cutting through complexity

Thy customer, where are thou?

KPMG's Indoor (Wi-Fi) Tracking: From simulation to implementation

TopQuants 2014

11 November 2014

Outline

- Management Summary
- Theory
- Toy Monte Carlo
- Proof – of – concept
- Summary

Management Summary

Short and sweet

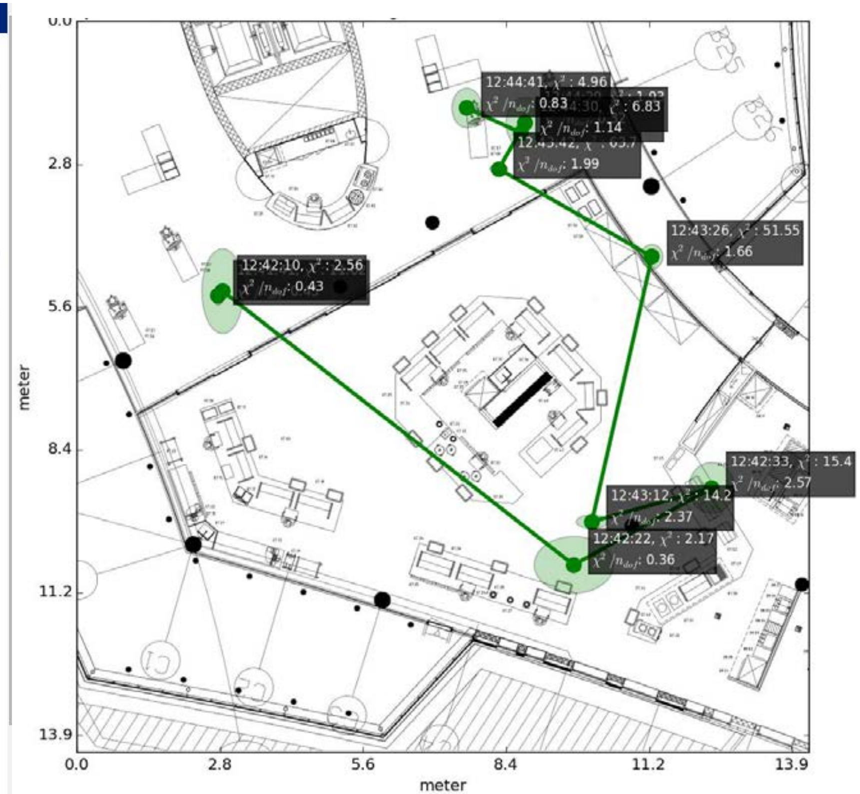
Why?

- Crowd monitoring
- Customer flow
- Dwell times
- Occupancies
- Waiting times at tills
- Scalability: no clickers, no pen and paper
- Passive, less intrusive
- …

→

- Crowd (flow) prediction/monitoring
- Improve layout of store
- “Measure” sales performance
- Staff at the right time and place
- Improved shopping experience: who wants to wait at the till
- Know thy customer: Customer decision journey
- Sales staff intervention
- Correlate with purchases/sales
- …

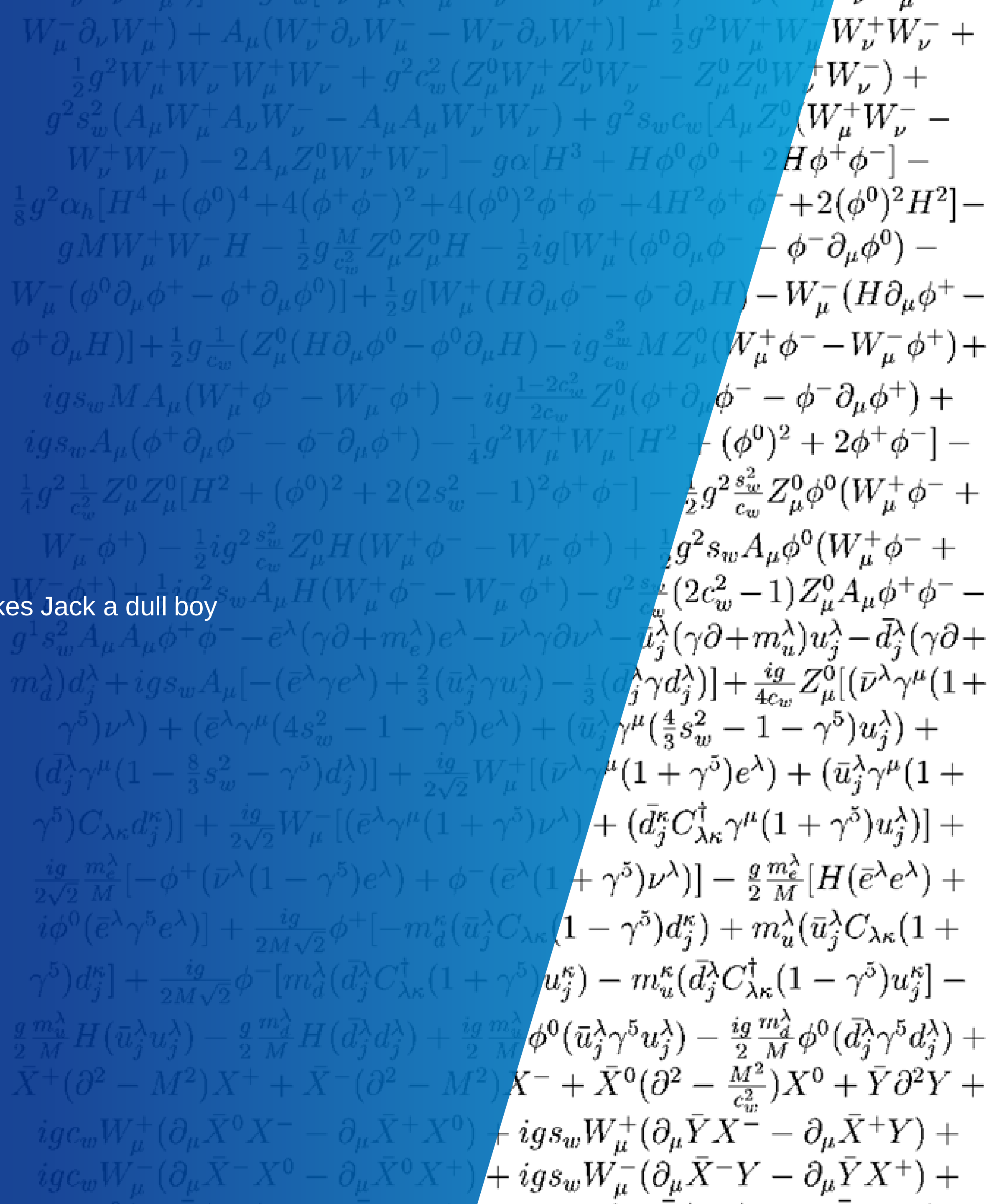
Take Home …



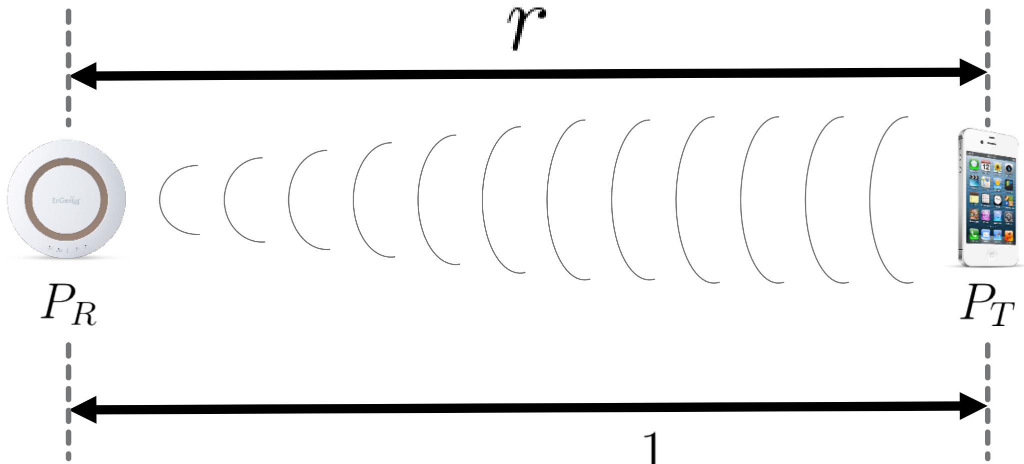
IT WORKS !!!!

Theory

All theory and no play makes Jack a dull boy



Friis Free Space Transmission Equation



$$P_R\,[W] \propto \frac{1}{r^2}$$

P_R – transmission power at reciever
P_T – transmission power
$n = 2$ – line of sight, i.e. no obstacles
c – speed of light
f – frequency, 2.4 GHz or 5 GHz
r – distance from transmitter to reciever

This is going to be our model

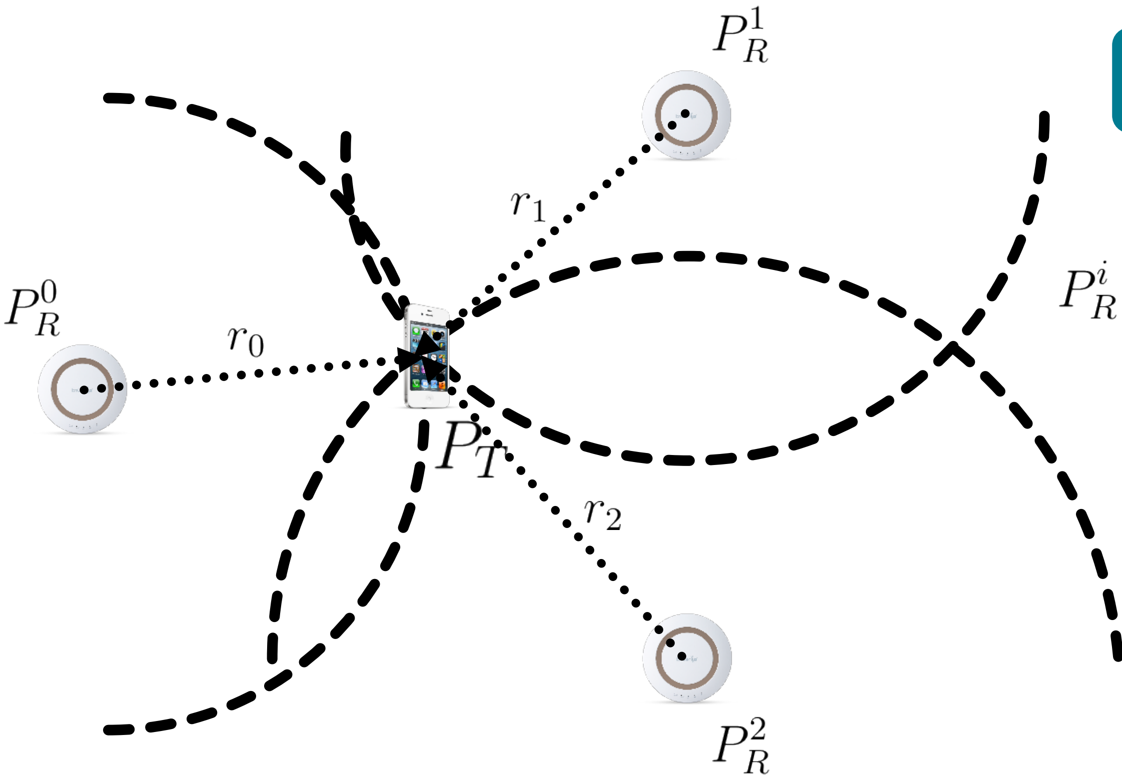
$$P_R\,[dBm] = P_T + 10\log\left(\frac{c}{4\pi f r}\right)^{\boxed{2}}$$

Valid for:

- Line – of – sight: No obstacles that may lead to reflections, refractions, etc.
- $r > 0.4\ m$

Theory

Trilateration



Use this to convert measured signal strength to a distance, i.e. ranging.

$$P_R^i = P_T + 10\log\left(\frac{c}{4\pi f}\right)^2 + 10\log\left(\frac{1}{r_i}\right)^2$$

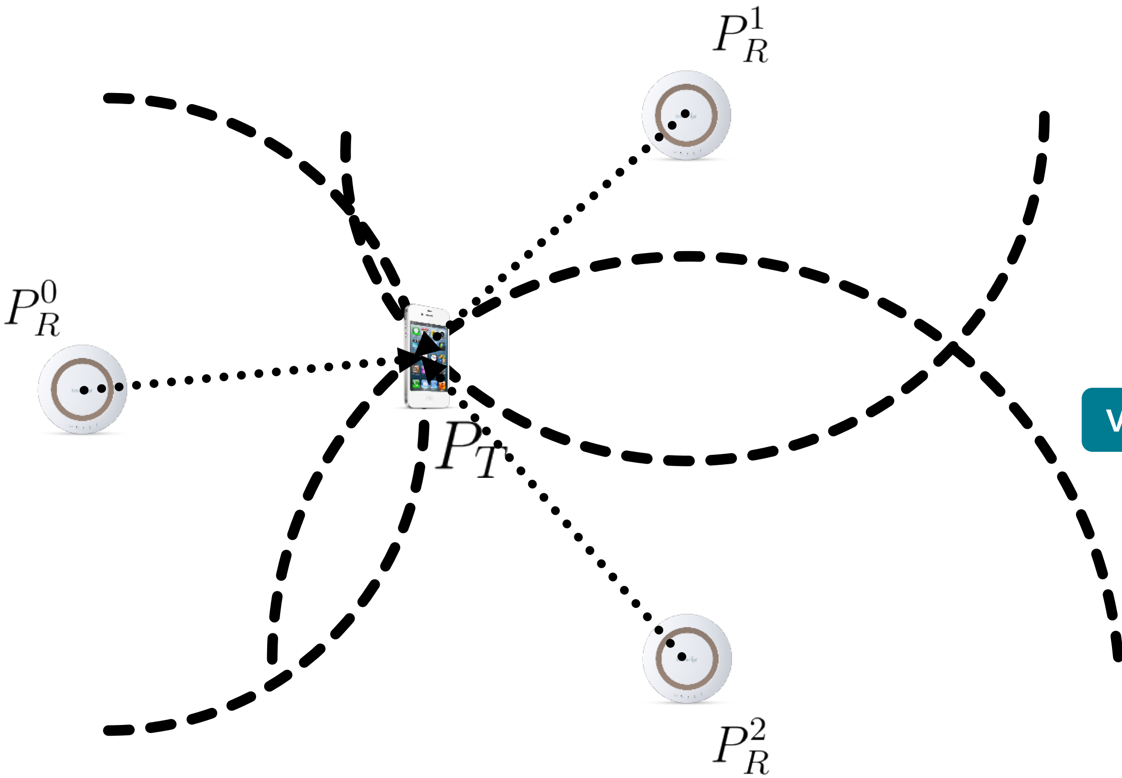
$$\boxed{r_i^2 = \vec{r}_i \cdot \vec{r}_i}$$

$$\vec{r}_i = \vec{x}_T - \vec{x}_R^i$$

Coordinates of the receivers are known. The coordinate of the transmitter is unknown. This is what we want to determine

Solve three equations simultaneously to determine intersection of circles, i.e. coordinates of the transmitter.

χ^2 minimisation aka GLS method I



Measurement model, e.g. Friis equation, which gives signal strength for a given set of parameters p.

i-th measured signal strength

Vector of residuals

$$r_i = m_i - h_i(\vec{p})$$

$$\chi^2 = \vec{r}^T V^{-1} \vec{r}$$

Covariance matrix of measurements, i.e. measurement resolution. For independent measurements

$$V = \sigma_m^2 I$$

Goal is to find the optimal set of parameters p, e.g. the coordinates of the transmitter, that minimises the χ^2 …

χ^2 minimisation aka GLS method II (Linear Measurement Model)

N.B. Linear in the parameters that we want to determine, i.e.

$$\frac{d^2 h(\vec{p})}{d\vec{p}^2} = 0$$

Measurement model can linearised using Taylor expansion about a suitable point

Goal is to find optimal set of parameters p, e.g. the coordinates of the transmitter, that minimises the χ^2, i.e.

$$\frac{d\chi^2}{d\vec{p}} \equiv 0$$

Estimated parameters are given by LSE

$$\hat{p} = (H^T V^{-1} H)^{-1} H^T V^{-1} \vec{m}$$

NxM (Projection or Design) matrix

$$H = \begin{pmatrix} \frac{\partial h_1}{\partial p_1} & \cdots & \frac{\partial h_1}{\partial p_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_N}{\partial p_1} & \cdots & \frac{\partial h_N}{\partial p_M} \end{pmatrix}$$

N measurements

M parameters

MxM Covariance matrix of parameters

$$C \equiv Cov(\hat{p}) = (H^T V^{-1} H)^{-1}$$

$H^T V^{-1} H$ is not invertible when:

- More parameters than measurements, *i.e.* the system is underdetermined.
- Parameters can be expressed as (linear) combinations of other parameters, so-called weak modes.

χ^2 minimisation aka GLS method III (Non – Linear Measurement Model)

Friis model is not linear in the parameters that we want to determine

$$P_r = \rho - 10 \log r(x_t, y_t)^2$$

Use Newton – Raphson

$$\vec{p}_{i+1} = \vec{p}_i - \frac{f(\vec{p})}{f'(\vec{p})}$$

to find roots of

$$f(\vec{p}) = \frac{d\chi^2}{d\vec{p}}$$

iteratively.

Estimated parameters are given by

$$\hat{p} = \vec{p}_0 - \left[\left. \frac{d^2\chi^2}{d\vec{p}^2} \right|_{\vec{p}_0} \right]^{-1} \left. \frac{d\chi^2}{d\vec{p}} \right|_{\vec{p}_0}$$

where

$$\frac{1}{2}\frac{d\chi^2}{d\vec{p}} = H^T V^{1} \vec{r}$$

and

$$\frac{1}{2}\frac{d^2\chi^2}{d\vec{p}^2} = H^T V^{-1} H$$

Note

$$C = 2\left[\frac{d^2\chi^2}{d\vec{p}^2}\right]^{-1}$$

Ignore higher order terms, which may introduce instabilities

Master equations and algorithm

$$\vec{p}_0$$

Calculate derivatives

$$\frac{1}{2}\frac{d\chi^2}{d\vec{p}}\bigg|_{\vec{p}_o} \text{ and } \frac{1}{2}\frac{d^2\chi^2}{d\vec{p}^2}\bigg|_{\vec{p}_o}$$

Solve

$$\frac{1}{2}\frac{d^2\chi^2}{d\vec{p}^2}\bigg|_{\vec{p}_o}\delta\vec{p} = \frac{1}{2}\frac{d\chi^2}{d\vec{p}}\bigg|_{\vec{p}_o}$$

Change in χ2

$$\Delta\chi^2 = -\frac{1}{2}\Delta\vec{p}^T\left[\frac{1}{2}\frac{d^2\chi^2}{d\vec{p}^2}\bigg|_{\vec{p}_o}\right]\Delta\vec{p} < 1$$

Yes

Done

No

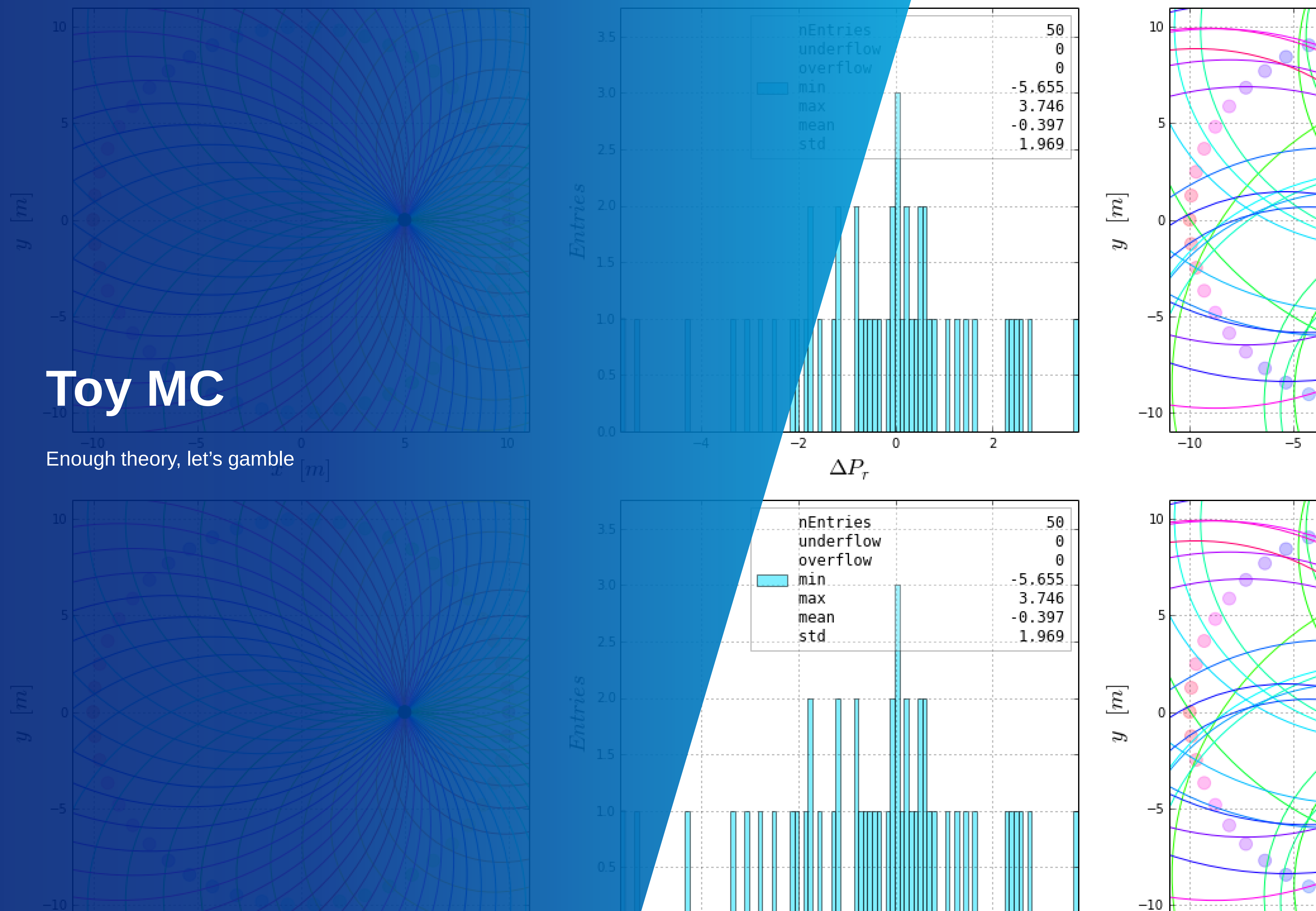
Update

$$\vec{p}_{n+1} = \vec{p}_0 + \delta p$$

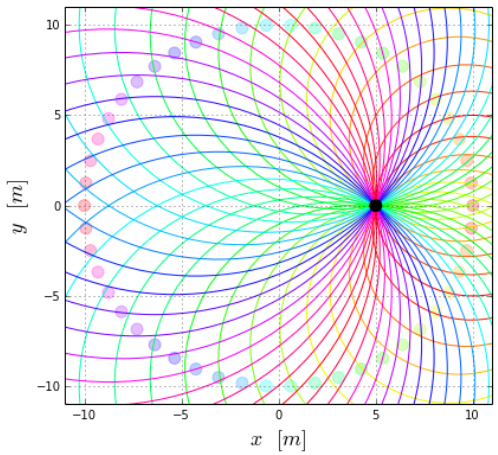
- Suitable expansion point is, e.g. centre – of – gravity
- Can also use a combination of gradient descent and Newton – Rapshon. In essence a regularisation of the covariance matrix to ensure stability

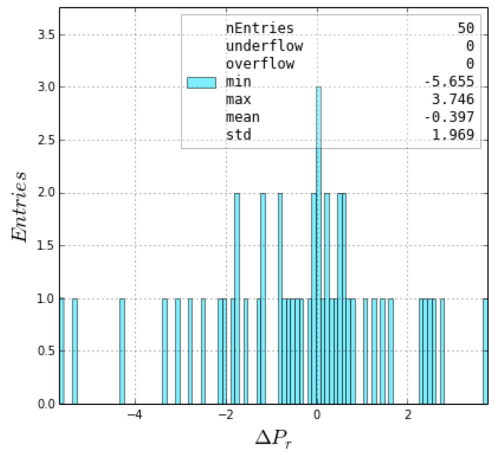
Toy MC

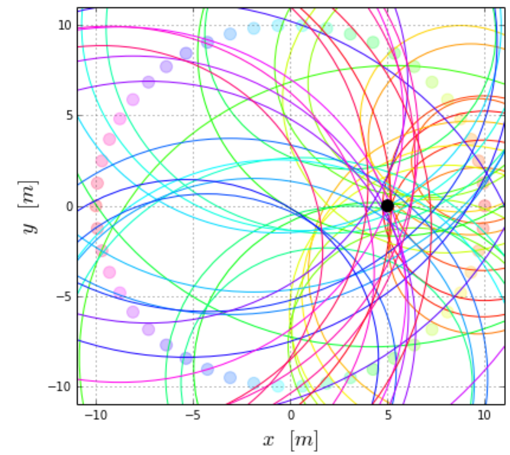
Enough theory, let's gamble



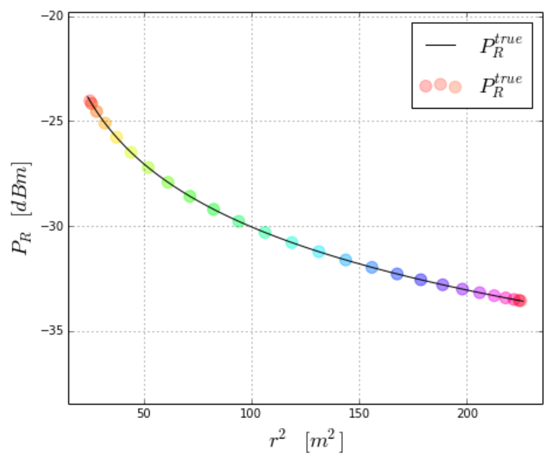
Toy MC
1 Transmitter and 50 Receivers

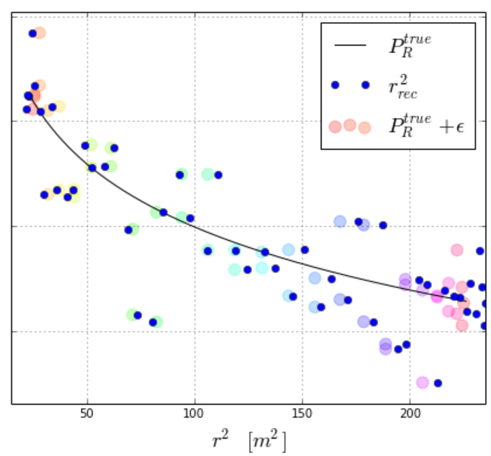






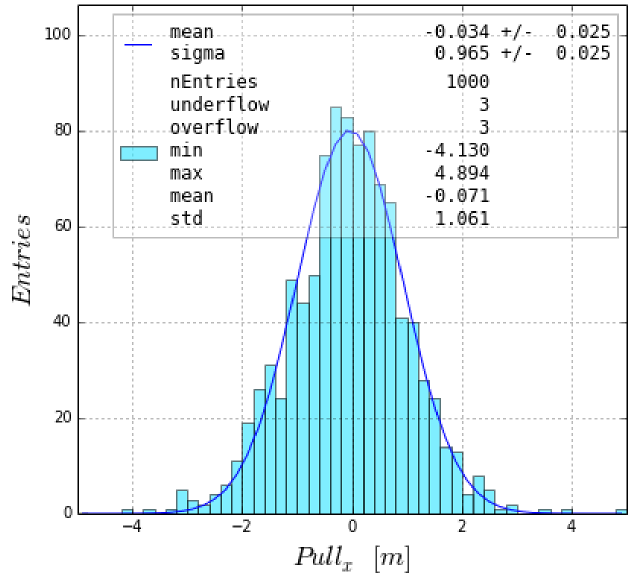
Smear signals with a measurement resolution of 2 dBm

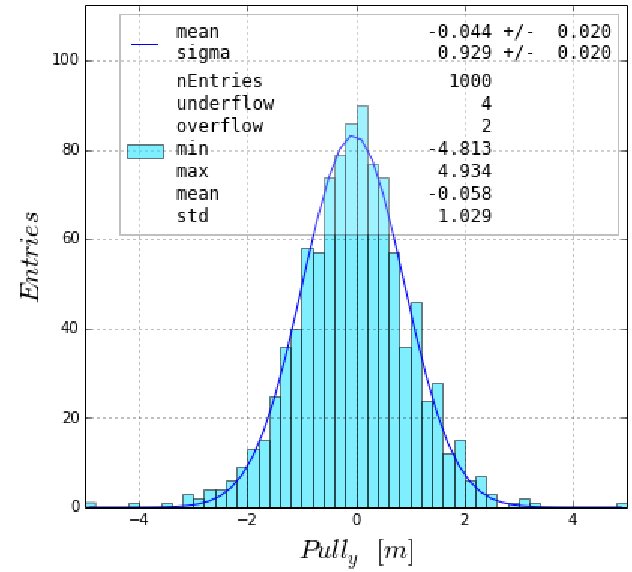




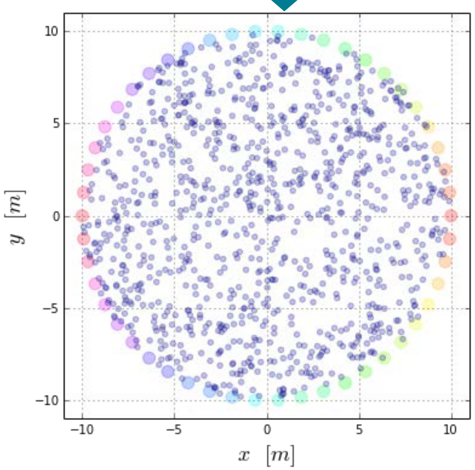
- Line – of – sight detection
- Only source of “noise” is the precision with which the sensors can measure a signal, i.e. measurement resolution
- Ignoring obstacles, random transmissions, dead time, faulty detectors, mis – alignments of recievers
- Sole purpose is to test the algorithm

Toy MC
Pull distributions





1000 randomly distributed transmitters and 50 receivers

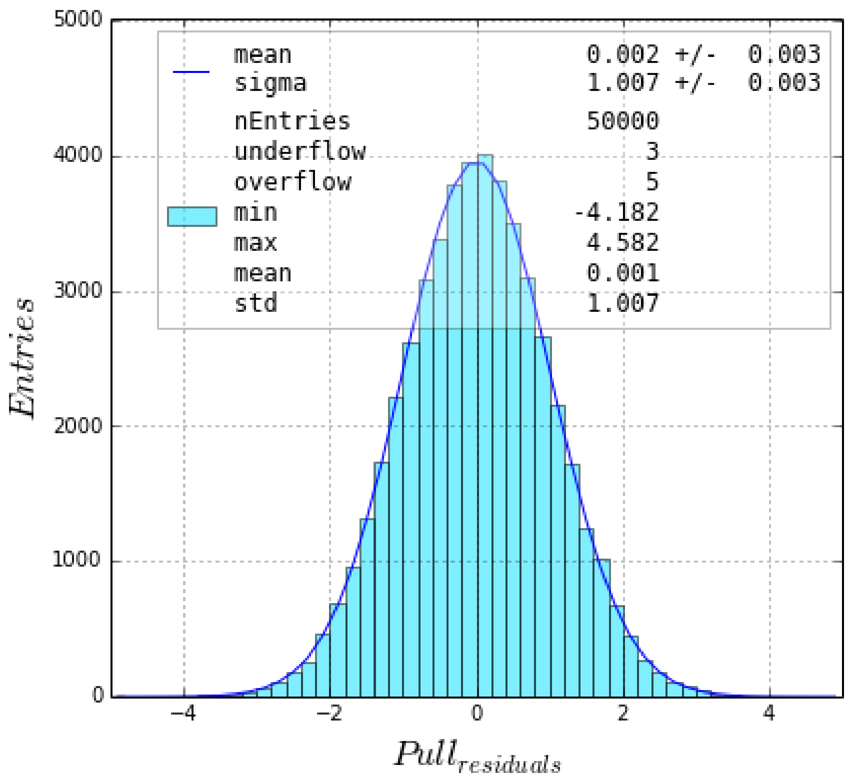


$$pull(p) = \frac{p_{rec} - p_{true}}{\sigma_{p_{rec}}}$$

- Mean should be 0 and sigma should be 1
- Non-zero mean indicates there is a bias
- Sigma > 1 indicates that the errors on the parameters are underestimated
- Sigma < 1 indicates that the errors on the parameters are overestimated

Toy MC

Pull of residuals



$$\vec{r} = \vec{m} - \vec{h}(\vec{p})$$

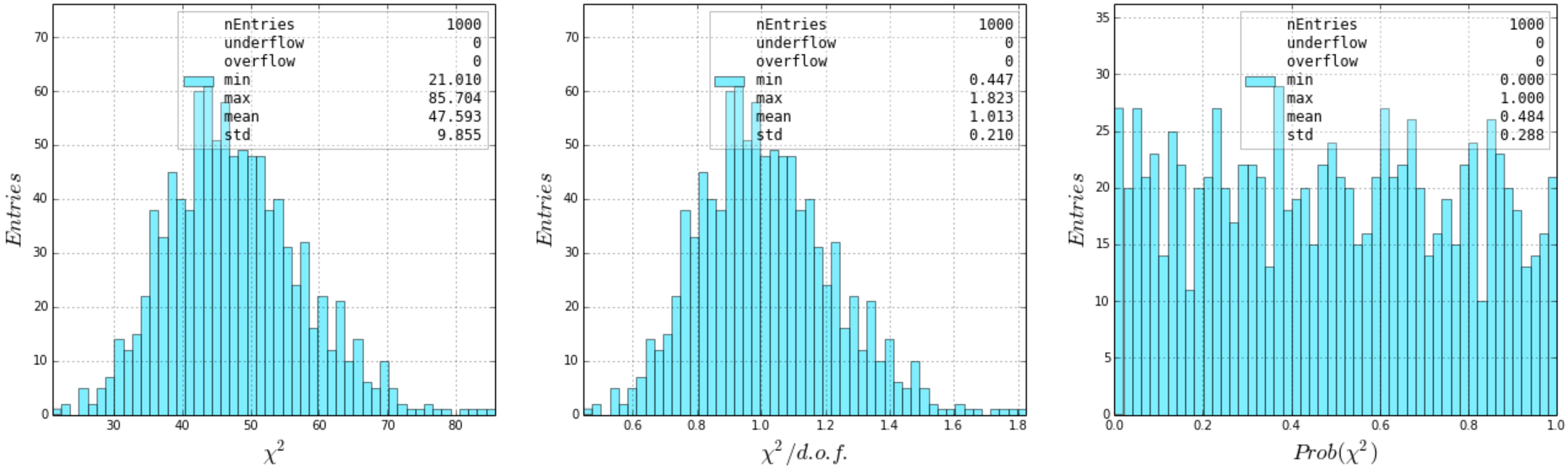
Correlated with covariance of parameters

$$Cov(\vec{r}) \equiv R = V - HCH^T$$

Covariance of parameters in measurement space

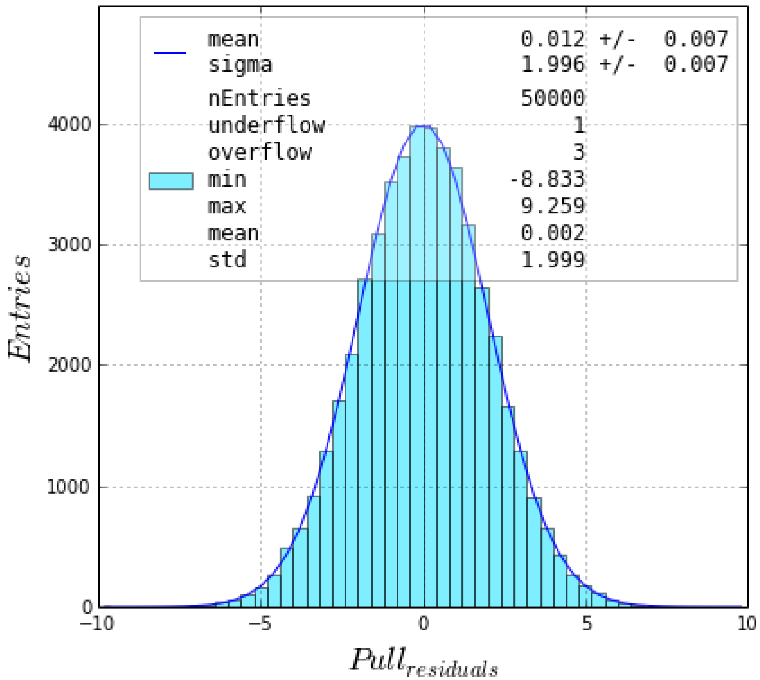
- In reality we don’t have any truth information
- Look at pull of residuals
- Mean should be 0 and sigma should be 1
- Non-zero mean indicates there is a bias
- Sigma > 1 indicates that the errors on the parameters are underestimated
- Sigma < 1 indicates that the errors on the parameters are overestimated

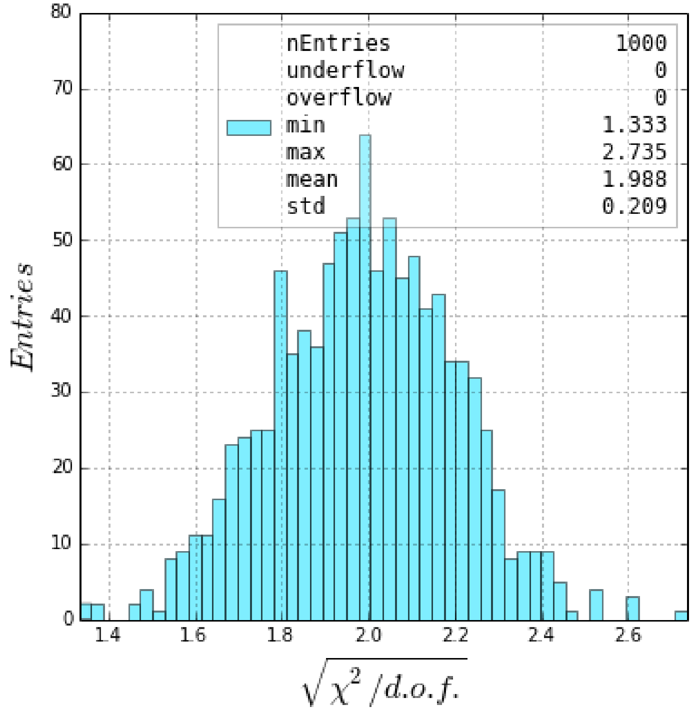
Toy MC
χ^2 goodness of fit



- E[χ^2] gives the number of degree of freedom: 50 measurements – 3 parameters = 47
- E[χ^2/d.o.f.] = 1
- Prop(χ^2) is the probability of finding a χ^2 that is equal or worse than this χ^2. Indicates how well our model describes the data.
  - Peak at zero means that our model describes the data “poorly”
  - Peak at one means that our model is too “good”

Estimating the measurement resolution





- Assume we do not know the measurement resolution
- Assuming it is the only source of noise and our model is correct:
  - We can estimate it from the width of the pull of the residuals
  - Or from $E[\chi^2/\text{d.o.f.}]$
- In this example the estimated measurement resolution is 2 dBm

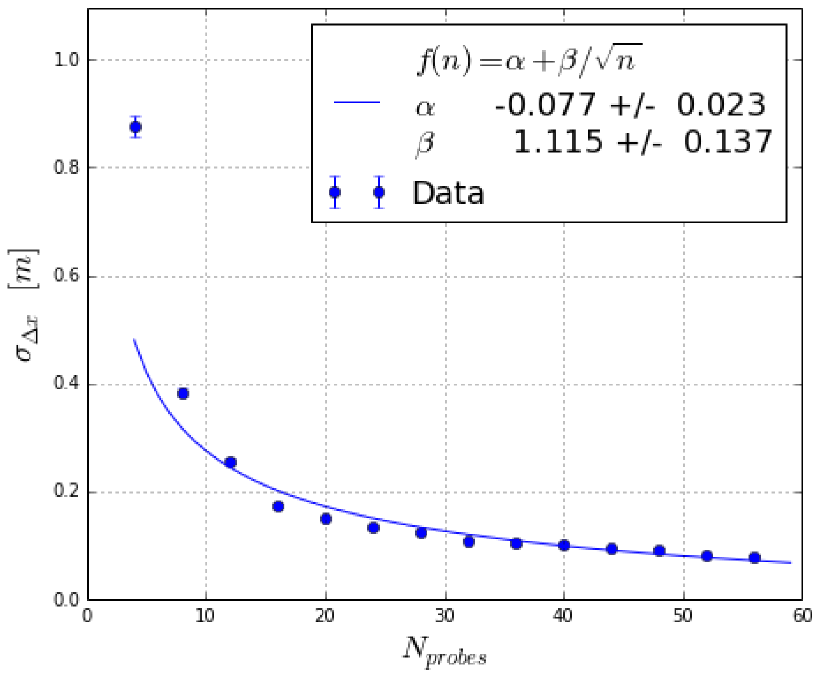
$$\hat{\sigma}_m = \sqrt{\frac{\chi^2(\sigma_m = 1)}{\nu}}$$

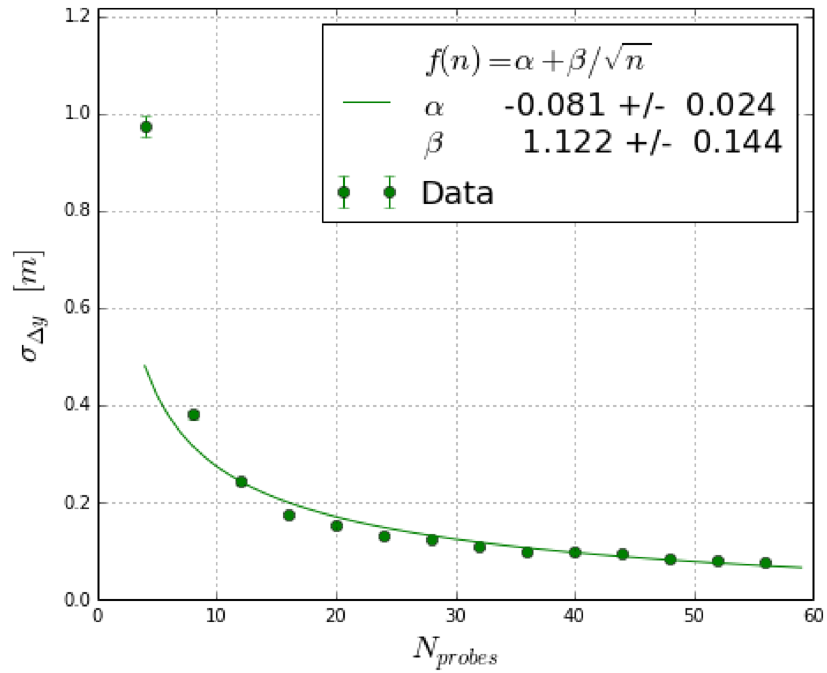
$$\nu = \text{Measurements} - \text{Parameters}$$

Number of degrees of freedom

Toy MC

Parameter resolution vs number of measurements



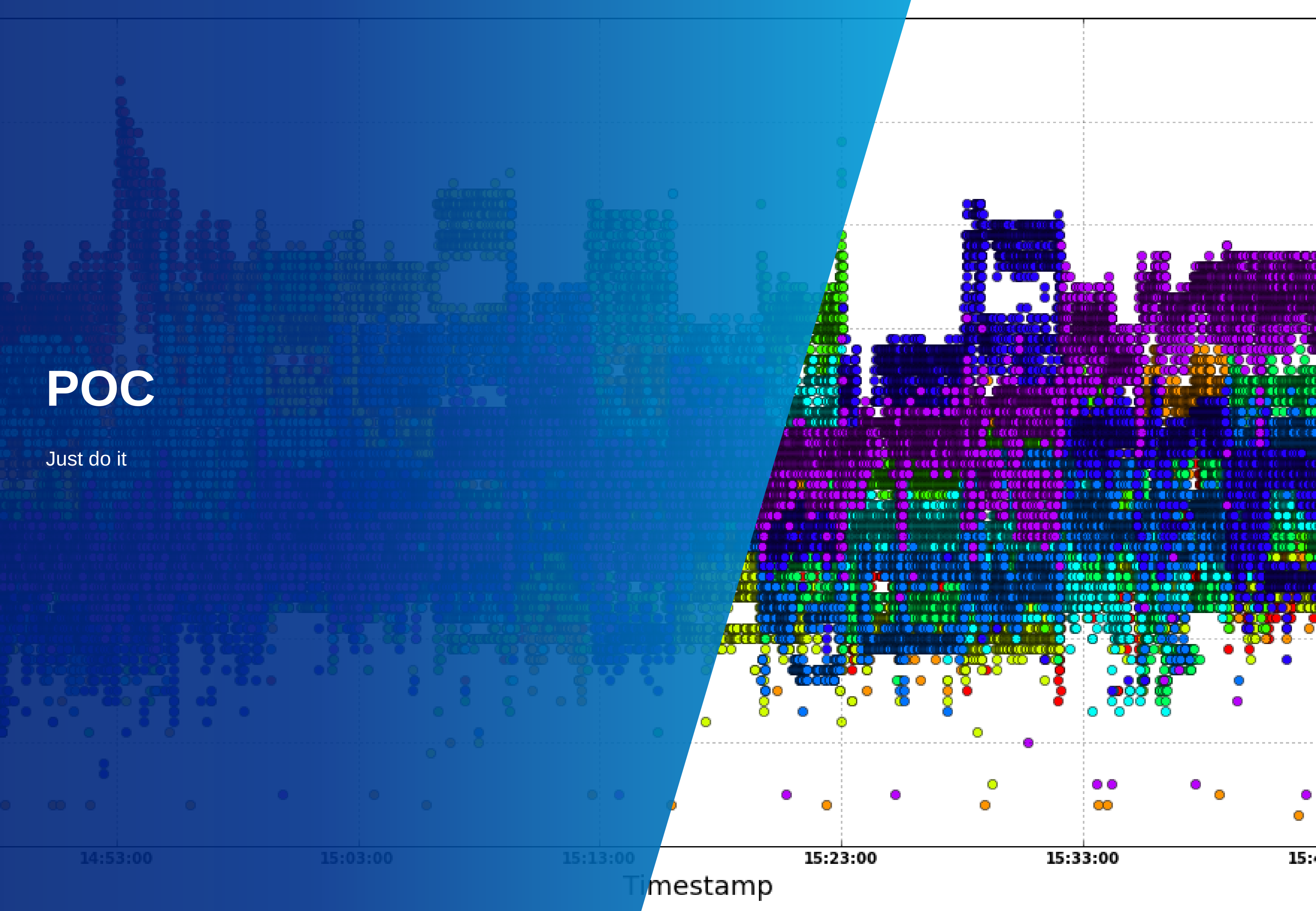


- To improve resolution by a factor 2 need 4 times as many measurements
- No significant gain beyond 16 devices
- Of course depends on:
  - Environment
  - Sensor density, i.e. number of sensors per square metre
  - Measurement resolution

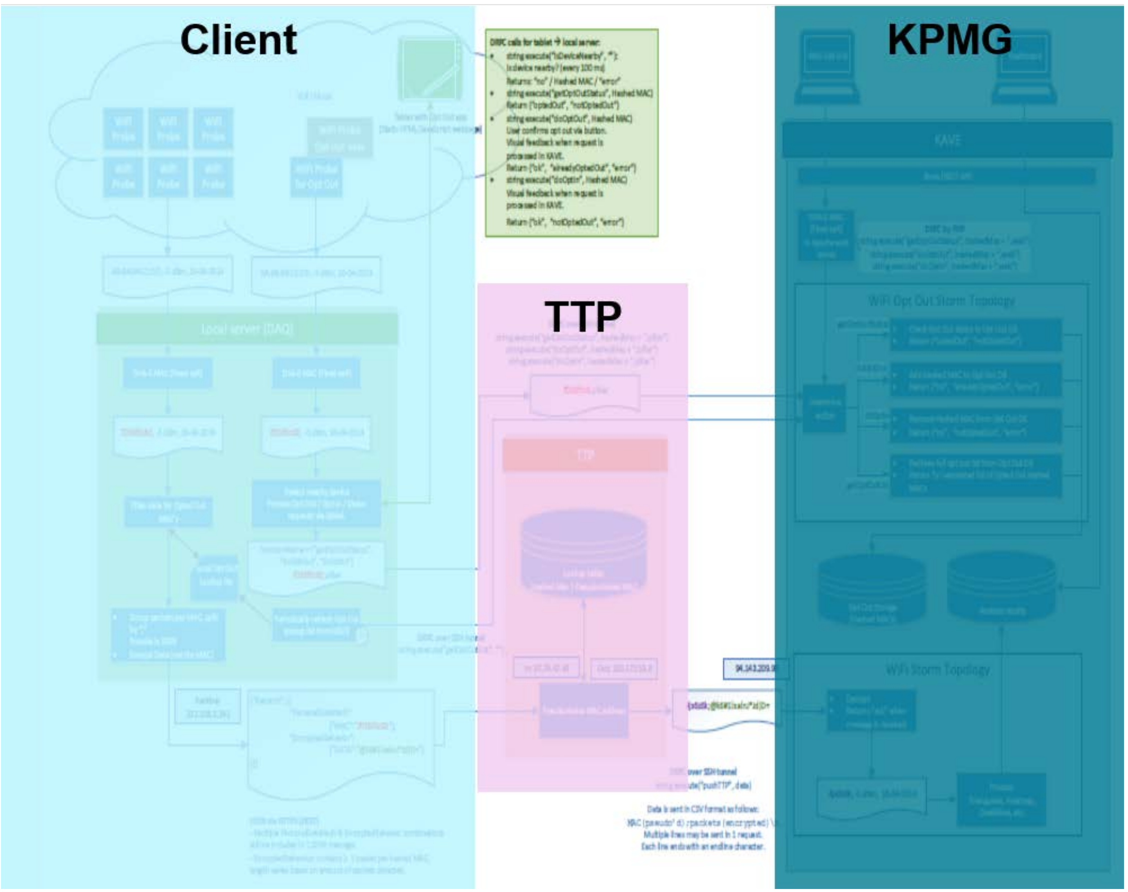
$$\sigma_p \propto \frac{1}{\sqrt{N}}$$

POC

Just do it



Platform



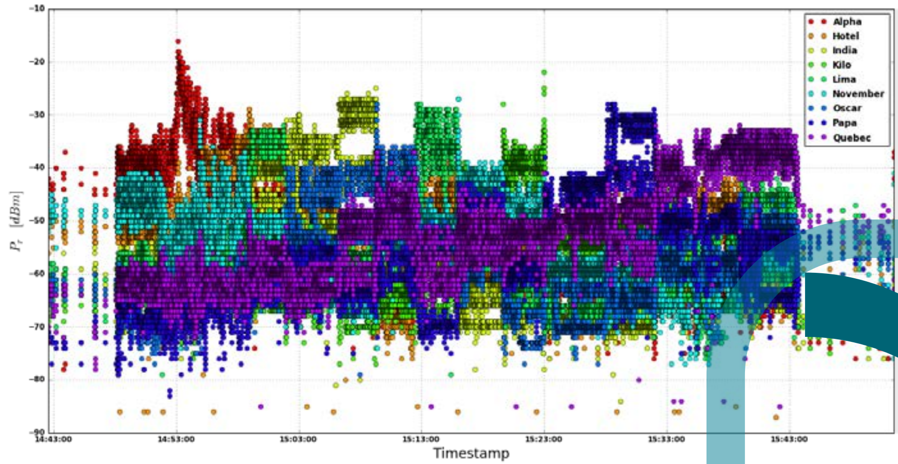
Hardware

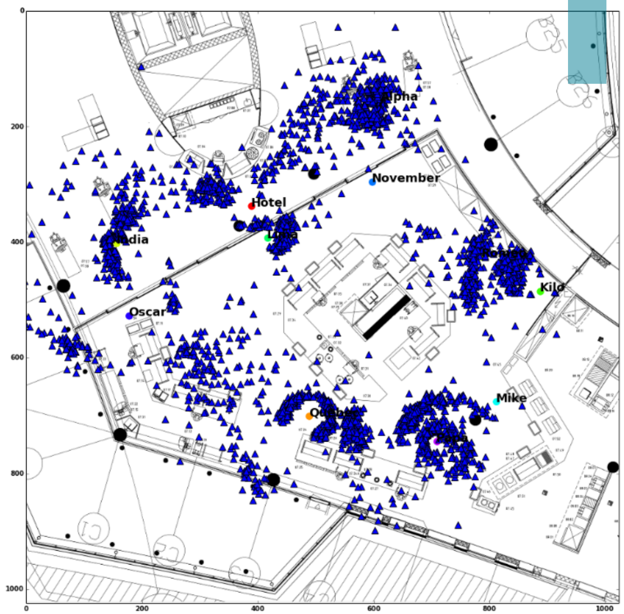
- Nothing fancy, of the shelf Wi-Fi routers
- Promiscuous mode, i.e. they only listen and do not communicate with Wi-Fi enabled devices. They are passive
- DAQ and Opt – out servers
- iPad opt – out pillars
- TTP, Trusted Third Party. Not really hardware … For anonymisation of hardware.

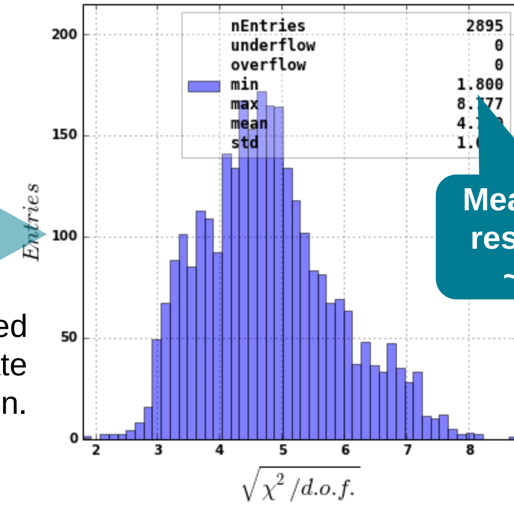
Software/Application Stack

- KAVE (KPMG Analytics and Visualisation Environment)
- Storm for real time processing
- Mongo and Hadoop for data storage, latter is used in batch processing
- Collection of algorithms/analyses written in Java and Python

POC
A calibration run



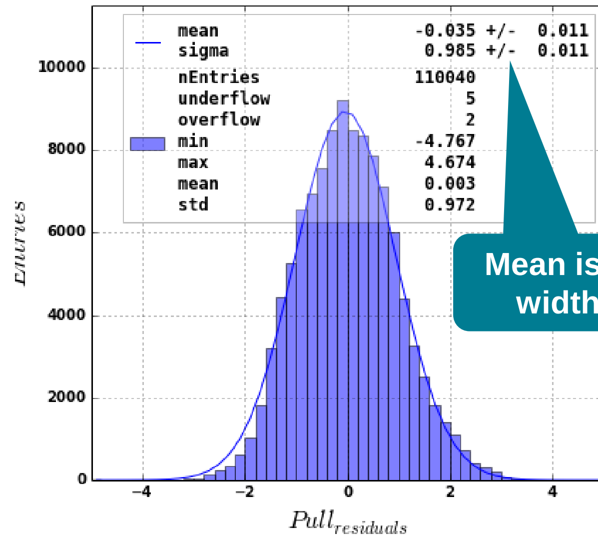




Measurement resolution is ~ 5 dBm

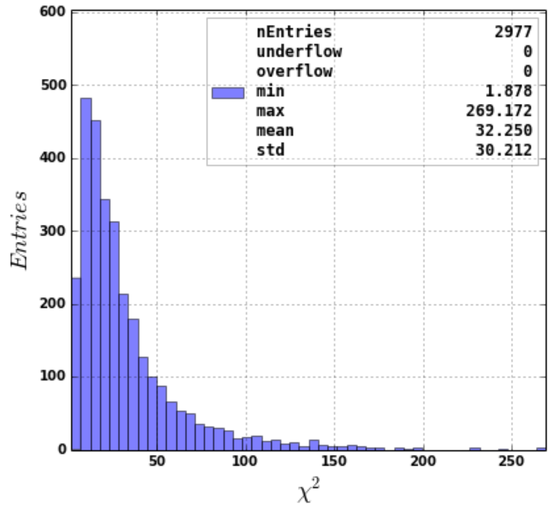
Use reconstructed events to estimate resolution.

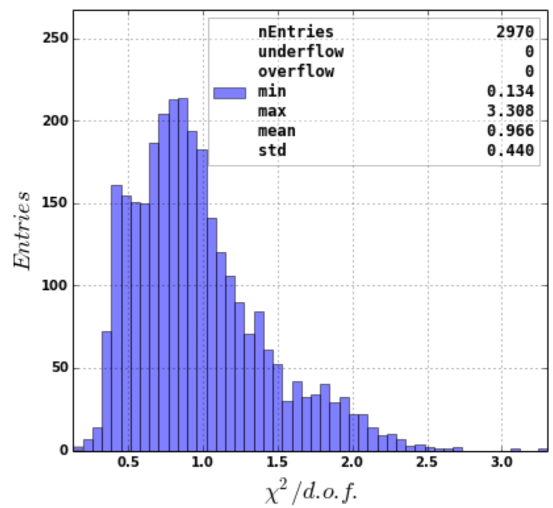
Use measurements to reconstruct coordinates of device

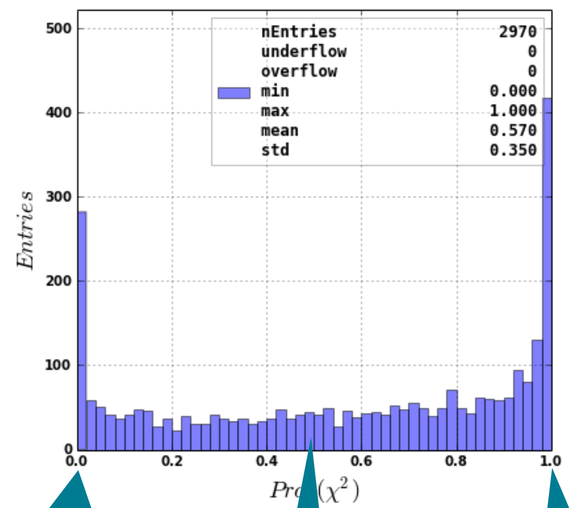


Mean is ~ 0 and width is ~ 1

A calibration run - χ^2 goodness of fit







Model doesn't describe data well

Too well …

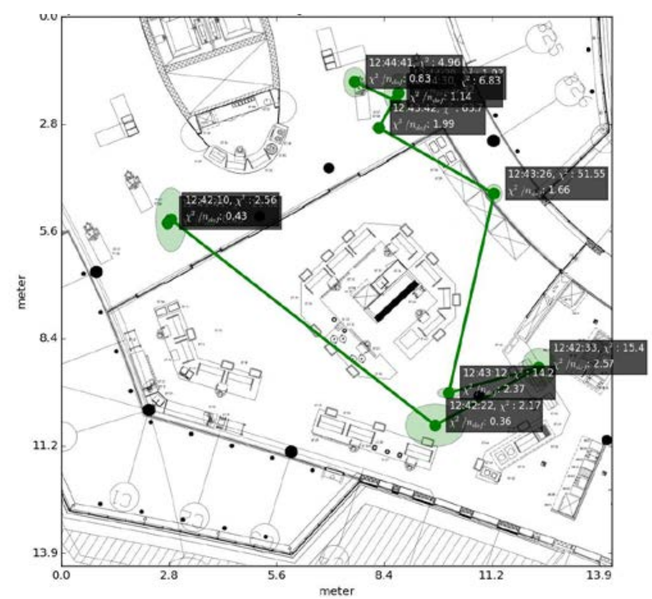
Uniform, OK

Possible reasons why fit is sometimes "off":

- Assume line – of – sight measurements, i.e. ignore scattering, reflections. These can lead to destructive/constructive interference
- Some sensors are more equal than others
- Sensors are mis-aligned, i.e. their coordinates are "off"
- Directionality of the sensors
- …

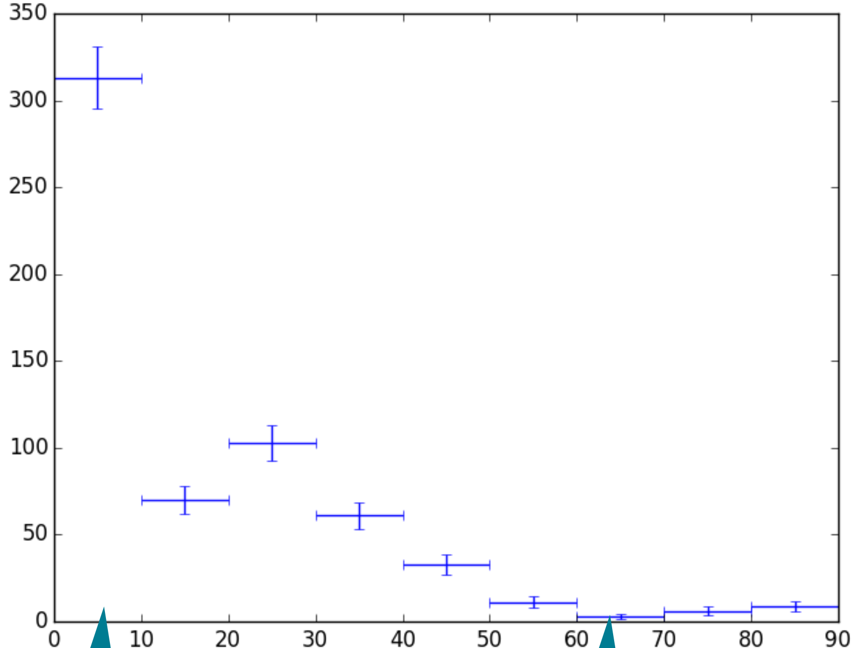
POC

Heat map – Time for lunch or one more quick email?



POC
Some “KPI’s”

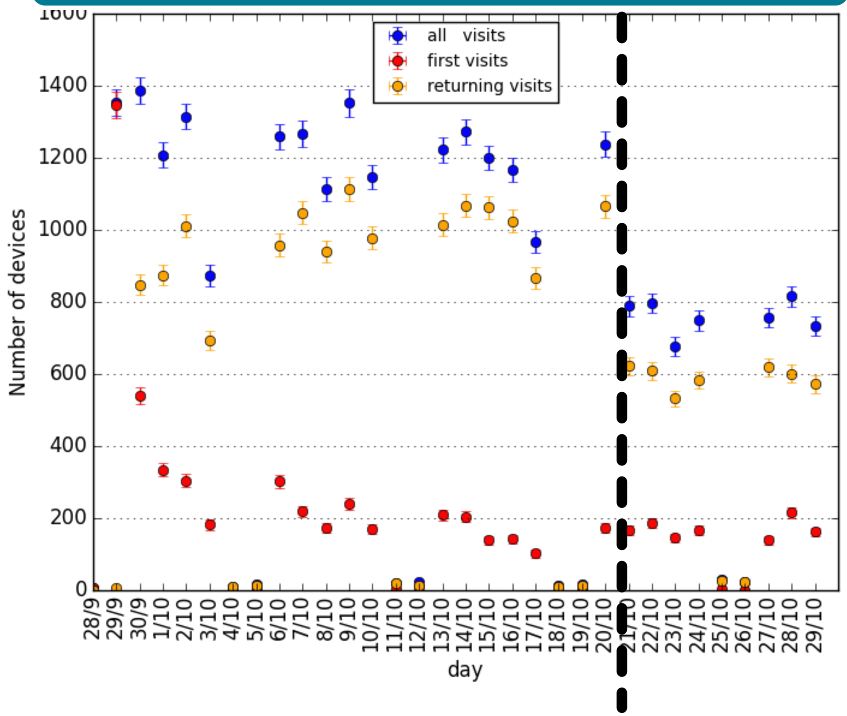
Dwell time; difference between t_{first} and t_{last} seen



Passing through

Lingering

Number of visits per day



Approx. factor 2 drop in number of visits. Were also reconstructing “repeater” packets before, i.e. “double counting”

Summary

The all good things must come an end

The end is neigh

- It works, but room for improvement:
  - Calibration of devices
  - Hit selection/collection, e.g. signal strength, coincidences
  - Outlier removal/refitting
  - Tuning/calibration of KPI's
- Next steps:
  - Kalman Filter Fitter, same stuff they use to track the Space Shuttle
  - Crowd prediction models/algorithms

cutting through complexity