

→ 20th NKG General Assembly, Vilnius, Lithuania, 2026

→ 14-17 September 2026

Outlier detection and quality control – paradoxes with some existing approaches and suggested alternatives.

Ola Øvstedal, Department of geomatics – NMBU (presenter)

Narve Kjørsvik, Field Geospatial

Jon Glenn Gjevestad, Department of geomatics – NMBU

Outlook

- What is the problem?
- What is meant with «an observation»?
- Testing for outliers in observations
- Paradoxes with “a single-error, scalar one-dimensional approach”
- A suggested alternative approach
 - Testing of observations
 - Internal and external reliability

What is the problem?

Many geomatic observations are multi dimensional by nature, e.g.

- derived GNSS observations like baselines and RTK point observations (x,y,z)
- photogrammetric image observations (x,y)

In testing of observations however, a “single-error, one-dimensional scalar approach” is still commonly used, leading to some paradoxes when applied also for multi dimensional observations.

LSQ Adjustment – the working horse

Used in numerous geomatic applications

- geodetic networks
- cadastral surveys
- photogrammetric ground control points (GCP)
- etc.

Parameter estimation and analysis

- LSQ adjustment of a redundant system of observations
- Quality control
 - Statistical testing of e.g. observations
 - Calculate quality numbers in the form of e.g. internal and external reliability

Quality control – the DIA concept

Detection of gross errors.

How well do the observations control each other (Internal reliability).

What is the effect of any undetected errors on estimated parameters (External reliability).



LSQ Adjustment using the parametric method; what is an observation?

The observations are expressed as functions of unknown parameters.

Observations are e.g.

- 1-D: distance, directions, zenith angles from total stations, height differences from levelling.
- 2-D: image observations in photogrammetry.
- 3-D: GNSS RTK point observations, GNSS baselines.

Derived GNSS observations are **multidimensional** and **indivisible**.

- E.g. x-component does not exist without information on y- and z- component.
- Reflected by fully occupied covariance matrix for precision of observations.



A multidimensional observation is frequently denoted “a tuple”.

Test of observations for outliers

- Requires redundant observation with internal control.
- Shall not be influenced by external constraint from e.g. existing control points.
- Normally carried out iteratively, following the standard “goodness-of-fit” test.
 - For every “observation” a hypothesis is tested:
→ H_0 : no gross-error against H_1 : gross error
 - For each iteration the most probable gross error is identified, the corresponding observation is removed if most probable gross error is significant.
 - The search for gross errors ends when no alternative hypotheses are significant.

Multiple t-testing – commonly used in Norway

Subsequently testing each observation, $i = 1, 2, \dots, n$:

- Estimate possible “gross error” as additional nuisance parameter
 - ∇
- Compute estimated standard deviation of unit weight,
 - $s_0 = \sqrt{\frac{v^T P v}{(n-e-1)}}$, degrees-of-freedom $f = (n - e - 1)$
- Compute standard deviation of estimated ∇
 - $s_{\nabla} = s_0 \sqrt{q_{\nabla\nabla}}$
- Compute test statistic
 - $t = \left| \frac{\nabla}{s_{\nabla}} \right|$

The test statistic t is T-distributed with f degrees of freedom and adopted significance level (adjusted for multiple testing).

The observation with largest computed t is considered to be the most probable gross error and is checked against threshold value from the T-distribution.



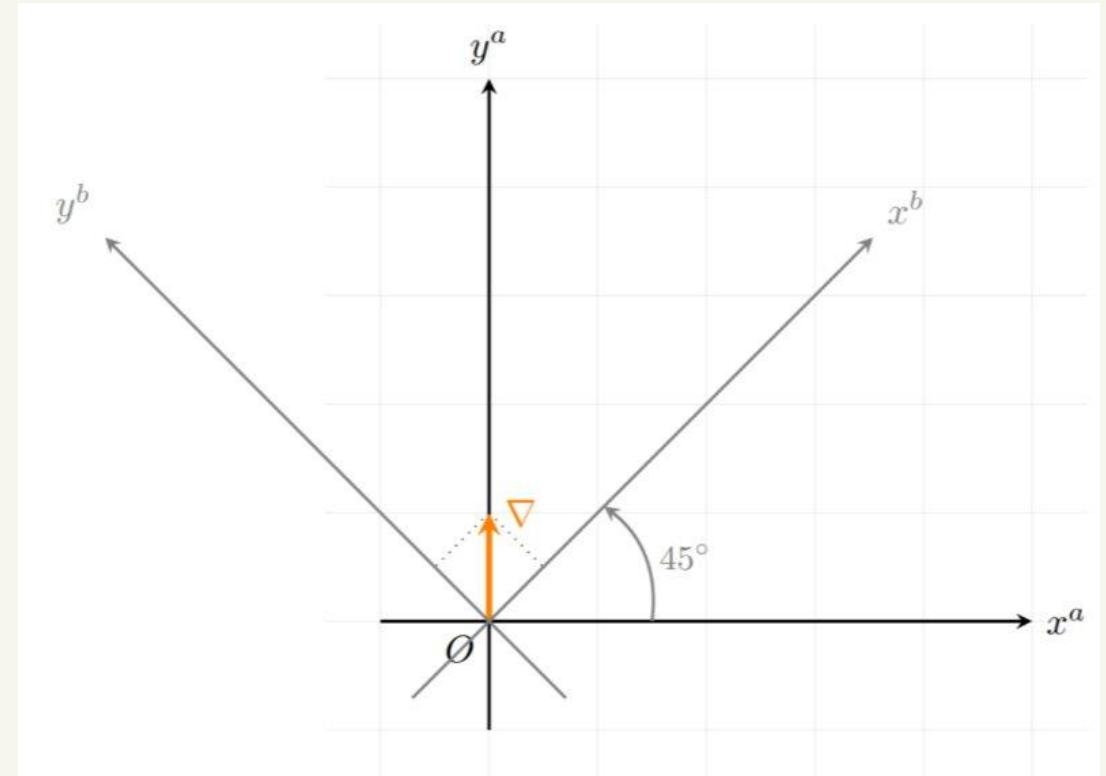
In the more commonly used data-snooping approach , a priory σ_0 is replacing the estimated s_0 .

Paradox I with one-dimensional approach

- the outcome depends on orientation of coordinate system-

Frequent approach:

- Scalar one-dimensional hypothesis.
- Assumes 1D-observations, also for tuples of observations.
- Example with 1:1-relationship between observations and unknowns.
- Estimation and quality control in given map projection.



Theoretical 2D-example with most unfavorable orientation

Real example: GNSS RTK-observations tested for outliers using multiple t-testing assuming one-dimensional coordinates. Successive rotation of coordinate system.

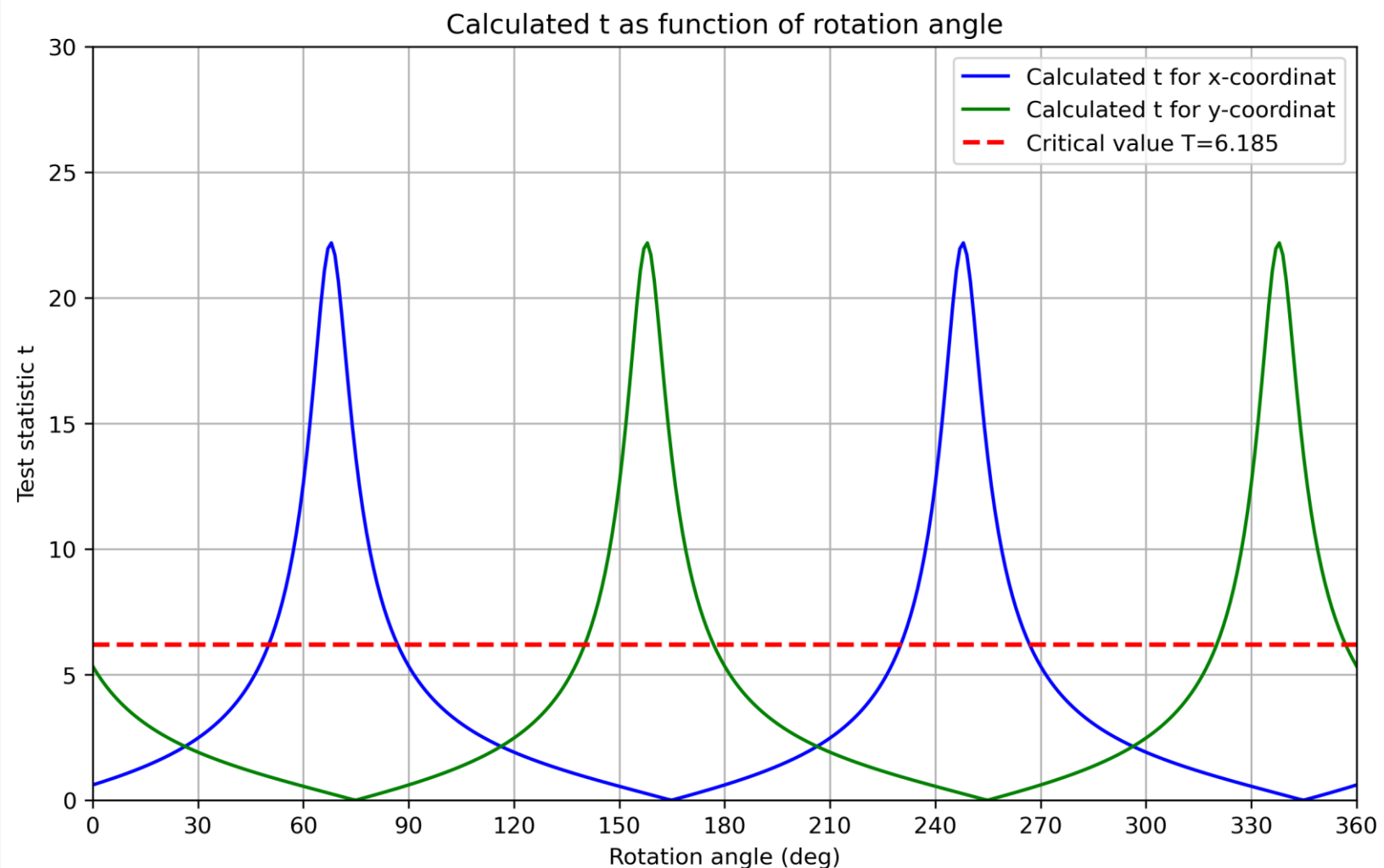
Analysis in 2D only.

GNSS points observations (x and y) carried out in one station. Demonstrated using one of total three point observations.

For each one-dimensional observation to be tested (x or y), test statistic t is calculated as:

$$t = \left| \frac{\bar{v}}{s_{\bar{v}}} \right|$$

1. Initially no significant gross error.
2. Successive rotation of coordinate system demonstrates that the outcome depends on choice of coordinate system



Paradox I with scalar one-dimensional approach

- For derived GNSS observations, the outcome of gross error detection depends on choice of coordinate system.
 - If carried out in a map projection like UTM, dependency of zone/central meridian.
 - Different results if analysis is carried out in earth centered or topocentric system.
 -
- Internal and external reliability based on the same one-dimensional hypothesis also gives different results for different coordinate systems.
- Violates “established professional practice for surveyors”?

Paradox II with:

- scalar one-dimensional approach
- multiple t-testing

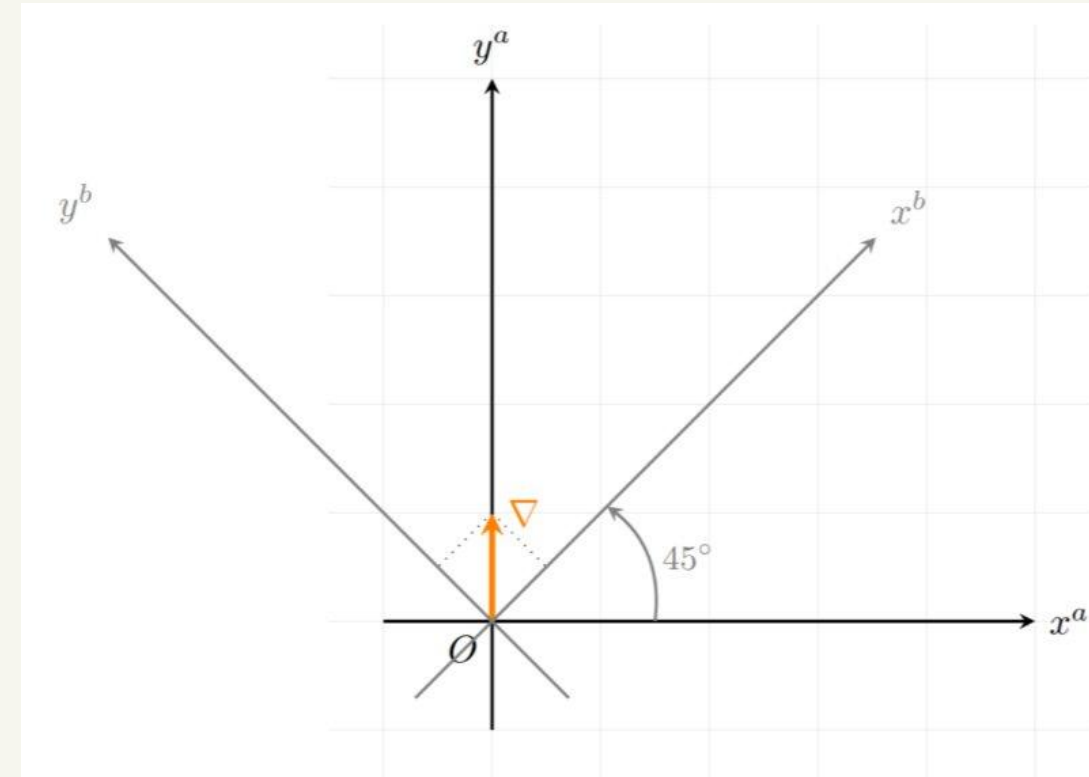
In multiple t-testing, the test statistic is calculated as:

$$t = \left| \frac{\nabla}{s_{\nabla}} \right|, \text{ where } s_{\nabla} = s_0 \sqrt{q_{\nabla \nabla}}$$

With unfavorable orientation of coordinate system, it can be shown that:

$$t \rightarrow \text{constant as } \nabla \rightarrow \infty$$

- Due to rotation of coordinate system, un-absorbed part of ∇ will contribute to estimated s_0 .



Theoretical 2D-example with most unfavorable orientation

Adaption

- as already implemented in some operational software

- When the observations are indivisible, the hypothesis must be adapted accordingly.
- A tuple of observations must be tested as a whole.
- The F-test replaces the t-test (identical result for 1D-hypothesis).
- The F-test uses the **statistical length** of the gross error as statistics: Invariant for choice of coordinate system.

(If σ_0 is assumed known, the chi-square test replaces the F-test.)

Critical value for the F-test

Note:

x, y and z in observation domain, not necessarily coordinates.

a, b and c are constants.

1D: Confidence interval: $\pm x$ gives the same test value

- $x^2/a^2 = 1$

2D: Confidence ellipse : constant test value along the edge of the ellipse

- $(x^2/a^2 + y^2/b^2 = 1)$

3D: Confidence ellipsoid : constant test value along the surface of the ellipsoid

- $(x^2/a^2 + y^2/b^2 + z^2/c^2 = 1)$

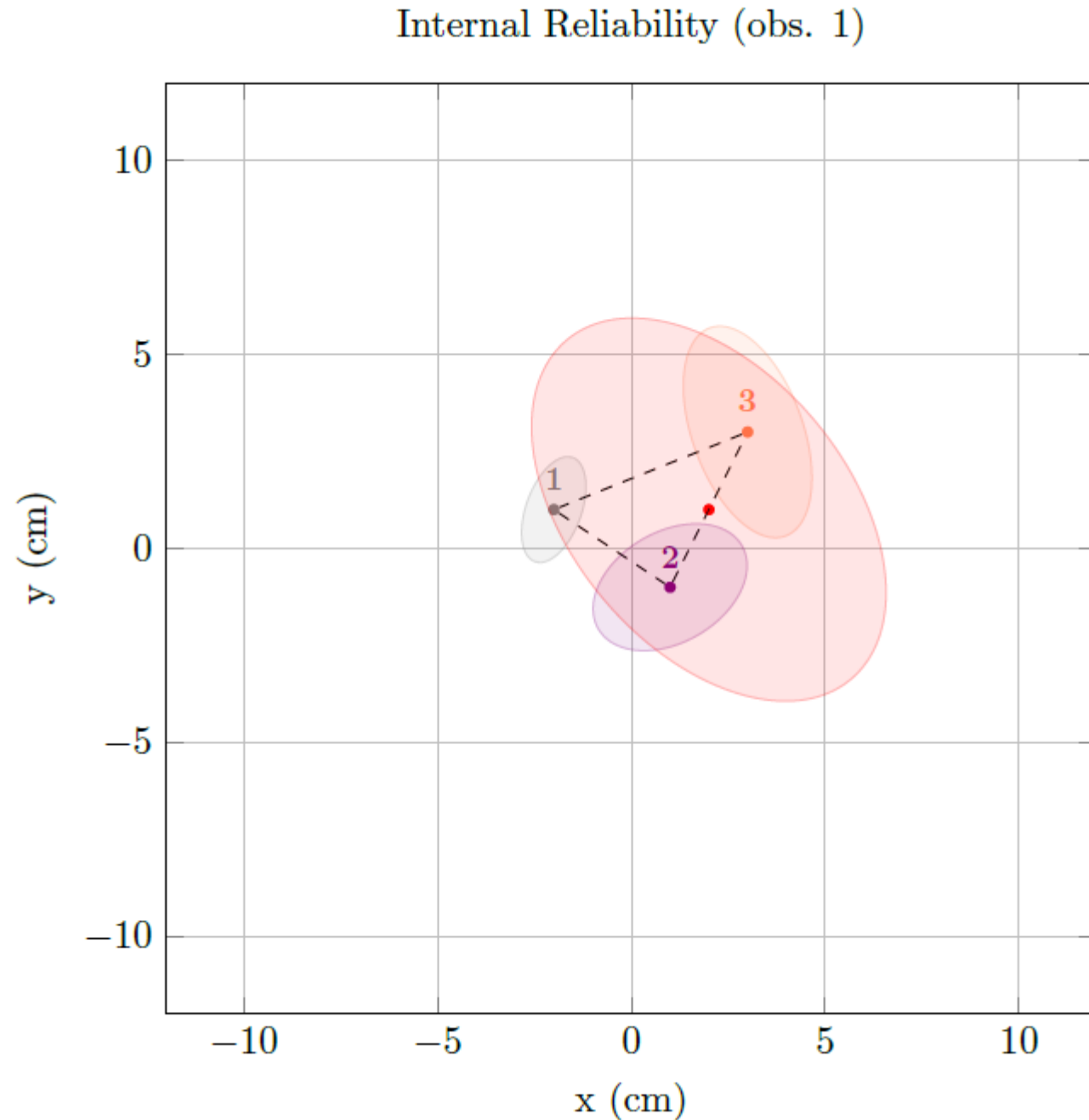
nD: Confidence hyperellipsoid: constant test value along the “surface” of the hyperellipsoid

- $(x^2/a^2 + y^2/b^2 + z^2/c^2 + \dots = 1)$

Example in 2D: GNSS point observations, test of obs. 1

Comment:

The shape of the rejection region is determined by how well measurement 1 is controlled by the other measurements, that is, by both the geometry and the measurement precision.



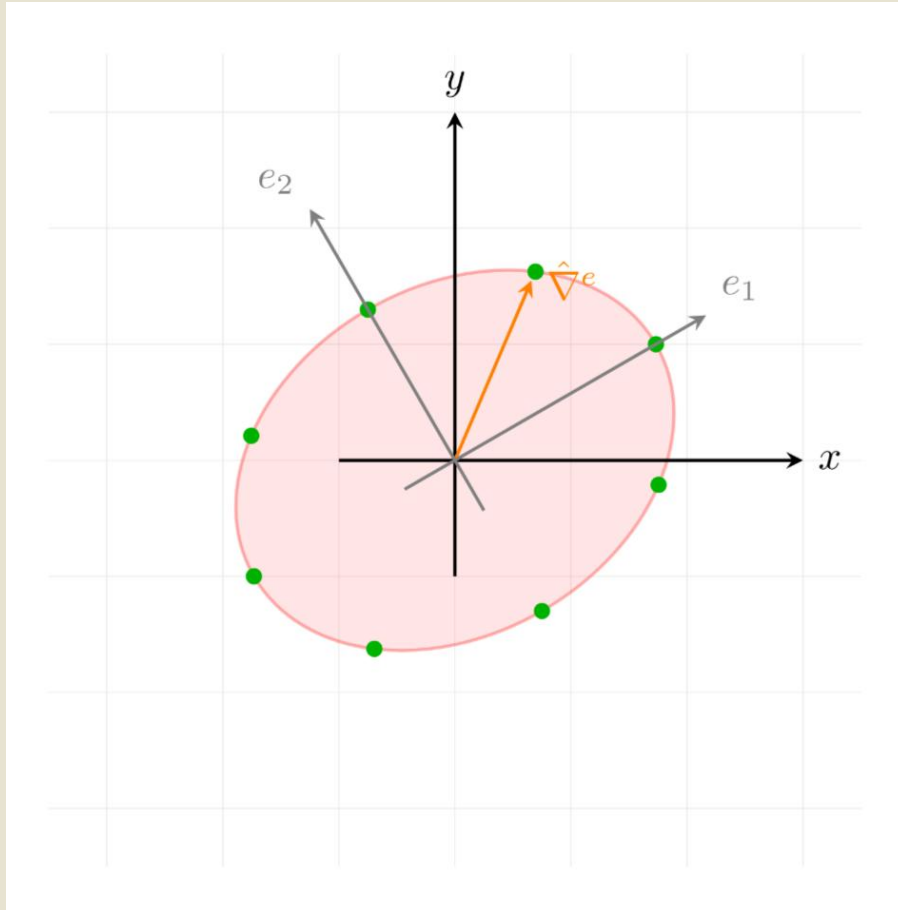
Challenge: what about External Reliability?

- External Reliability quantifies the potential impact of any undetected gross errors on the estimated parameters.
- The impact must be computed based on the same hypotheses that were used for gross error detection.
- A computed number for External Reliability has no universal meaning.
 - The computed number is only valid for the chosen hypothesis.
 - Several hypotheses are both possible and meaningful.

Computation of External Reliability?

- When Internal Reliability (confidence value for detection of gross error) changes from one-dimensional scalar to ellipse, ellipsoid or hyper ellipsoid, external Reliability must be computed in another way.
- In a nested process, we must evaluate all possible realizations of the one gross error that lies on the same critical boundary.
 - **Example in 2D:** The impact is evaluated for a range of points (possible gross errors) along the confidence ellipse. A representative measure is then selected, such as the maximum effect, the circumscribed circle, or another suitable value.
- The numerical evaluation can be carried out efficiently using eigenvalue decomposition and polar coordinates.
 - The same approach can be applied to gross errors of arbitrary dimension.

Trick: eigenvalue decomposition in 2D



For **each hypothesis** (2D gross error):

- Change of basis from (x,y) to (e_1, e_2)
- Distribute points representatively (in the new basis) using polar coordinates.
- Change all points back to (x,y)
- Propagate all points through the normal equations
- Save possible impact on all parameters

For **every parameter**:

- Derive External Reliability from all possible impacts.

→ E.g. maximum deformation/distortion

Example 2D RTK point-observations

External reliability is formed by the envelope of the contributions from the individual observations.

The resulting shape can be quite complex.

Shown here together with the circumscribing circle.

