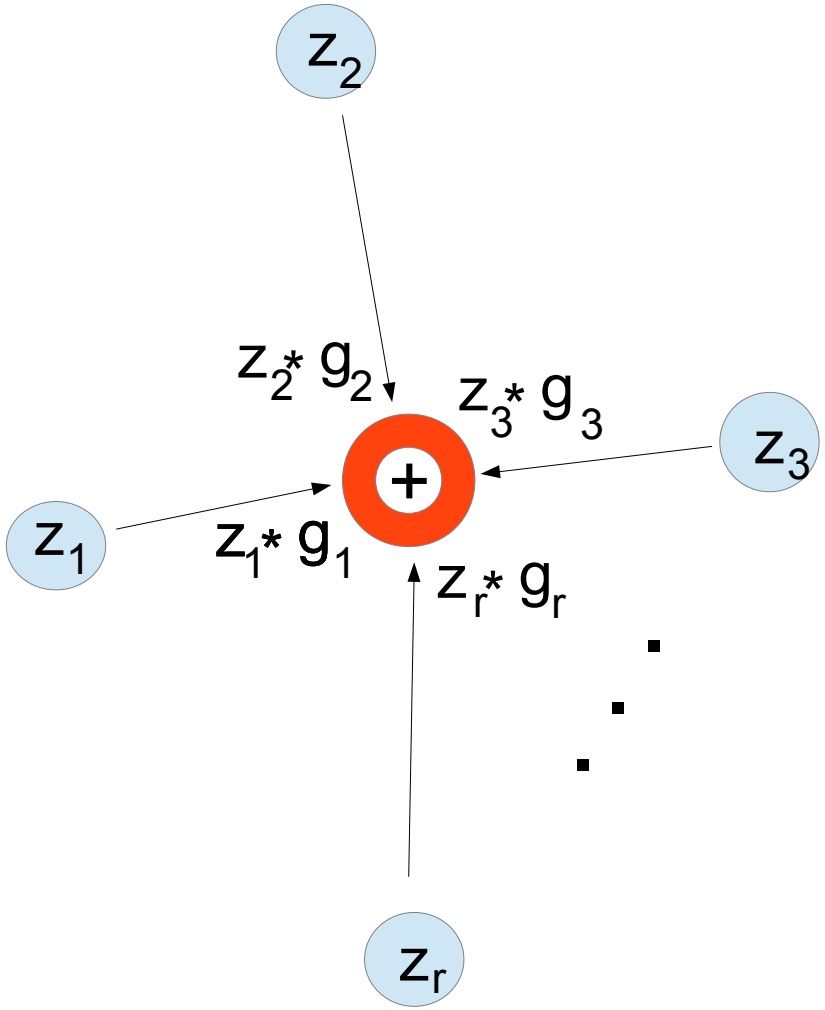
A. Paulraj:

> *"Channel estimation is one of the key problems for 5G – it is the elephant in the room."*

If channel estimation is such a problem, maybe we can eliminate it altogether and thus also reduce signal overhead?

# Blind deconvolution meets blind demixing

Suppose we are given $r$ sensors, each one sends a function $z_i$ (e.g. a signal or image) to a receiver common to all $r$ sensors. During transmission each $z_i$ gets convolved with a function $g_i$.

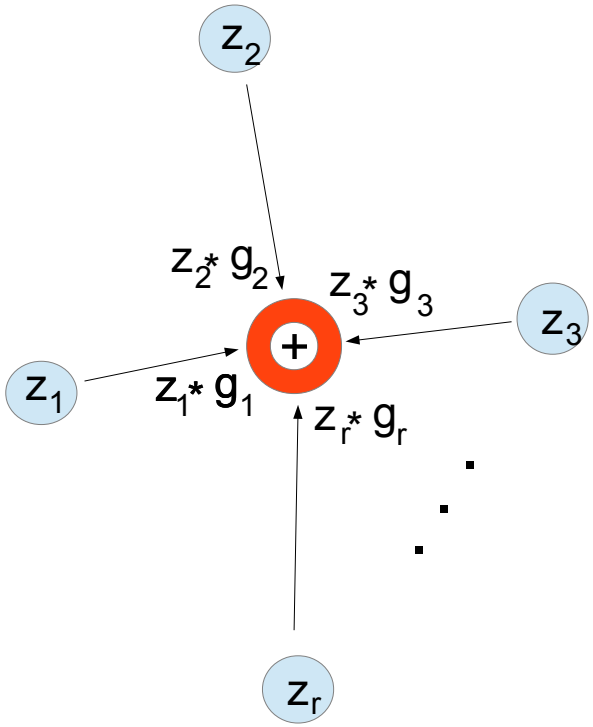


# Blind deconvolution meets blind demixing

The receiver measures the signal $y$, consisting of the sum of all these convolved signals:

$$y = \sum_{i=1}^{r} z_i * g_i + w,$$

where $w$ is additive noise.




Assume the receiver does neither know the $z_i$ nor the $g_i$. When and with which algorithm is it possible to recover all individual signals $z_i$ and $g_i$ from one received signal $y$?

# Toward an efficient solver for blind deconvolution

Can we construct a blind deconvolution algorithm that is:

- numerically efficient,
- robust against noise,
- and comes with rigorous convergence guarantees?

We want to have it all: Rapid, Robust, Reliable!

Want a general framework that does not only work for images, but for other types of applications as well!

# Blind Deconvolution

For concreteness, assume we are given

$$y = g * z + w,$$

where $y, g, z, w \in \mathbb{C}^L$ and only $y$ is known.
Without any further assumptions this problem is ill-posed.
Essentially any arbitrary function $z$ can serve as solution.

# Blind Deconvolution

For concreteness, assume we are given

$$y = g * z + w,$$

where $y, g, z, w \in \mathbb{C}^L$ and only $y$ is known.
Without any further assumptions this problem is ill-posed.
Essentially any arbitrary function $z$ can serve as solution.

Assumption: $g$ as well as $z$ belong to known subspaces:
There exist known matrices $A \in \mathbb{C}^{L \times K}$ and $B \in \mathbb{C}^{L \times N}$ such that

$$g = Bh, \qquad z = Ax,$$

for some unknown functions $h \in \mathbb{C}^N, x \in \mathbb{C}^K$.

# Subspace assumption

Subspace assumption provides flexibility and concreteness. Choice of subspaces depends on specific application.

## Subspace for blurring function:

For example, the blurring function $g$ is often a low-pass filter and all we need to know about $g$ is an upper bound on its bandwidth. This is sufficient to determine (or estimate) $B$.

In communications all we need to know is the so-called maximum delay spread

## Subspace for signal of interest:

In communications, $A$ corresponds to the encoding matrix or spreading matrix, which is often known at base station.

In image processing this is more challenging.

# Blind deconvolution and nonconvex optimization

Using basic Fourier analysis (and assuming circulant convolution and abusing notation) we can rewrite the blind deconvolution problem as:

$$y = \operatorname{diag}(Bh)\overline{Ax} + w,$$

We can attempt to solve for $x$ and $h$ via

$$\min_{(h,x)} \| \operatorname{diag}(Bh)\overline{Ax} - y\|^2.$$

This is a non-linear least squares problem.

Denote

$$F(h,x) := \| \operatorname{diag}(Bh)\overline{Ax} - y\|^2.$$

# Lifting and blind deconvolution

Following recent ideas from phase retrieval, (PhaseLift, [Candes-Strohmer-Voroninski, 2013]), we define the matrix-valued linear operator $\mathcal{A}$ via

$$\mathcal{A} : \mathcal{A}(Z) = \{b_l^* Z a_l\}_{l=1}^L$$

where $b_l$ denotes the $l$-th column of $B^*$ and $a_l$ is the $l$-th column of $A^*$. Hence:

$$y = \mathcal{A}(h_0 x_0^*),$$

and $F(h, x) = \| \operatorname{diag}(Bh)\overline{Ax} - y\|^2$ becomes

$$F(h, x) = \|\mathcal{A}(hx^*) - y\|^2.$$

# Blind deconvolution as semidefinite program

Using this lifting trick, Ahmed, Recht, and Romberg [2014] proposed to relax the blind deconvolution problem into

$$\min_Z \|Z\|_* \qquad \text{subject to} \quad \|\mathcal{A}(Z) - y\|^2 \leq \eta,$$

where $\|Z\|_*$ is the sum of the singular values of $Z$.
Thus, similar to phase retrieval, blind deconvolution can be recast as semidefinite program.
Ahmed, Recht, and Romberg showed that under certain conditions the solution to this problem gives $Z = h_0 x_0^*$.

# Blind deconvolution as semidefinite program

Using this lifting trick, Ahmed, Recht, and Romberg [2014] proposed to relax the blind deconvolution problem into

$$\min_Z \|Z\|_* \qquad \text{subject to} \quad \|\mathcal{A}(Z) - y\|^2 \leq \eta,$$

where $\|Z\|_*$ is the sum of the singular values of $Z$.
Thus, similar to phase retrieval, blind deconvolution can be recast as semidefinite program.
Ahmed, Recht, and Romberg showed that under certain conditions the solution to this problem gives $Z = h_0 x_0^*$.

Advantage: Convex optimization, nice theoretical framework.

Drawback: Not suitable for medium- or large scale problems, since solving an SDP is computationally expensive.

# Toward an efficient solver for blind deconvolution

Main issue: $F(h, x)$ has many local minima, thus a simple alternating minimization algorithm or gradient descent will likely get stuck in some suboptimal solution.

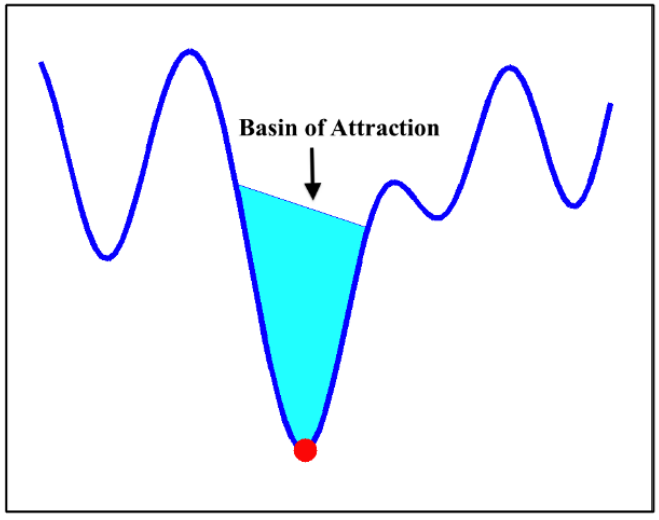
Candes, Li, and Soltanolkotabi [2015] proposed a two-step approach for the nonlinear phase retrieval problem:

1. Use a good starting point obtained via spectral initialization (based on lifting trick)
2. Use gradient descent to find the true solution.

Their Wirtinger-Flow framework yields an efficient algorithm to solve the phase retrieval problem.

# Toward an efficient solver for blind deconvolution

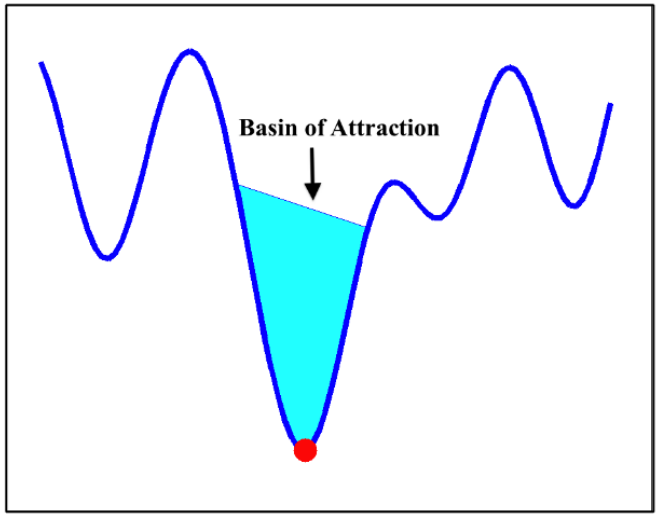
We follow a similar general two-step philosophy: We use some form of spectral initialization to construct a starting point that puts us in the “basin of attraction” and some form of gradient descent to compute the true solution.




Surprisingly, however, the actual algorithm and the theoretical analysis are quite different from the Wirtinger-Flow approach.

# The initial guess

Observation 1: We need to carefully choose the initial guess that puts us in the “basin of attraction”, otherwise the algorithm will get stuck in a local minimum.




Need to define some neighborhood to quantify nearness to true solution.

# (Non)uniqueness of solution

Observation 2: If the pair $(g_0, z_0)$ is a solution to $y = g * z$, then so is the pair $(\alpha g_0, \alpha^{-1} z_0)$ for any $\alpha \neq 0$.

Thus the blind deconvolution problem always has infinitely many solutions of this type.

Recovery of the true solution $(g_0, z_0)$ therefore always means modulo such trivially assoicated solutions $(\alpha g_0, \alpha^{-1} z_0)$.

However, from a numerical viewpoint we need to prevent that $\|g\| \to 0$ and $\|z\| \to \infty$ (or vice versa), while $\|g\| \cdot \|z\| = c$.

# The Importance of Being Incoherent

Observation 3: Our numerical experiments have shown that the algorithm's performance depends on how much the rows of $B$ and the vector $h_0$ are correlated.

Let $\mu_h^2 = \frac{\|Bh_0\|_\infty^2}{\|h_0\|^2}$. The smaller $\mu_h$, the better.

# The Importance of Being Incoherent

Observation 3: Our numerical experiments have shown that the algorithm's performance depends on how much the rows of $B$ and the vector $h_0$ are correlated.

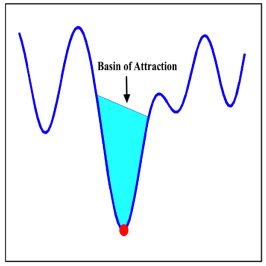
Let $\mu_h^2 = \frac{\|Bh_0\|_\infty^2}{\|h_0\|^2}$. The smaller $\mu_h$, the better.

This incoherence property reminds us of matrix completion: The left and right singular vectors of the solution matrix cannot be "too aligned" with those of the measurement matrices.

This suggests to introduce a term that controls the incoherence between $B$ and the iterates $h_k$.

# Building a basin of attraction

Based on the three observations above, we define the three neighborhoods (denoting $d_0 = \|h_0\|\|x_0\|$):




$$
\begin{aligned}
\mathcal{N}_{d_0} &:= \{(h,x) : \|h\| \le 2\sqrt{d_0}, \|x\| \le 2\sqrt{d_0}\} \\
\mathcal{N}_{\mu} &:= \{h : \|Bh\|_\infty \le 4\mu\sqrt{d_0/L}\} \\
\mathcal{N}_{\epsilon} &:= \{(h,x) : \|hx^* - h_0x_0^*\|_F \le \varepsilon d_0\}.
\end{aligned}
$$

We first obtain a good initial guess $(u_0, v_0) \in \mathcal{N}_{d_0} \cap \mathcal{N}_{\mu} \cap \mathcal{N}_{\epsilon}$, which is followed by regularized gradient descent.

Regularization forces iterates $(u_k, v_k)$ inside $\mathcal{N}_{d_0} \cap \mathcal{N}_{\mu} \cap \mathcal{N}_{\epsilon}$.

# Regularized Gradient Descent

We consider

$$\widetilde{F}(h,x) = F(h,x) + G(h,x)$$

where the regularization function $G(h,x)$ is the form of

$$G(h,x) = \rho\Big[G_0\left(\frac{\|h\|^2}{2d}\right) + G_0\left(\frac{\|x\|^2}{2d}\right) + \sum_{l=1}^{L} G_0\left(\frac{L|b_l^* h|^2}{8d\mu^2}\right)\Big]$$

with $G_0(z) = \max\{z-1, 0\}^2$.

# Regularized Gradient Descent

We consider

$$\widetilde{F}(h,x) = F(h,x) + G(h,x)$$

where the penalty function $G(h,x)$ is the form of

$$G(h,x) = \rho\Big[\underbrace{G_0\left(\frac{\|h\|^2}{2d}\right) + G_0\left(\frac{\|x\|^2}{2d}\right)}_{K_{d_0}} + \underbrace{\sum_{l=1}^{L} G_0\left(\frac{L|b_l^* h|^2}{8d\mu^2}\right)}_{K_\mu}\Big]$$

with $G_0(z) = \max\{z-1, 0\}^2$.

# The RGD algorithm

## Phase 1: Initialization via spectral method and projection:

Compute $\mathcal{A}^*(y)$.
Find the leading singular value, left and right singular vectors of $\mathcal{A}^*(y)$, denoted by $d$, $\hat{h}_0$ and $\hat{x}_0$ respectively.
Set $v_0 = \sqrt{d}\hat{x}_0$ and solve:

$$u_0 := \text{argmin}_z \|z - \sqrt{d}\hat{h}_0\|^2, \quad \text{subject to} \quad \sqrt{L}\|Bz\|_\infty \leq 2\sqrt{d}\mu$$

This ensures that $(u_0, v_0) \in \mathcal{N}_{d_0} \cap \mathcal{N}_\mu \cap \mathcal{N}_\epsilon$.

## Phase 2: Gradient descent with constant stepsize:

**Initialization:** obtain $(u_0, v_0)$ via Phase 1.
**for** $k = 1, 2, \ldots,$ **do**

$$u_k = u_{k-1} - \eta \nabla \widetilde{F}_h(u_{k-1}, v_{k-1})$$
$$v_k = v_{k-1} - \eta \nabla \widetilde{F}_x(u_{k-1}, v_{k-1})$$

**end for**

**Theorem:** [Li-Ling-Strohmer-Wei, 2016].

Consider $$y = \mathrm{diag}(Bh_0)\overline{Ax_0} + w,$$

where $B$ is an $L \times N$ low-pass Fourier matrix, $A$ is an $L \times K$ Gaussian random matrix and the noise $w \sim \mathcal{N}(0, \frac{\sigma^2 d_0^2}{L})$.
If the number of measurements satisfies

$$L \geq C(\mu_h^2 + \sigma^2) \max\{K, N\} \log^2(L)/\varepsilon^2,$$

(i) then the initialization $(u_0, v_0) \in \frac{1}{\sqrt{3}}\mathcal{N}_{d_0} \bigcap \frac{1}{\sqrt{3}} K_\mu \bigcap K_{\frac{2}{5}\varepsilon}$;
(ii) our algorithm creates a sequence $(u_k, v_k) \in \mathcal{N}_{d_0} \cap \mathcal{N}_\mu \cap \mathcal{N}_\epsilon$, which converges linearly to $(h_0, x_0)$; i.e., with high probability:

$$\max\{\sin\angle(u_k, h_0), \sin\angle(v_k, x_0)\} \leq \frac{1}{d_k}\left((1-\alpha)^{k/2}\varepsilon d_0 + 40\|\mathcal{A}^*(w)\|\right)$$

where $\alpha = \mathcal{O}(\frac{1}{(K+N)\log^2 L})$ and $d_k := \|u_k\|\|v_k\| \to \|h_0\|\|x_0\|$.

# Proof

The proof relies on four key properties:

1. Local Regularity Condition: establishes that objective function decreases.

$$\|\nabla \widetilde{F}(h,x)\|^2 \geq \omega \widetilde{F}(h,x)$$

2. Local Restricted Isometry Property: allows us to transfer convergence of objective function to convergence of iterates.

$$\frac{3}{4}\|hx^* - h_0x_0^*\|_F^2 \leq \|\mathcal{A}(hx^* - h_0x_0^*)\|^2 \leq \frac{5}{4}\|hx^* - h_0x_0^*\|_F^2$$

holds uniformly for all $(h,x) \in \mathcal{N}_{d_0} \cap \mathcal{N}_\mu \cap \mathcal{N}_\epsilon$.

# Proof

3. Local Smoothness Condition: governs rate of convergence. There exists a constant $C$ such that

$$\|\nabla f(z + t\Delta z) - \nabla f(z)\| \leq Ct\|\Delta z\|, \quad \forall\, 0 \leq t \leq 1,$$

for all $\{(z, \Delta z) : z \in \mathcal{N}_\epsilon \cap \mathcal{N}_{\widetilde{F}}, z + t\Delta z \in \mathcal{N}_\epsilon \bigcap \mathcal{N}_{\widetilde{F}}\}$
i.e., the whole line segment connecting $z$ and $z + \Delta z$ belongs to the nonconvex set $\mathcal{N}_\epsilon \bigcap \mathcal{N}_{\widetilde{F}}$.

4. Robustness Condition: provides stability against noise.

$$\|\mathcal{A}^*(w)\| \leq \frac{\varepsilon d_0}{10\sqrt{2}}$$

# Numerical simulations

Nonconvex vs. convex optimization: How does Regularized Gradient Descent compare to Nuclear Norm Minimization?

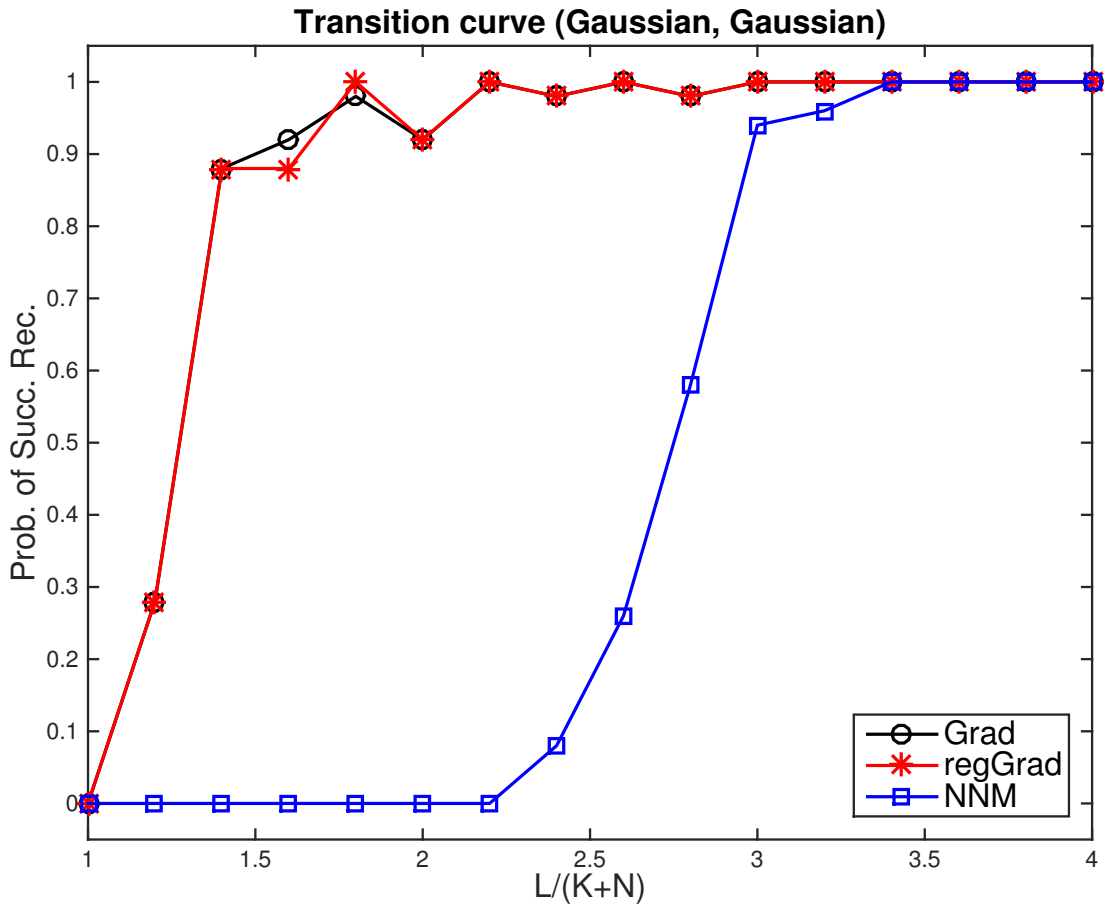
$$\min_{h,x} \widetilde{F}(h,x)$$

vs

$$\min_{Z} \|Z\|_* \text{ s.t. } \|\mathcal{A}(Z) - y\|^2 \leq \eta,$$

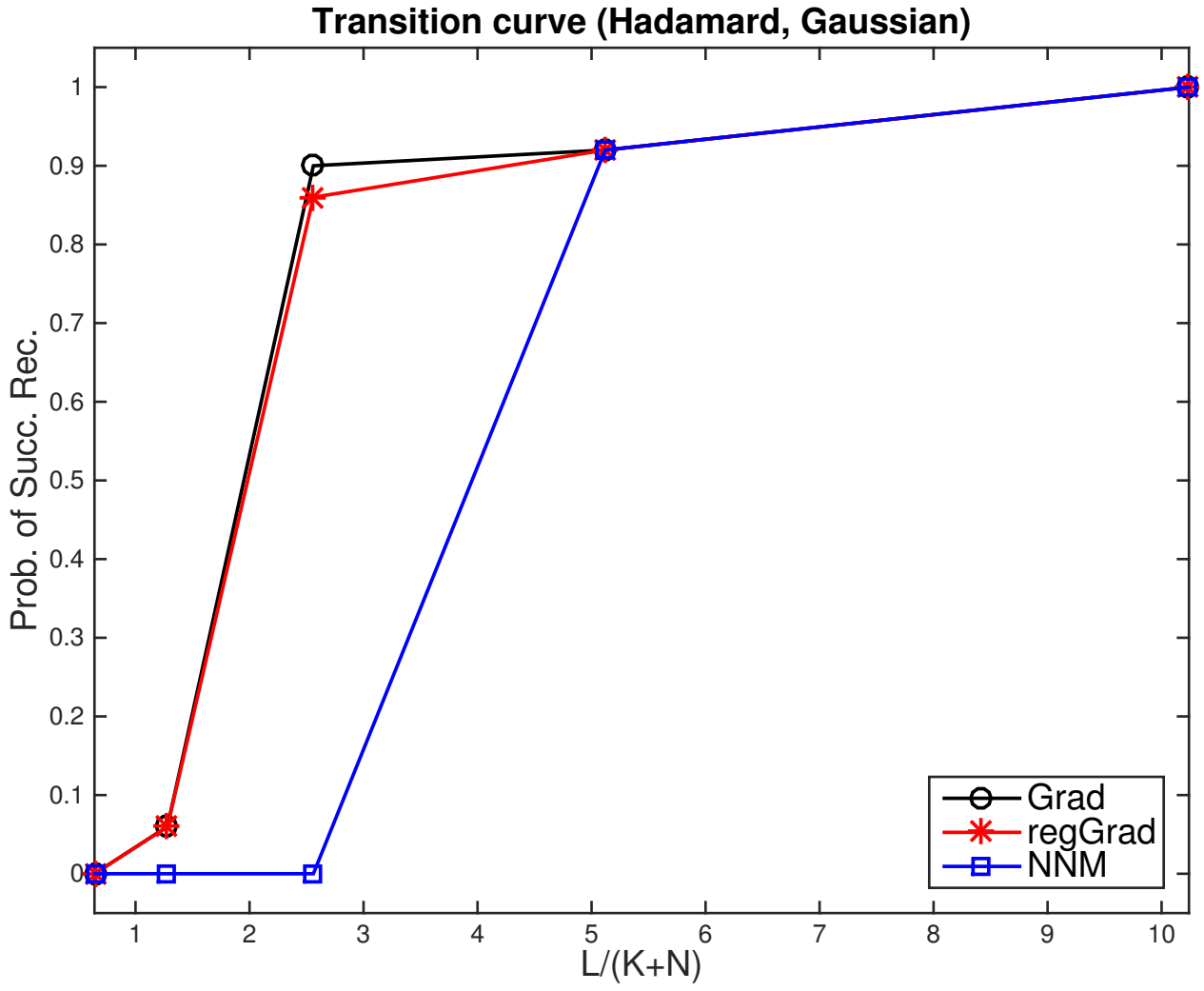
# Nonconvex vs. convex optimization

Empirical phase transition curves when $A$ is random Gaussian.

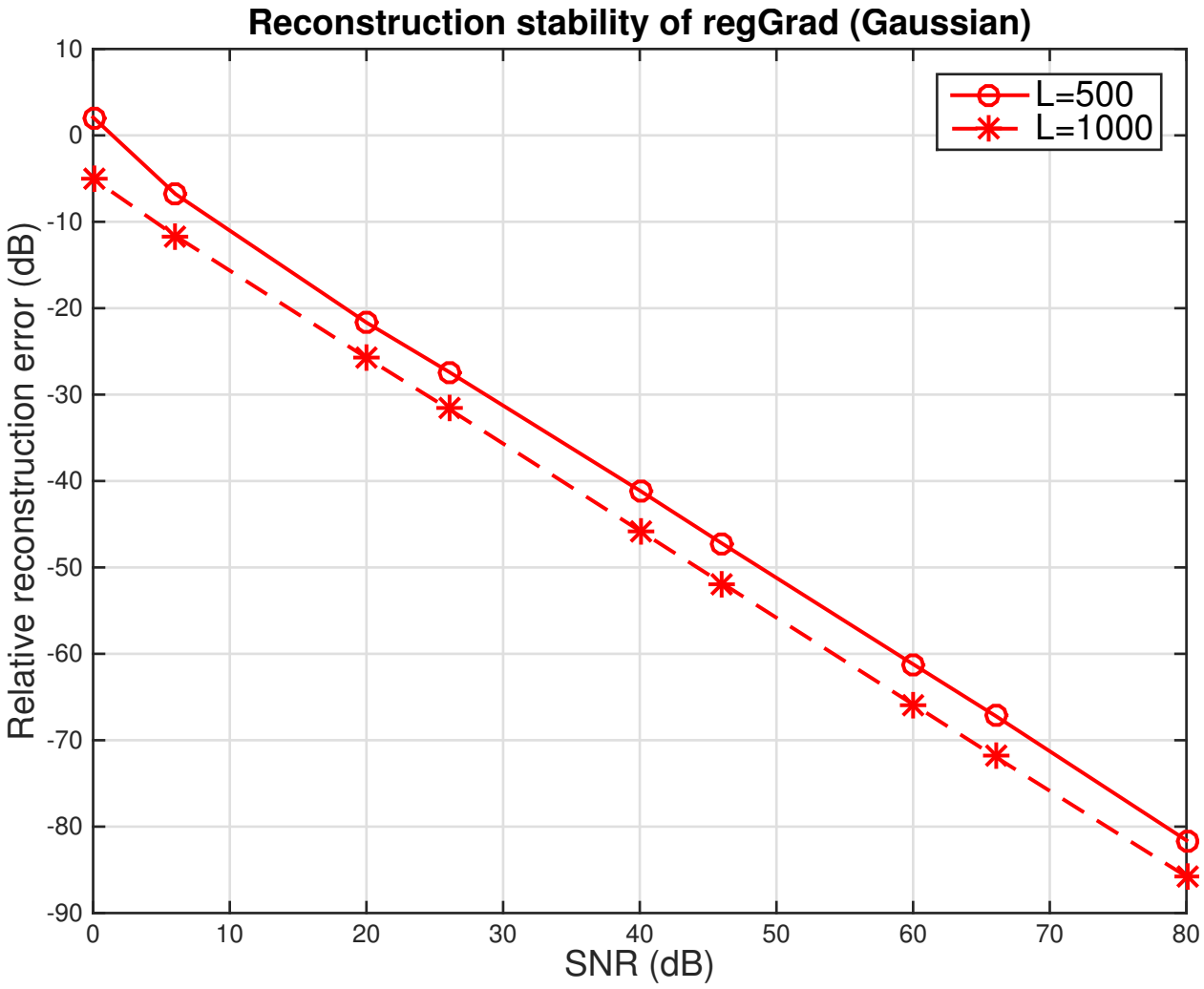


# Nonconvex vs. convex optimization

Empirical phase transition when $A$ is random Hadamard matrix.



# Stability against noise



# Unknown subspaces

In the previous examples the subspaces $A$ and $B$ are known (but of course, the signal and the blurring function are unknown!)

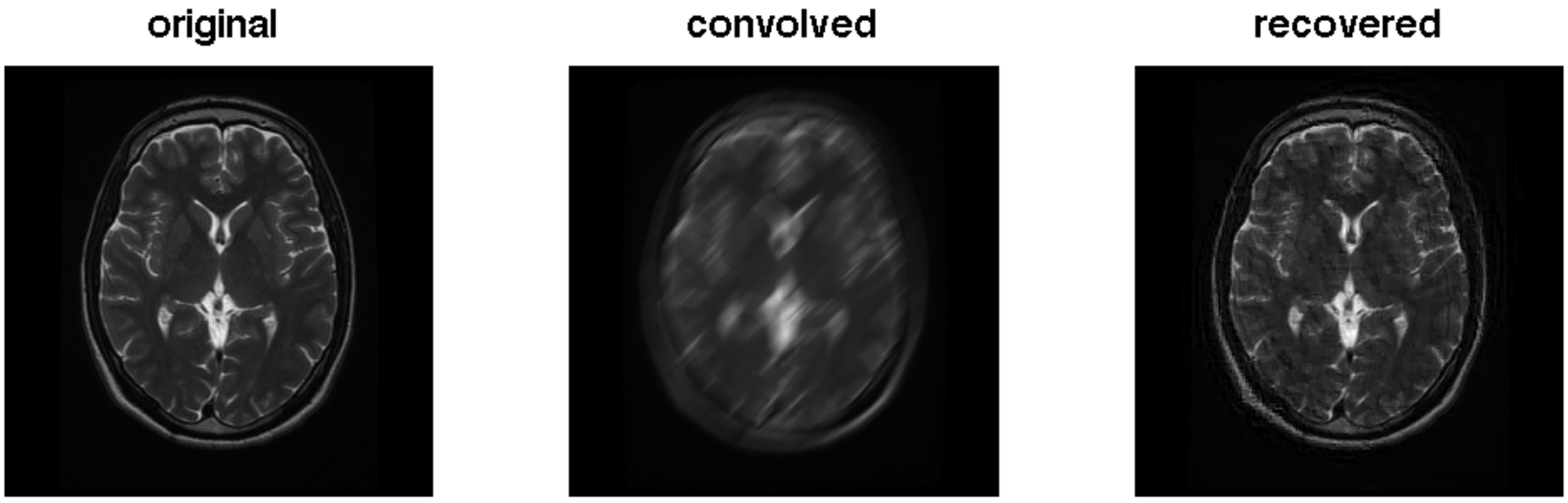
This is a justified assumption in certain applications, such as wireless communications.

In various image processing applications we may neither know $A$ or $B$ (in addition to not knowing the blurring function and the signal).

For images, we might assume that the subspace $A$ belongs to some larger wavelet space, but we may or may not know $A$.

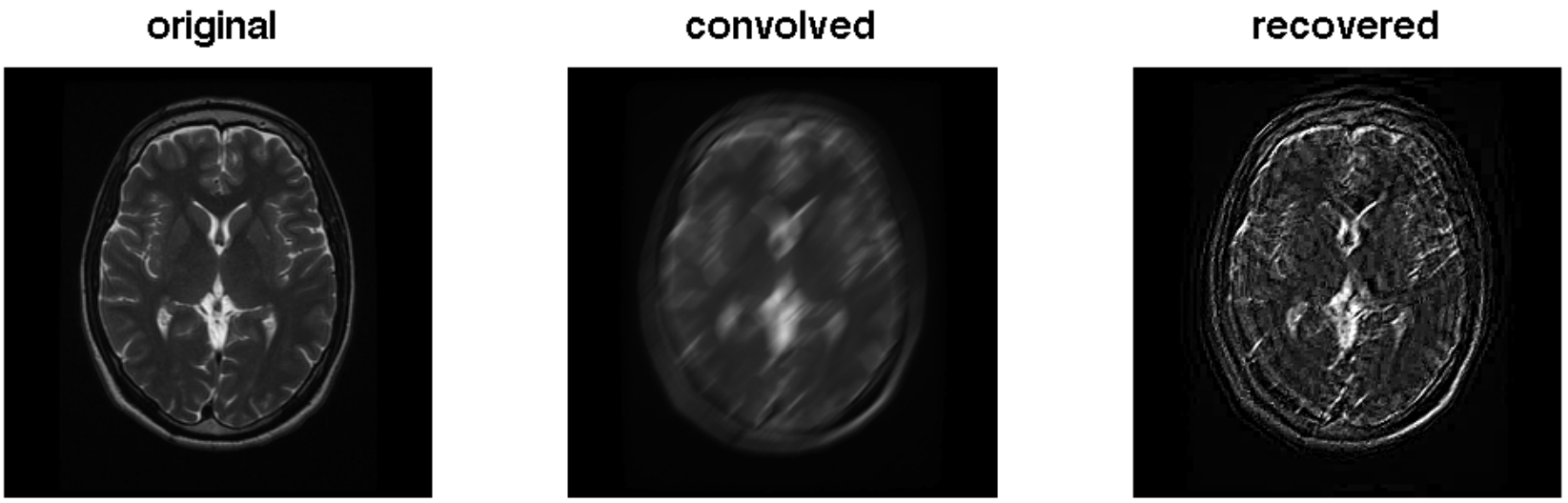
# Blind deconvolution with known blurring subspace

Subspace *B* is known, but subspace *A* is unknown.



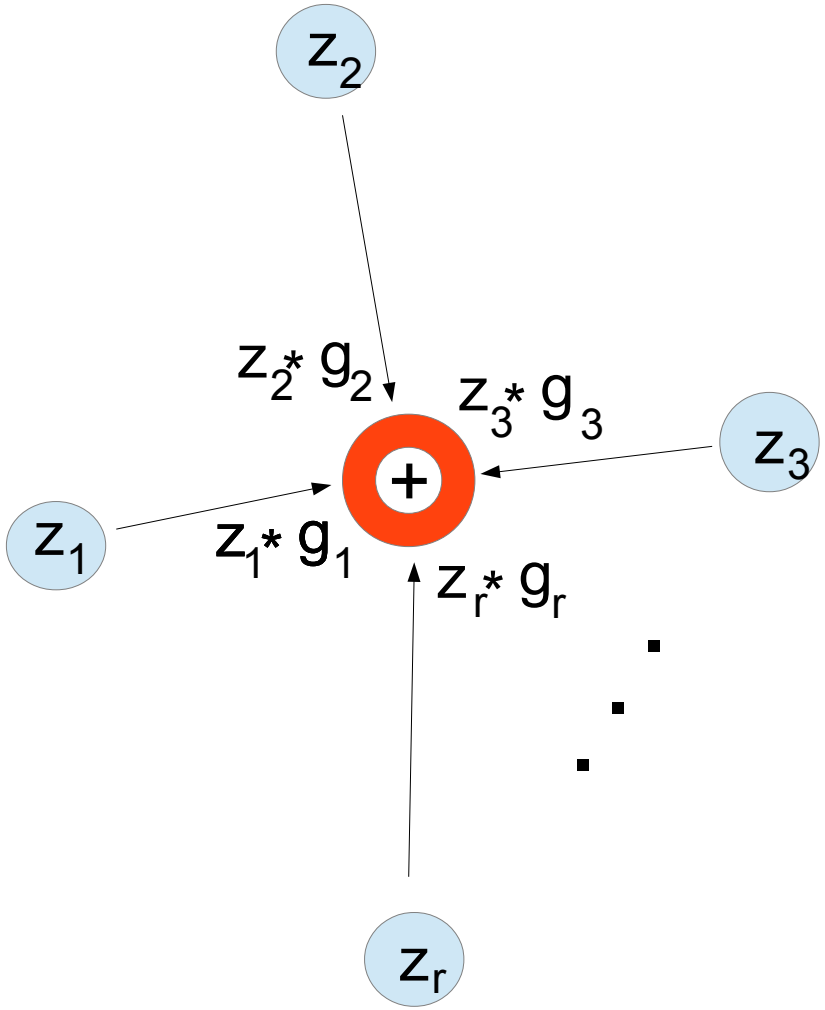
# Blind deconvolution with unknown subspaces

Subspace $B$ is unknown and subspace $A$ is unknown.



# Blind deconvolution meets blind demixing

Suppose we are given $r$ sensors, each one sends a function $z_i$ (e.g. a signal or image) to a receiver common to all $r$ sensors. During transmission each $z_i$ gets convolved with a function $g_i$.



# The basic model

We are given $y \in \mathbb{C}^L$ with

$$y = \sum_{i=1}^{r} z_i * g_i + w,$$

and we want to find all $z_i$ and $g_i$.
This problem is highly underdetermined, thus we need some prior information about $z_i$ and $g_i$ to have a chance to solve it.

Often we can assume that the $g_i$ as well as the $z_i$ belong to some known (but perhaps for each $i$ different) subspaces:

$$z_i = A_i x_i \text{ for } A_i \in \mathbb{C}^{L \times N_i}, \quad \text{and} \quad g_i = B_i h_i \text{ for } B_i \in \mathbb{C}^{L \times K_i},$$

for known $A_i, B_i, i = 1, \ldots, r$.

# The basic model

We are given $y \in \mathbb{C}^L$ with

$$y = \sum_{i=1}^{r} z_i * g_i + w,$$

and we want to find all $z_i$ and $g_i$.
This problem is highly underdetermined, thus we need some prior information about $z_i$ and $g_i$ to have a chance to solve it.

Often we can assume that the $g_i$ as well as the $z_i$ belong to some known (but perhaps for each $i$ different) subspaces:

$$z_i = A_i x_i \text{ for } A_i \in \mathbb{C}^{L \times N}, \quad \text{and} \quad g_i = B_i h_i \text{ for } B_i \in \mathbb{C}^{L \times K},$$

for known $A_i, B_i, i = 1, \ldots, r$.

# The basic model

We are given $y \in \mathbb{C}^L$ with

$$y = \sum_{i=1}^{r} z_i * g_i + w,$$

and we want to find all $z_i$ and $g_i$.
This problem is highly underdetermined, thus we need some prior information about $z_i$ and $g_i$ to have a chance to solve it.

Often we can assume that the $g_i$ as well as the $z_i$ belong to some known (but perhaps for each $i$ different) subspaces:

$$z_i = A_i x_i \text{ for } A_i \in \mathbb{C}^{L \times N}, \quad \text{and} \quad g_i = B_i h_i \text{ for } B_i \in \mathbb{C}^{L \times K},$$

for known $A_i, B_i, i = 1, \ldots, r$.
We need at least $L \geq r(N + K)$ many measurements to have a theoretical chance to solve the problem.

Proceeding with the meanwhile well-established lifting trick, we let $X_i := h_i x_i^T \in \mathbb{R}^{K \times N}$ and define for $i = 1, \dots, r$

$$\mathcal{A}_i(Z) := \{b_{i,l}^* Z a_{i,l}\}_{l=1}^L.$$

We can lift the *non-linear vector-valued* equations to *linear matrix-valued* equations given by

$$y = \sum_{i=1}^r \mathcal{A}_i(X_i).$$

Since this system is highly underdetermined, we consider

$$\min \sum_{i=1}^r \|Z_i\|_* \quad \text{subject to} \quad \sum_{i=1}^r \mathcal{A}_i(Z_i) = y. \qquad \text{(SDP)}$$

If the solutions $Z_1, \dots, Z_r$ are all rank-one, we can easily extract $x_i$ and $h_i$ from $Z_i$ via a simple matrix factorization.

**Theorem:** [Ling-Strohmer 2015]
Let the $A_i$ be $L \times N$ i.i.d. Gaussian random matrices. Assume that the impulse responses $g_i$ satisfy $g_i(k) = 0$ if $k \geq K$.
Let $\mu_h$ be a certain "incoherence parameter" related to the measurement matrices. Suppose we are given

$$y = \sum_{i=1}^{r} g_i * (A_i x_i).$$

Then, as long as the number of measurements $L$ satisfies

$$L \gtrsim C r^2 \max\{K, \mu_h^2 N\} \log L^3 \log r,$$

(where $C$ is a numerical constant), all $x_i$ (and thus $z_i = A_i x_i$) as well as all $g_i$ can be recovered from $y$ with high probability by solving (SDP).

**Theorem:** [Ling-Strohmer 2015]
Let the $A_i$ be $L \times N$ i.i.d. Gaussian random matrices. Assume that the impulse responses $g_i$ satisfy $g_i(k) = 0$ if $k \geq K$.
Let $\mu_h$ be a certain "incoherence parameter" related to the measurement matrices. Suppose we are given

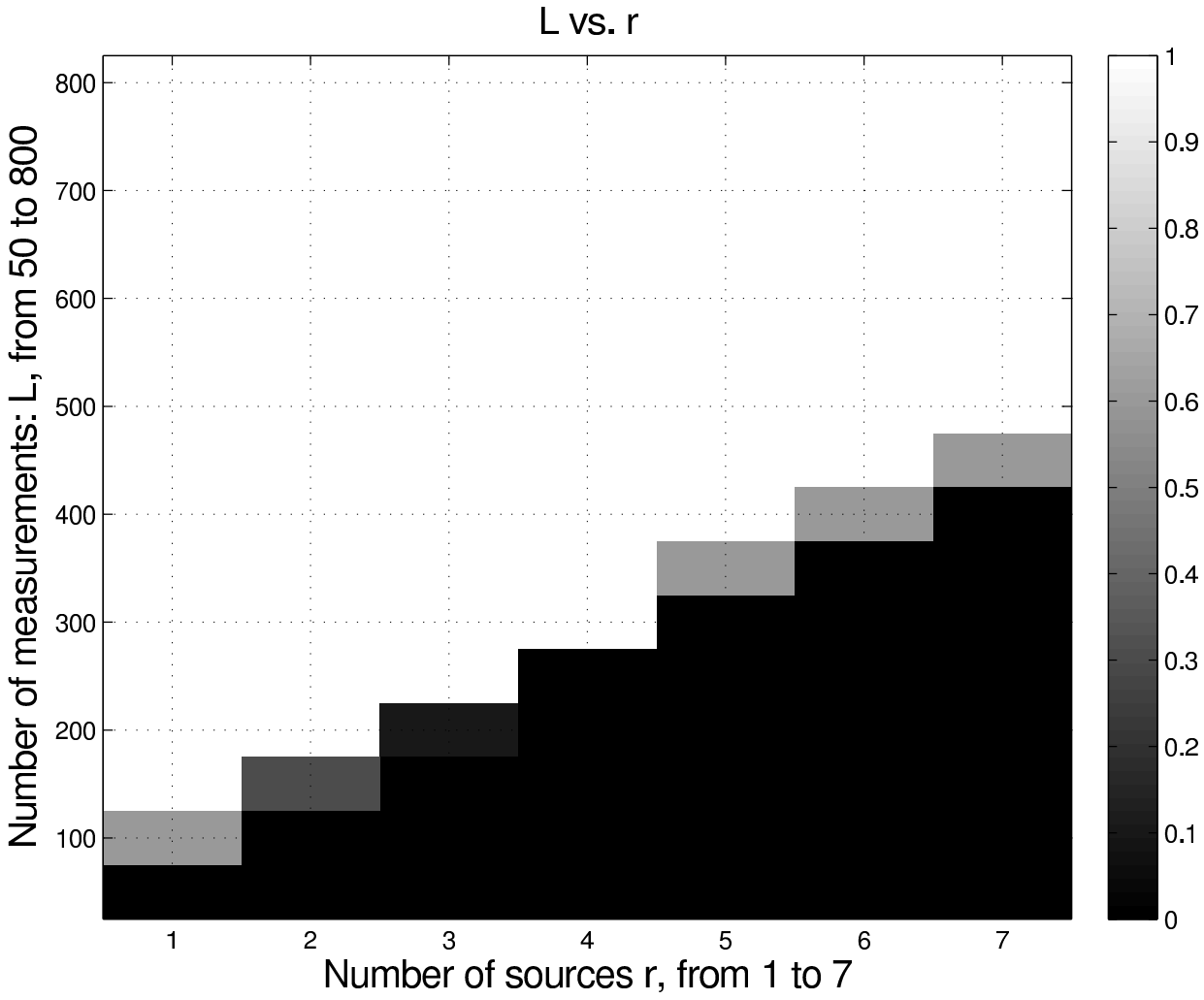
$$y = \sum_{i=1}^{r} g_i * (A_i x_i).$$

Then, as long as the number of measurements $L$ satisfies

$$L \gtrsim C r^2 \max\{K, \mu_h^2 N\} \log L^3 \log r,$$

(where $C$ is a numerical constant), all $x_i$ (and thus $z_i = A_i x_i$) as well as all $g_i$ can be recovered from $y$ with high probability by solving (SDP).

# Does $L$ really scale with $r^2$?



# Does $L$ really scale with $r^2$?

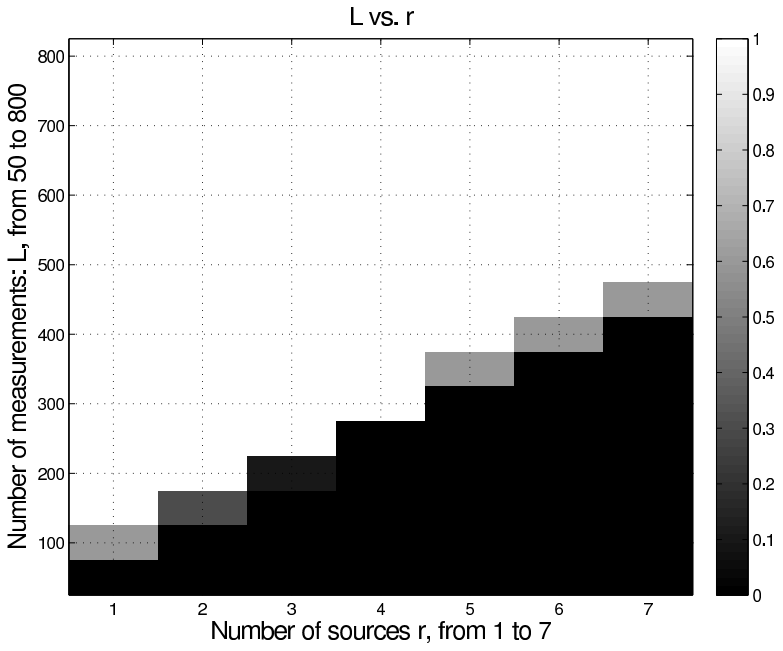
Numerical experiments show that $L$ scales with $r$.

# Does $L$ really scale with $r^2$?

Numerical experiments show that $L$ scales with $r$.

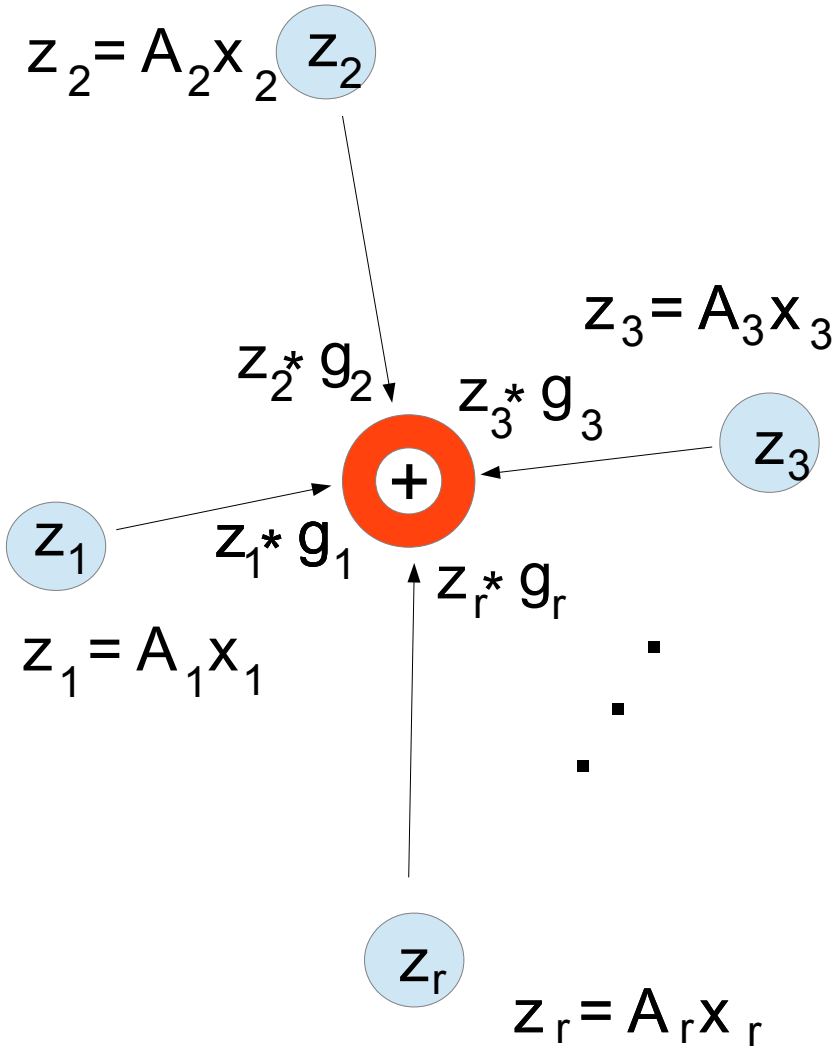
**Surprise:** $L \approx r(K+\mu_h^2 N)$ measurements seem to suffice, which is close to the information-theoretic limit.

In numerical example, e.g. $K = N = 30, r = 7$, hence $r(K+N) = 420$ and with $L > 480$ get exact recovery.



# Back to the Internet-of-Things

At each sensor the signal $x_k$ gets "precoded" by applying $A_k$.



# Back to the Internet-of-Things

A large number of simple sensors sporadically send a few data to a base station (and a few other sensors send frequently a lot of data).

Each sensor “precodes” the signal $x_k$ by applying $A_k$ (like in standard CDMA).

The base station only needs to know which of the users is active. Can be achieved via very-low rate feedback channel.

The base station can recover all signals without any channel estimation (unlike in standard CDMA) with only minimal overhead.

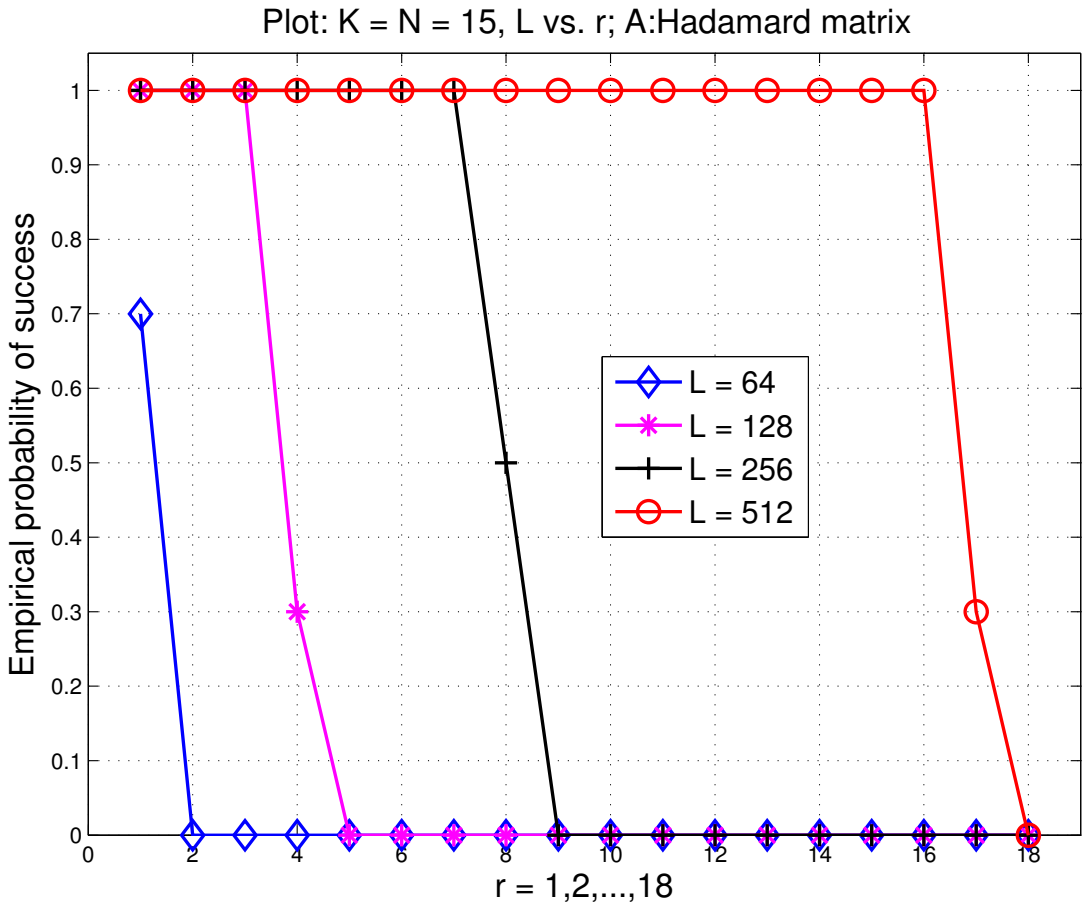
# A Toy Internet-of-Things Example

A practical choice for $A_k$ (although not yet covered by our theory):
Let $A_0$ be a tall subsection of a Hadamard matrix.
Define $A_k = \boldsymbol{D}_k A_0$, where $\boldsymbol{D}_i$ is a diagonal matrix with entries $\pm 1$.

Can either
(i) choose diagonal entries $\pm 1$ randomly.
or
(ii) choose diagonal entries $\pm 1$ deterministically according to design of mutually unbiased basis (for how to do this see Calderbank-Cameron-Kantor-Seidel 1995).

Can easily demix dozens of users without channel estimation.

# A Toy Internet-of-Things Example




Number of measurements for empirical recovery is close to information-theoretic limit in these examples.
E.g. for $r = 16 : r(K + N) = 480$ and $L = 512$ gives successful recovery!

# Calibration issues everywhere

- Femtosecond X-ray Nanocrystallography
- Compressive microscopy: calibration of coded aperture (similar to structured illuminations with random masks)
- Radar: calibration of antenna arrays
- Time-interleaved ADCs: unknown time-offsets and gains
- Astronomy: E.g. James Webb Space Telescope
- Chemical cloud detection
- Exploration geophysics

# Conclusion and Open Problems

- Efficient framework for blind deconvolution.
- Extension to blind deconvolution-blind demixing.
- Could play a role in the future Internet-of-Things
- Open problem: include sparsity in signal or blurring fct.
- Open problem: for image processing need better way to determined unknown subspace $A$.

# Conclusion and Open Problems

- Efficient framework for blind deconvolution.
- Extension to blind deconvolution-blind demixing.
- Could play a role in the future Internet-of-Things
- Open problem: include sparsity in signal or blurring fct.
- Open problem: for image processing need better way to determined unknown subspace $A$.
- Open problem: A murderer is still on the loose ....


**50000 €**