Ausgleichs- und Fehlerrechnung

Dr. Uwe Stöhr

Fehlerfortpflanzung

Definitionen

- Fehlerfortpflanzung nur dann anwendbar, wenn Fehlergrößen voneinander unabhängig sind
- Unsicherheit u:

$$u = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^{N} (y_i - \bar{y})^2}$$

$\bar{y}$:= Mittelwert

y_i := Datenwert

Definitionen

- Fortpflanzung von Fehlergrenzen:

$$\Delta F = \left|\frac{\partial F}{\partial x_1}\right| \Delta x_1 + \left|\frac{\partial F}{\partial x_2}\right| \Delta x_2 + \cdots$$

- Fortpflanzung von Unsicherheiten:

$$u_F = \sqrt{\left(\frac{\partial F}{\partial x_1} u_1\right)^2 + \left(\frac{\partial F}{\partial x_2} u_2\right)^2 + \cdots}$$

ΔF := Fehler von F

Δx_1 := Fehler von x_1

Bestimmung der Kapazität

Bestimmung der Kapazität einer Schaltung

Formel für Kapazität lautet:

$$C = \frac{2L}{R^2 + \omega^2 L^2}$$

Bestimmung der Kapazität

Fehler der Kapazität:

$$\Delta C = \left| \frac{-4\,LR}{\left(L^2\omega^2 + R^2\right)^2} \right| \Delta R + \left| \frac{-4\,L^3\omega}{\left(L^2\omega^2 + R^2\right)^2} \right| \Delta\omega$$
$$+ \left| \left(\frac{2}{L^2\omega^2 + R^2} - \frac{4\,L^2\omega^2}{\left(L^2\omega^2 + R^2\right)^2} \right) \right| \Delta L$$

Bestimmung der Kapazität

Messwerte 1

R = 10 Ω
L = 10 µH
ω = 50 1/s

Annahme

$\Delta R / R$ = 1 %
$\Delta L / L$ = 10 %
$\Delta\omega / \omega$ = 0.1 %

Fehlerterme 1

R-Term = 4e-9
L-Term = 2e-8
ω-Term = 1e-18

Bestimmung der Kapazität

Messwerte 1	Messwerte 2	Annahme
R = 10 Ω	R = 10 Ω	$\Delta R / R$ = 1 %
L = 10 µH	L = 1000 µH	$\Delta L / L$ = 10 %
ω = 50 1/s	ω = 50 1/s	$\Delta\omega / \omega$ = 0.1 %

Fehlerterme 1	Fehlerterme 2
R-Term = 4e-9	R-Term = 4e-7
L-Term = 2e-8	L-Term = 2e-6
ω-Term = 1e-18	ω-Term = 1e-12

Bestimmung der Flussrate im Rohr

Formel für Flussrate lautet:

$$\Upsilon = \frac{\pi R^4 p}{8 \eta l}$$

Bestimmung der Flussrate im Rohr

Formel für Flussrate lautet:

$$\Upsilon = \frac{\pi R^4 p}{8\eta l}$$

Relativer Fehler:

$$\frac{\Delta\Upsilon}{\Upsilon} = 4\frac{\Delta R}{R} + \frac{\Delta p}{p} + \frac{\Delta\eta}{\eta} + \frac{\Delta l}{l}$$

Ausgleichsrechnung

Definitionen

- Regression
  Untersuchung der Korrelation von Datenpunkten ohne Messfehler mit angenommenen Zusammenhang

- Fit
  wie Regression, jedoch mit Berücksichtigung von Messfehlern

Bestimmtheitsmaß R^2

- Beschreibt Anteil der Varianz einer abhängigen Variablen Y durch ein statistisches Modell
- Nur für lineare Regression eindeutig definiert als R^2 = Korrelation^2
- Kann bedingt zur Beschreibung der Güte einer Regression verwendet werden

Bestimmtheitsmaß

$$R^2 = \frac{\sum_{i=1}^{n} (Y_{i\,\mathrm{reg}} - \bar{Y})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

$\bar{Y}$:= Mittelwert

Y_i := Datenwert

$Y_{i\,\mathrm{reg}}$:= Wert der Regression

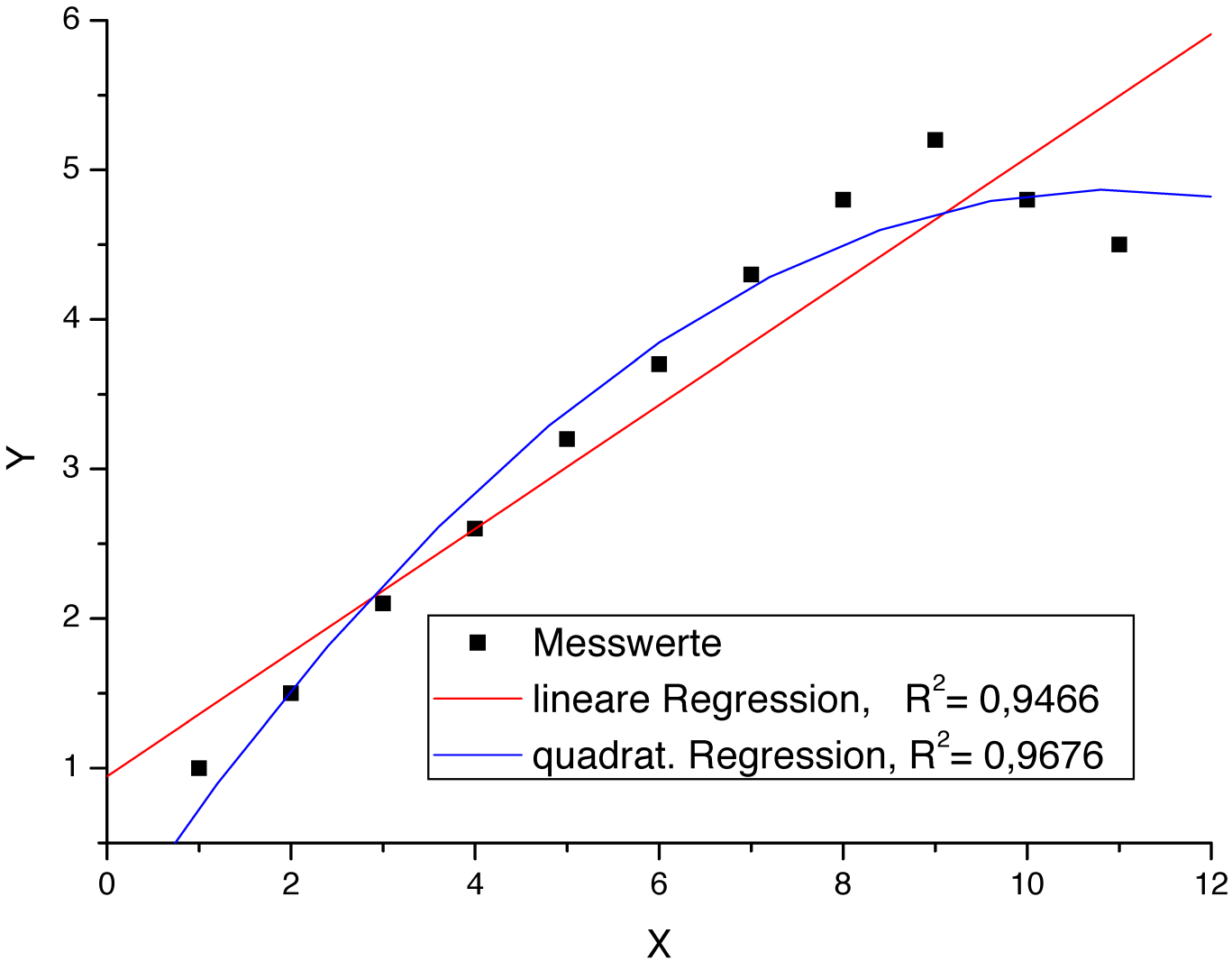
Bestimmtheitsmaß

Nachteile:

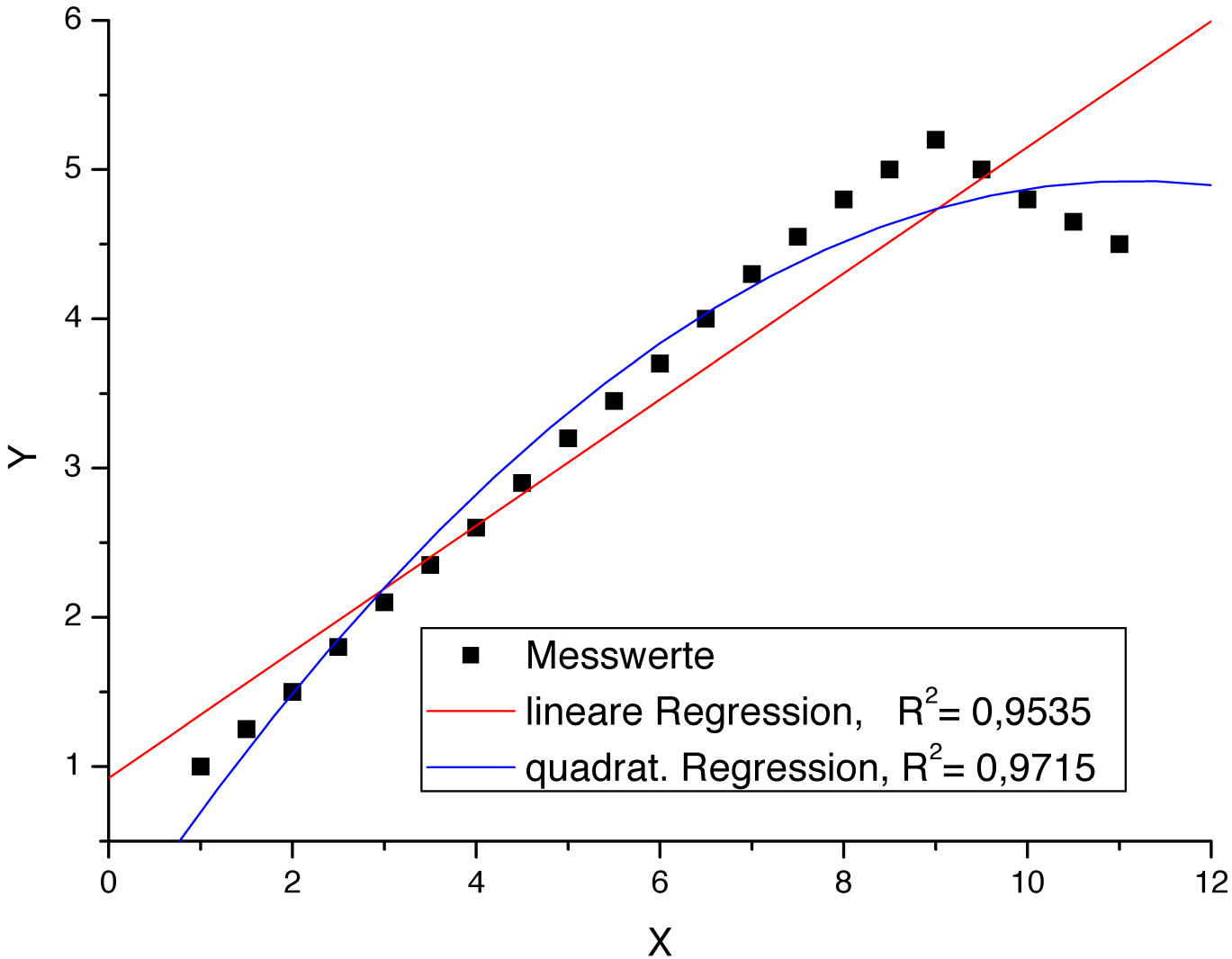
- Für mehr Messwerte steigt R^2 an, ohne dass Korrelation bzw. Regression besser wird

$$R^2 = \frac{\sum_{i=1}^{n} (Y_{i\,\mathrm{reg}} - \bar{Y})^2}{\boxed{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}}$$

Bestimmtheitsmaß



Bestimmtheitsmaß



Bestimmtheitsmaß

Nachteile:

- Für mehr Werte steigt R^2 an, ohne dass Korrelation bzw. Regression besser wird
  - Kann durch „korrigiertes“ R^2 behoben werden ($\bar{R}^2$)

$$\bar{R} = 1 - \left(1 - R^2\right) \frac{n-1}{n-p-1}$$

n := Anzahl der Werte

p := Anzahl der Variablen des Regressionsmodells

Bestimmtheitsmaß

Nachteile:

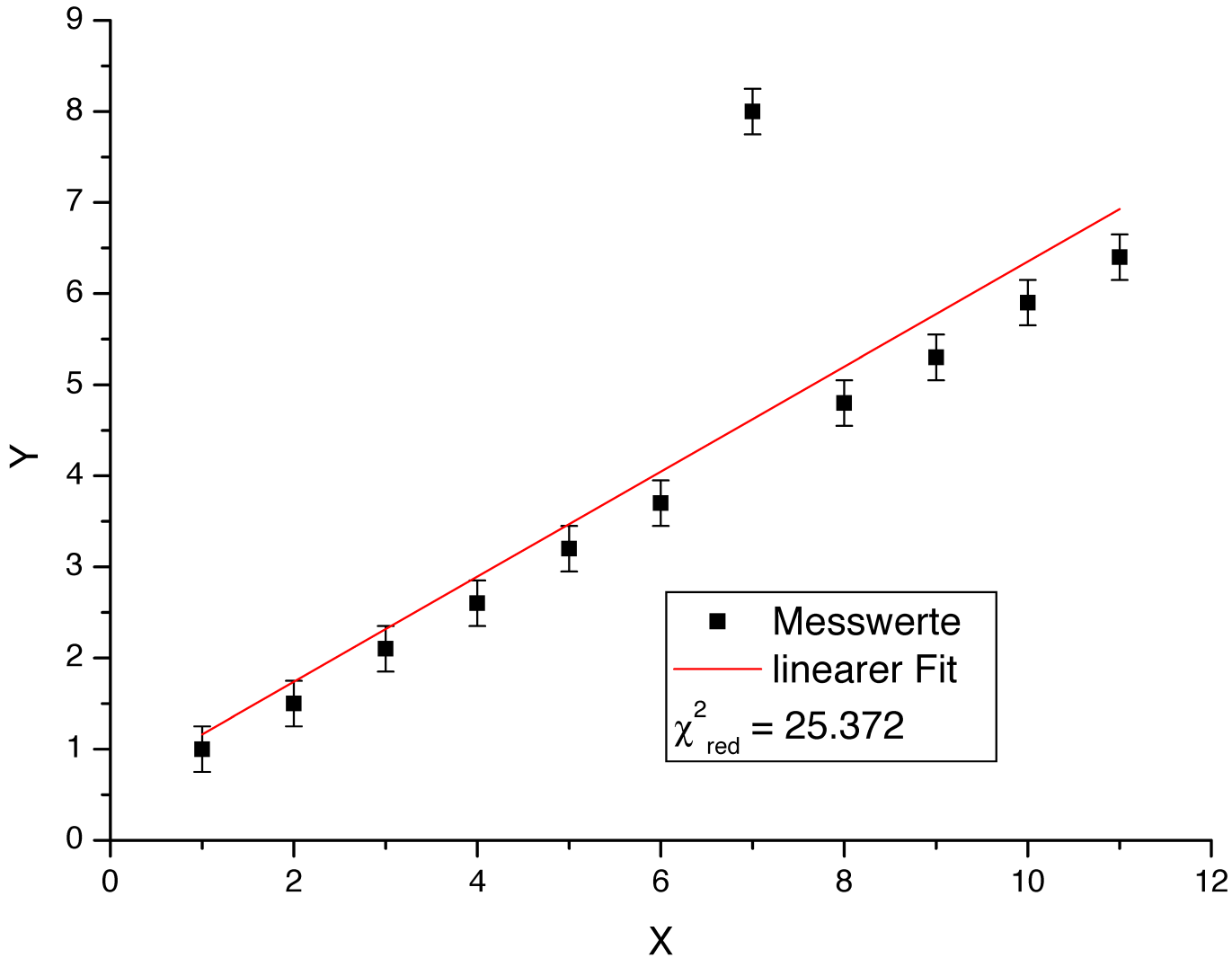
- Für mehr Werte steigt R^2 an, ohne dass Korrelation bzw. Regression besser wird
- Keine Aussage, ob korrektes Regressionsmodell verwendet wurde

Least Square Fit

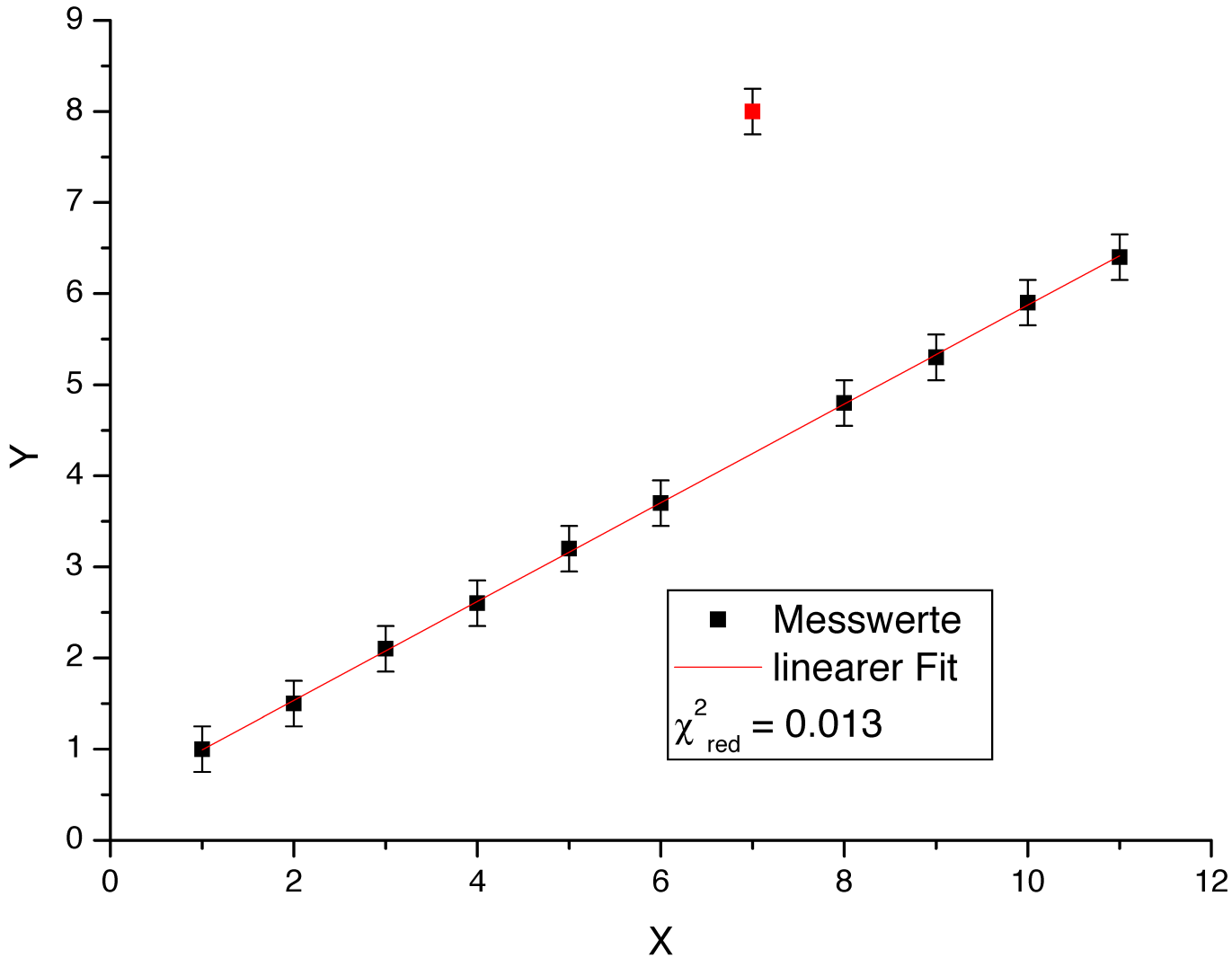
Eigenschaften

- Fit minimiert quadrat. Abweichung der Fitwerte von den Messwerten
- Ist anfällig für „Außreißer“

Ausreißer


Y
X
0
1
2
3
4
5
6
7
8
9
0
2
4
6
8
10
12
Messwerte
linearer Fit
χ^2_{red} = 25.372

Ausreißer

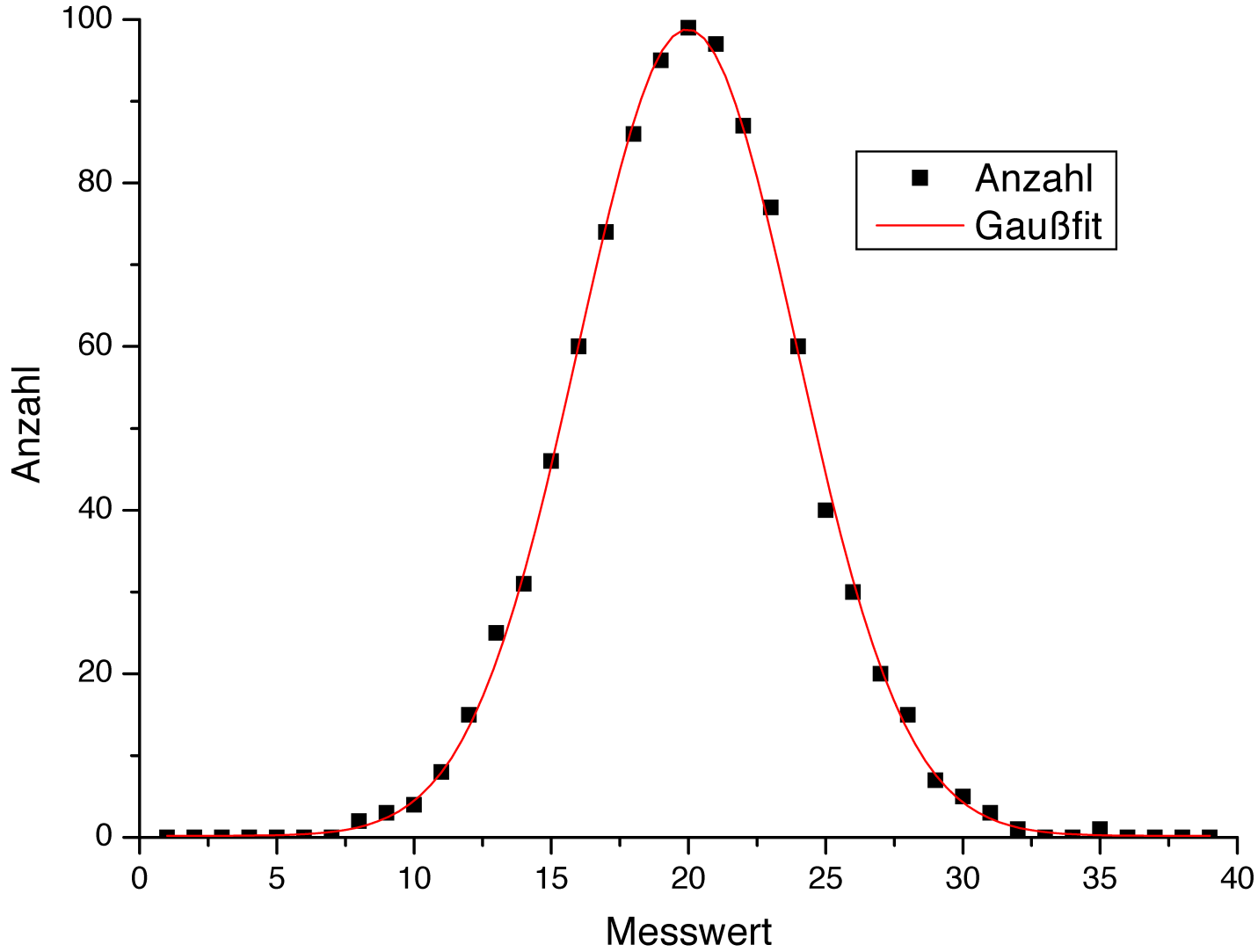

Y
X
0
1
2
3
4
5
6
7
8
9
0
2
4
6
8
10
12
Messwerte
linearer Fit
χ^2_{red} = 0.013

Least Square Fit

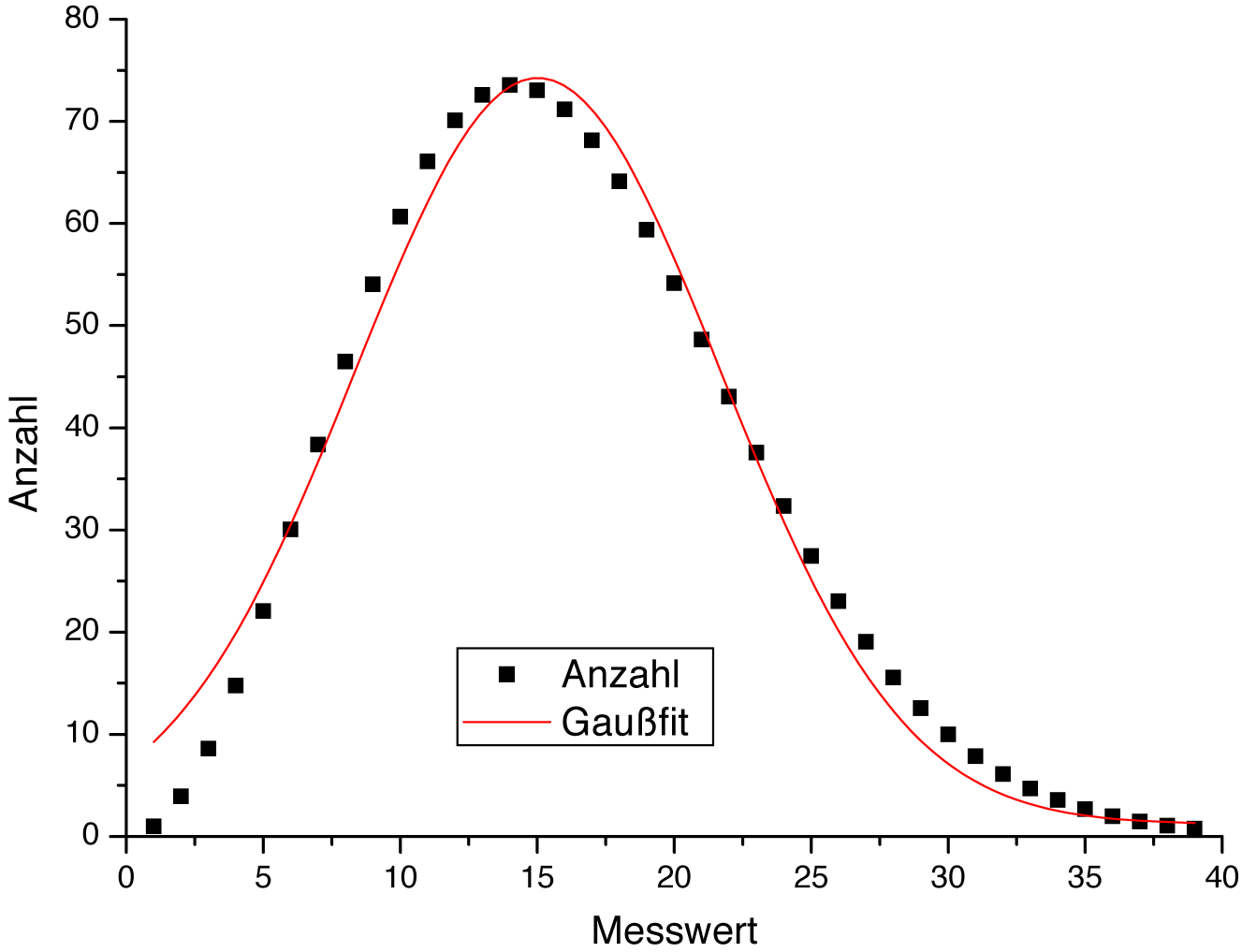
Eigenschaften

- Fit minimiert quadrat. Abweichung der Fitwerte von den Messwerten
- Ist anfällig für „Außreißer“
- Wenn Messwerte normalverteilt sind (zufällige Messfehler), kann Fitgüte durch Berechnung von χ^2_{red} ermittelt werden.

Zufälliger Messfehler



Nicht-zufälliger Messfehler



Z.B. durch zeitliche Drift des Messgeräts durch Temperaturerhöhung

Schlussfolgerung

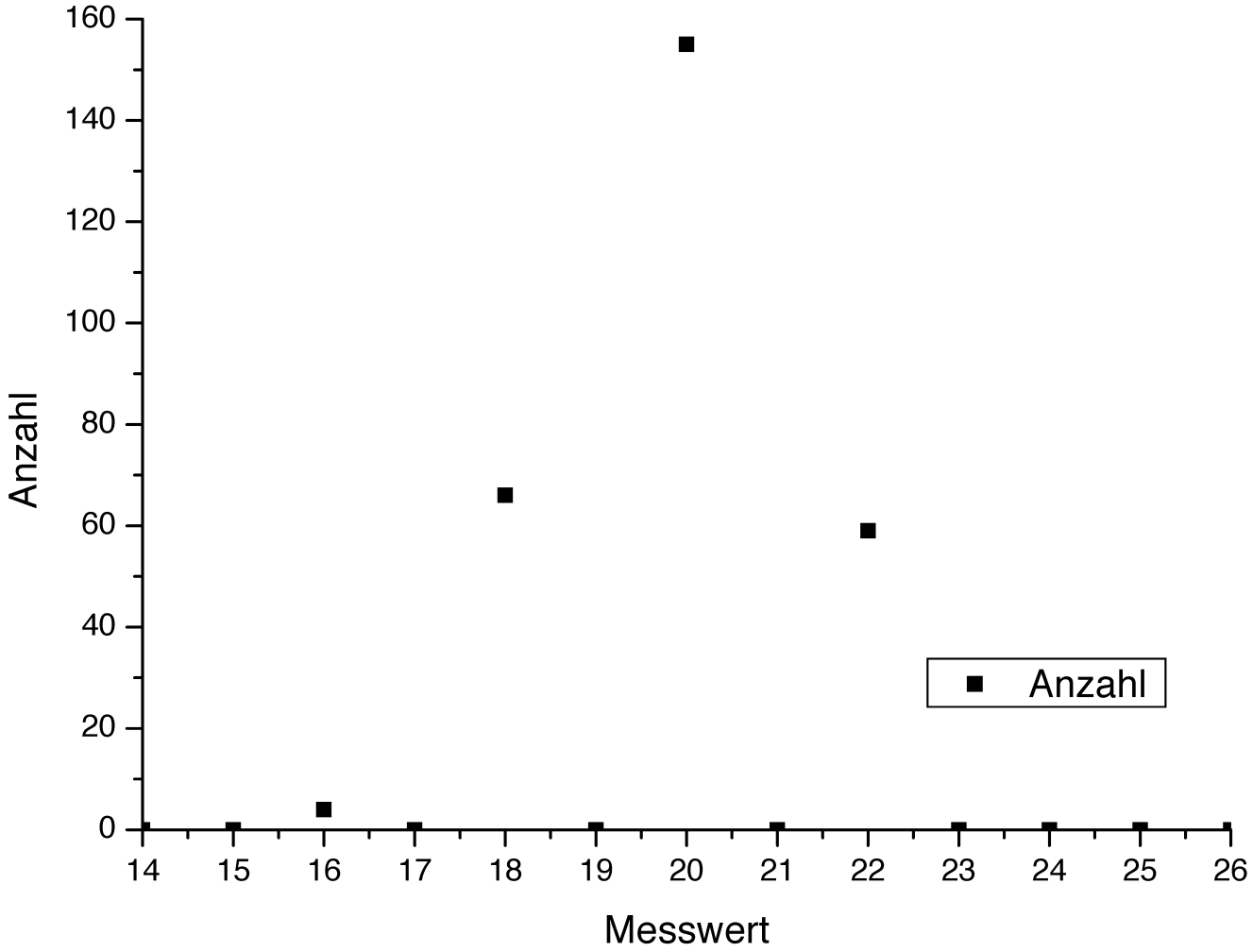
Drifts von Messgeräten müssen eliminiert werden.

Wenn das nicht möglich ist, muss Drift gemessen werden und Messwerte entsprechend der Drifts korrigiert werden.

Häufige Fragen

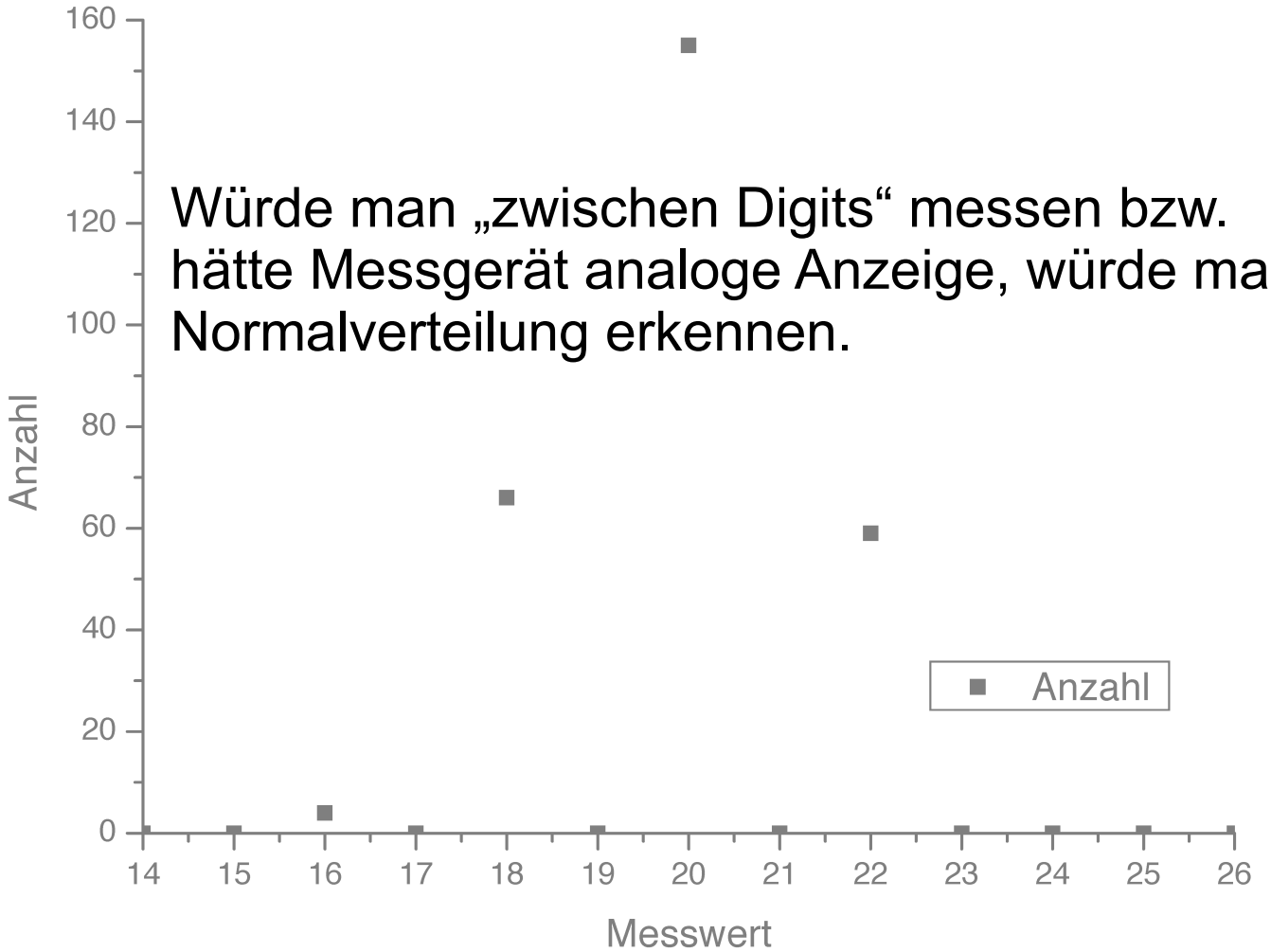
Sind digital gemessene Werte auch normalverteilt?

„Digitaler“ Messfehler



Messgenauigkeit des Geräts ist 2 (z.B. 2 Digits)

„Digitaler“ Messfehler



Würde man „zwischen Digits“ messen bzw. hätte Messgerät analoge Anzeige, würde man Normalverteilung erkennen.

Häufige Fragen

Sind digital gemessene Werte auch normalverteilt?

Antwort: Ja

Häufige Fragen

Was gibt man als Messfehler an?

Angabe des Messfehlers

Allgemein:
Messfehler = systematischer + zufälliger Fehler

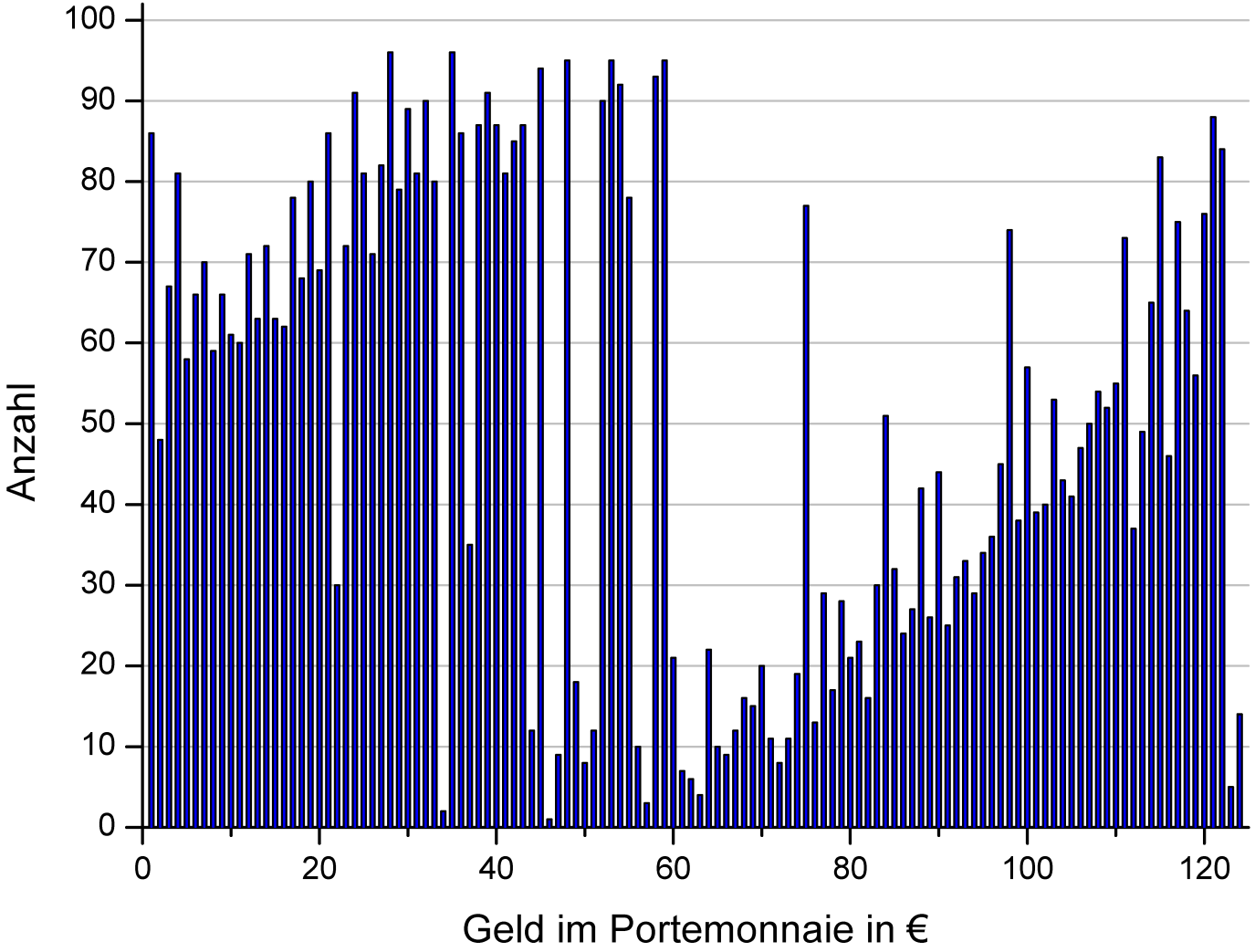
- Systematischer Fehler:
  Fehler des Messgeräts, Fehler die immer gleich gemacht werden
  (Herstellerangabe, mind. letztes Digit von Digitalanzeigen)

- Zufälliger Fehler:
  Fehler des Messenden
  (Ablesefehler, Fehler durch andere Umweltbedingungen)

Häufige Fragen

Gibt es Messungen ohne Fehler?

Messfehler?



7750 Leute wurden in der Fußgängerzone befragt, wie viel Geld sie dabei haben

Häufige Fragen

Gibt es Messungen ohne Fehler?

Antwort: Nein!

Gibt es keinen Fehler, ist es keine Messung

→ Fit ohne Fehler nicht möglich

X^2-Verteilung

Wenn X_i unabhängige, normalverteilte, zufällige Variablen sind, ist Summe ihrer Quadrate X^2-verteilt:

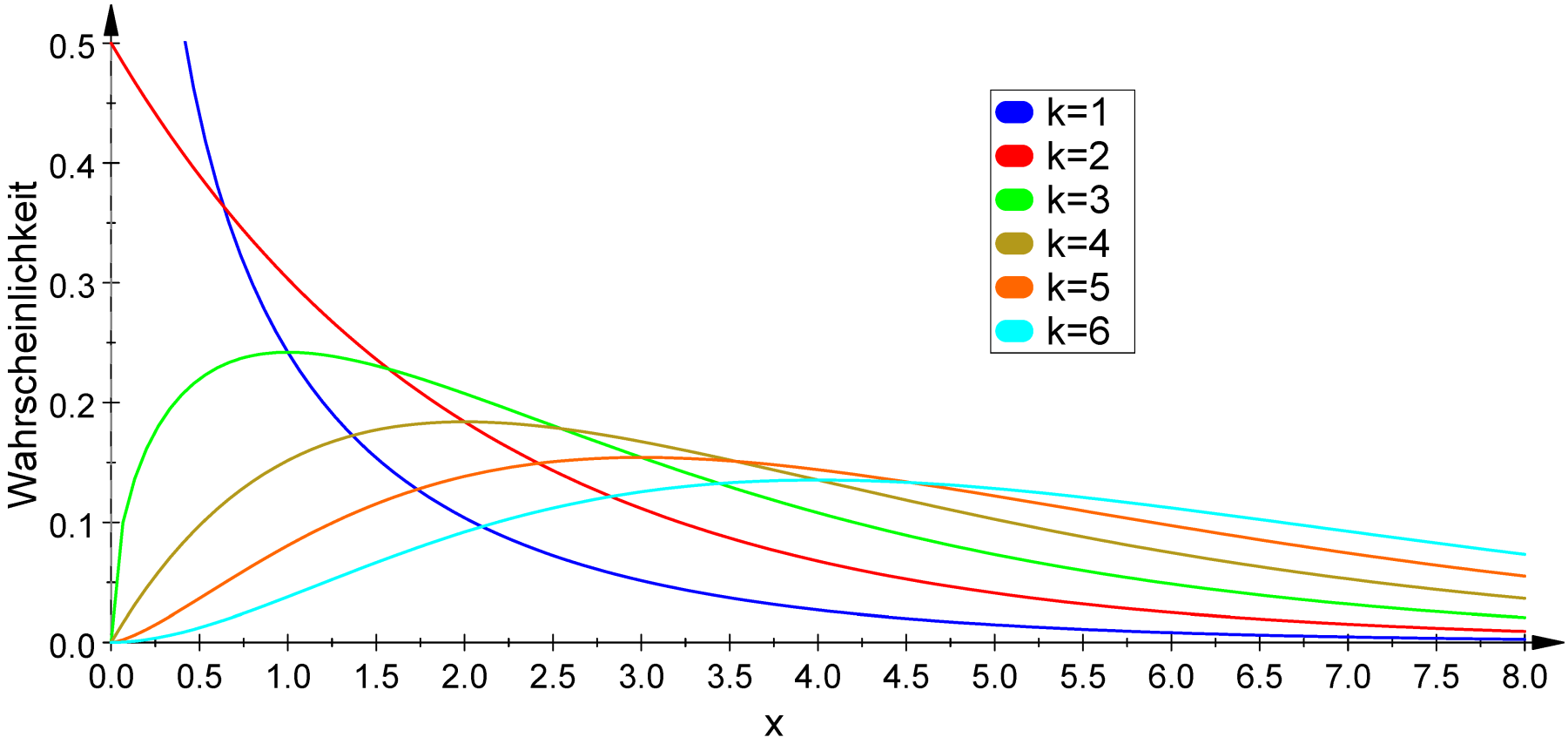
$$Q = \sum_{i=1}^{k} X_i^2 \sim \chi^2(k)$$

$$\chi^2(X,k) = \sum_{i=1}^{k} \frac{(X_i - \bar{X}_i)^2}{\sigma_i^2}$$

k := Anzahl der Freiheitsgrade

σ := Standardabweichung

χ^2-Verteilung



k := Anzahl der Freiheitsgrade

χ^2-reduziert

χ^2-reduziert: normiertes χ^2 mit Wichtung

$$\chi^2_{\mathrm{red}} = \frac{1}{k-p} \sum_{i=1}^{k} w_i (Y_i - Y_{i\,\mathrm{fit}})^2$$

k := Anzahl der Messwerte

p := Anzahl der Variablen der Fitfunktion

w := Wichtung

χ^2-reduziert

Wichtung ist entweder

- $1/\sigma^2$
  (z.B. $\sigma^2 = N$ für Poisson-verteilte Messwerte)

- 1/Messfehler^2
  (wenn σ der Messwerte unbekannt)

$$\chi^2_{\mathrm{red}} = \frac{1}{k-p} \sum_{i=1}^{k} w_i (Y_i - Y_{i\,\mathrm{fit}})^2$$

χ^2-reduziert

Eigenschaften

$$\chi^2_{\mathrm{red}} = \frac{1}{k-p} \sum_{i=1}^{k} w_i (Y_i - Y_{i\,\mathrm{fit}})^2$$

- Wegen Normierung mit k - p braucht man z.B. für linearen Fit mit 2 Variablen:
  $y = A(x) + y_0$
  min. 3 Messwerte

χ^2-reduziert

Eigenschaften

- $\chi^2_{red} \rightarrow 1$, wenn quadrat. Abweichung = σ^2 und $k >> p$
- χ^2_{red} gibt an, mit welcher Sicherheit der Fit die Messwerte besser beschreibt als zufällige Werte

χ^2-Verteilung

Sicherheit des Fits

Wahrscheinlichkeitsdichte der χ^2-Verteilung:

$$F_\chi(X,\, k) = \frac{1}{2^{k/2}\,\Gamma(k/2)}\, X^{k/2-1} e^{-X/2}$$

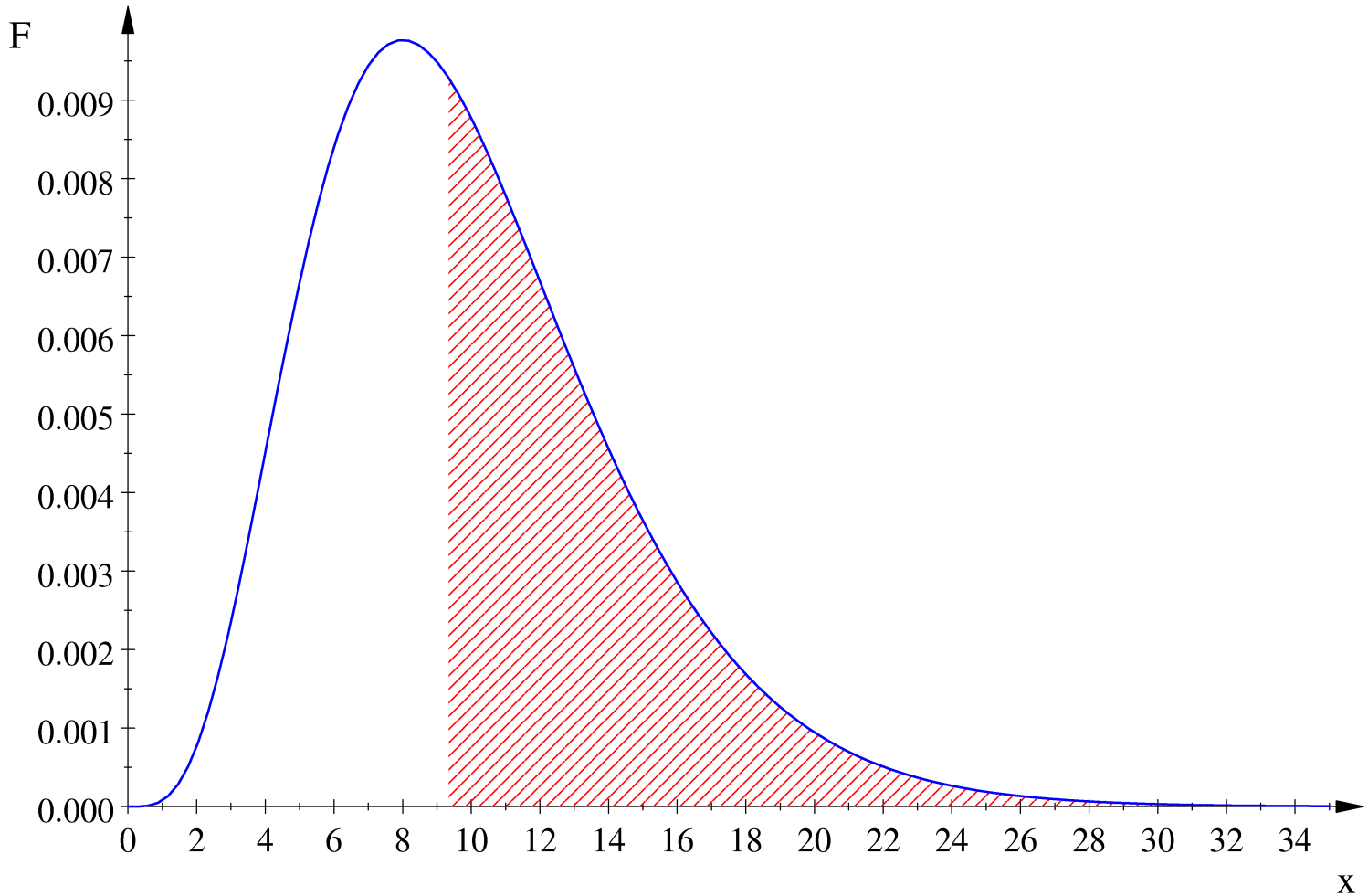
Wahrscheinlichkeit, dass χ^2_{red} von normalverteilten Zufallswerten größer ist als χ^2_{red} des Fits:

$$P\left(\chi^2_{\text{red}},\, D\right) = \int_{\chi^2_{\text{red}}\cdot D}^{\infty} F_\chi\left(X, k\right) \mathrm{d}X$$

Γ := Gammafunktion

$D = k - p$:= Anzahl der Freiheitsgrade

Sicherheit des Fits



Kurve: $F_\chi(x, 10)$; schraffiert: $P(\chi^2_{\mathrm{red}}, 10) = 0.5$

X^2-reduziert

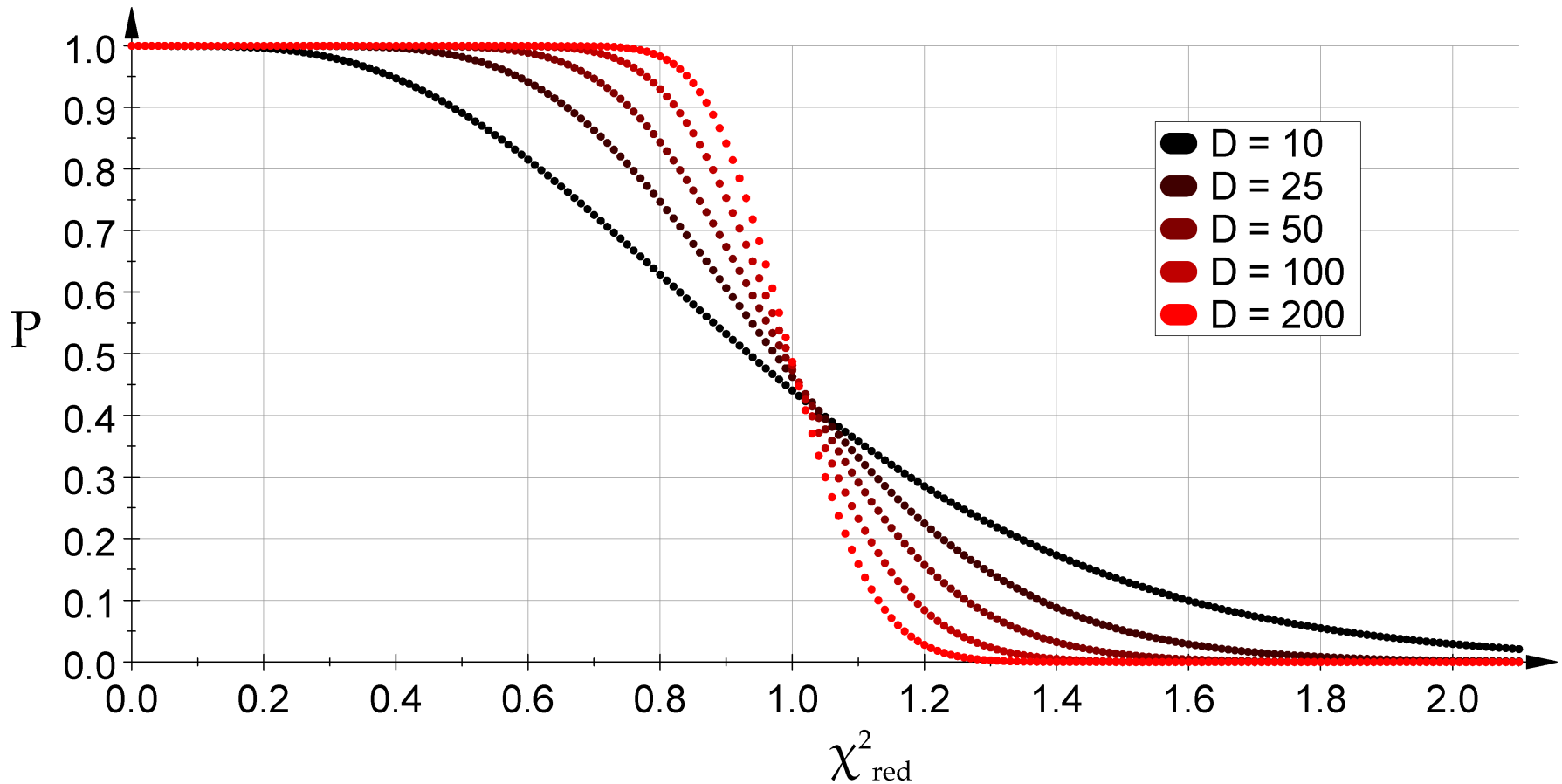
Wertetabelle von $P(X^2_{red}, D)$

$D \setminus P$	0.01	0.1	0.5	0.9	0.95
10	2.321	1.599	0.934	0.487	0.394
20	1.878	1.421	0.967	0.622	0.543
30	1.696	1.342	0.978	0.687	0.616
40	1.592	1.295	0.983	0.726	0.663
50	1.523	1.263	0.987	0.754	0.695
100	1.358	1.185	0.993	0.824	0.779
200	1.247	1.130	0.997	0.874	0.841

Daten auch verfügbar in Tab. C.4 in *Bevington, Robinson, 2002, „Data Reduction and Error Analysis for the Physical Sciences“*

χ^2-reduziert

Verlauf von $P(\chi^2_{red}, D)$



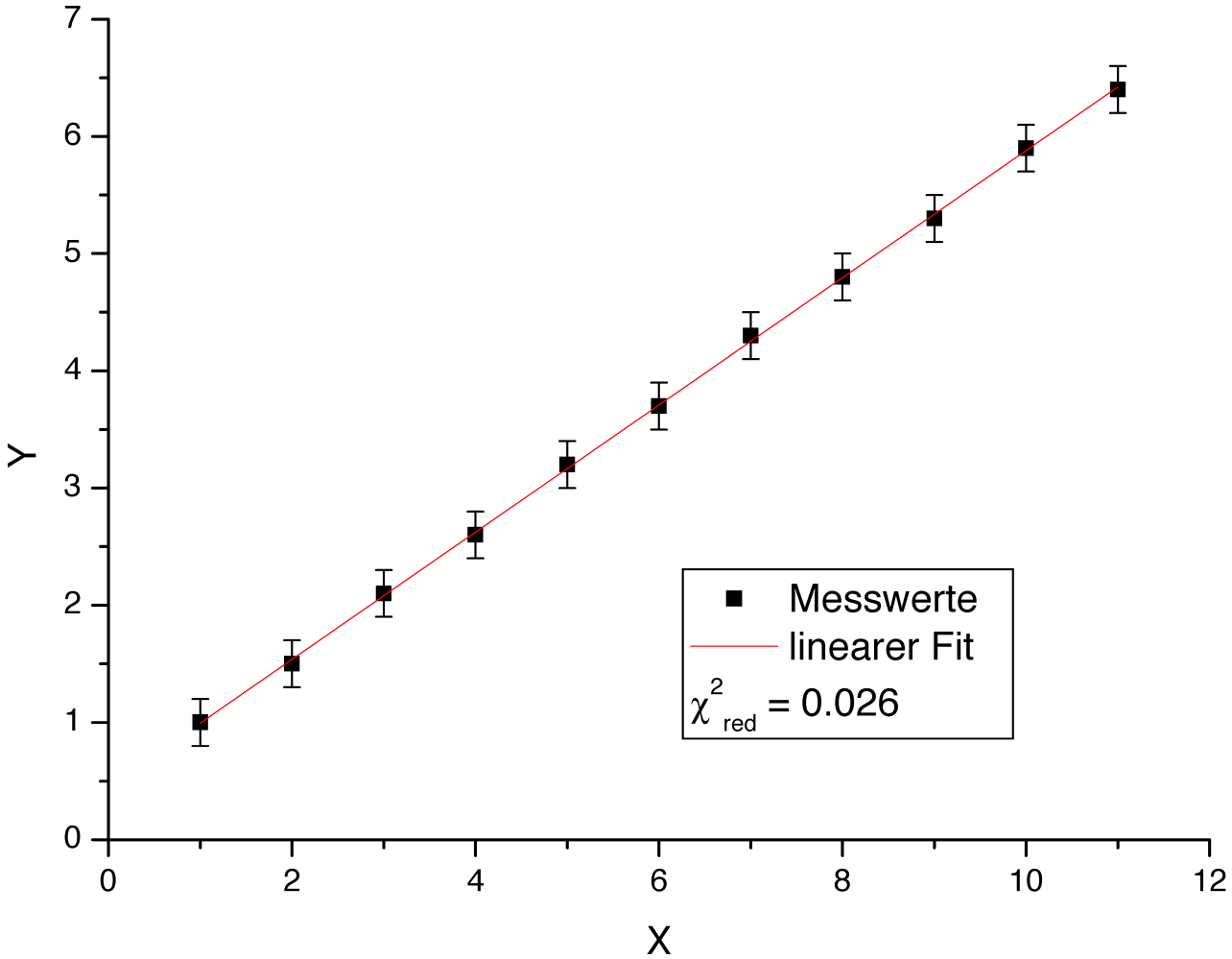
χ^2-reduziert

Eigenschaften

Faustregeln:

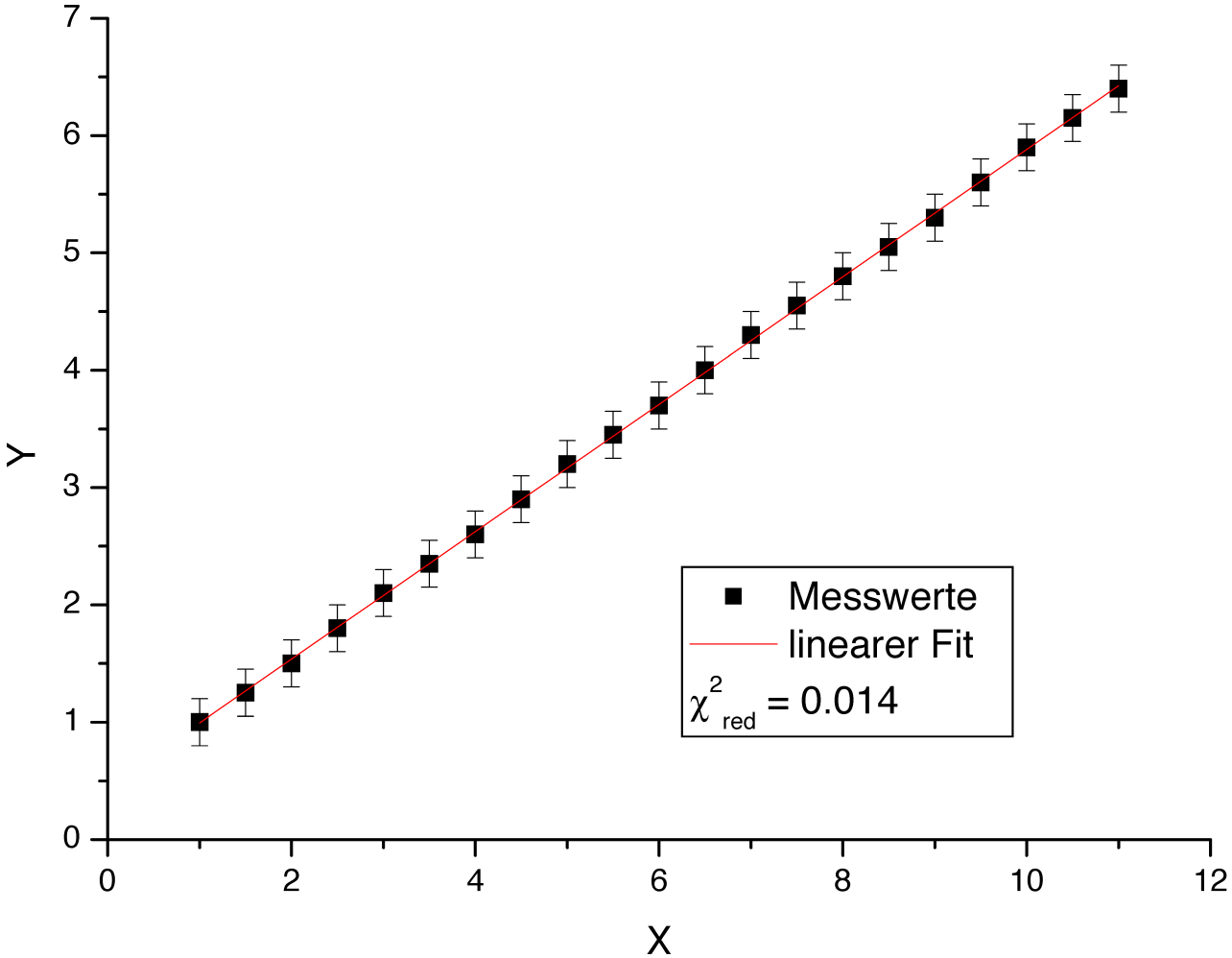
- Wenn $\chi^2_{red} > 2$, beschreibt Fit Messung nicht brauchbar oder der Messfehler ist kleiner gewählt als er ist
- Wenn $\chi^2_{red} << 1$, ist entweder Messfehler größer gewählt als er ist oder Wichtung wurde nicht oder falsch angegeben
- Solange $k/p \not> 10$, sinkt χ^2_{red} wenn k steigt

Beispiele



Fehler ist größer angenommen worden, als er tatsächlich ist

Beispiele



Kleineres χ^2_{red} durch mehr Messwerte

Anwendungen für Fits

- Bestimmen von Werten

Bestimmung von Werten

Anwendung:

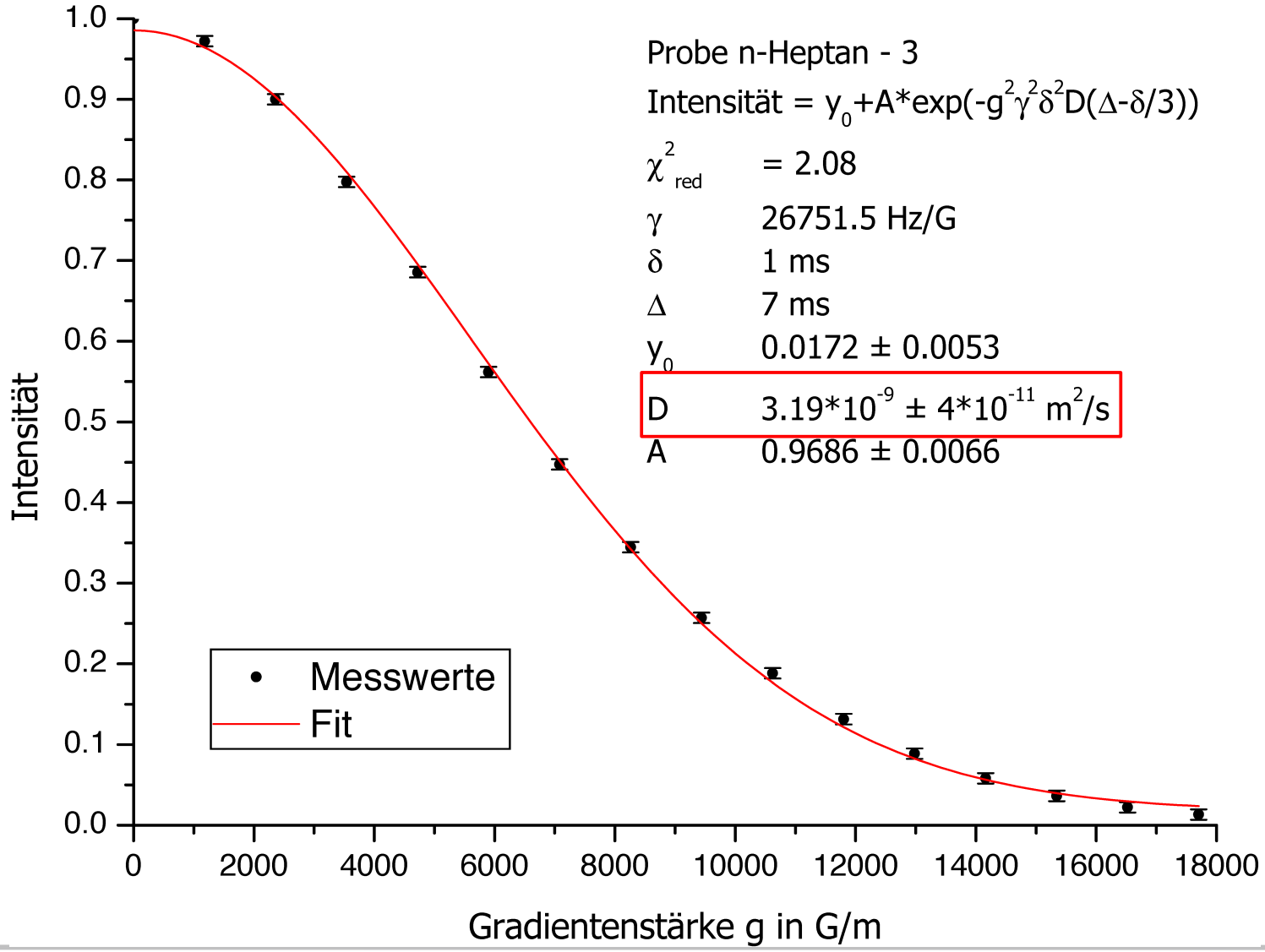
Bestimmung des Diffusionskoefizienten *D*
Mittels NMR-Pulsen kann *D* bestimmt werden

Dazu wird Intensität I gemessen, die folgender Gleichung gehorcht:

$$I = y_0 + A \exp(-g^2\gamma^2\delta^2 D(\Delta - \delta/3))$$

Messung der Intensität ergibt:

Messung der Intensität



Bestimmung von Werten

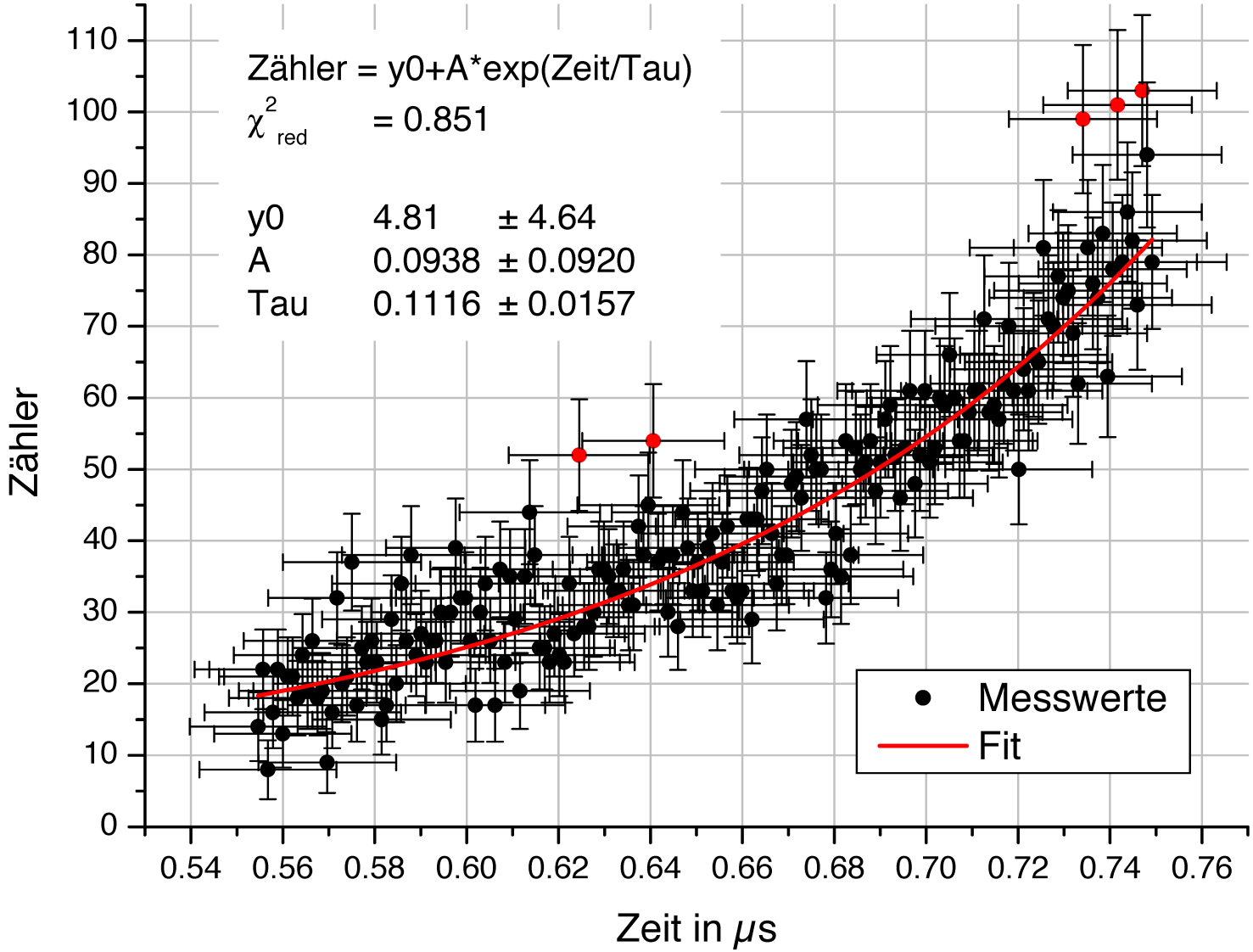
Anwendung:

Bestimmung der Zerfallskonstante τ von $^{57}Fe_{14.4\ keV}$
mit Methode der verzögerten Koinzidenzen
Zählrate N gehorcht der Gleichung:

$$N = y_0 + A \exp(t/\tau)$$

Messung ergibt:

Messung der Zählrate



Bestimmung von Werten

Ergebnis:

χ^2_{red} = 0.85 bei D = 179

$\rightarrow P$ = 0.92

Problem:

- x-Messfehler wird für Berechnung des χ^2_{red} nicht berücksichtigt

Fitten mit x-Fehler

Prozedur:

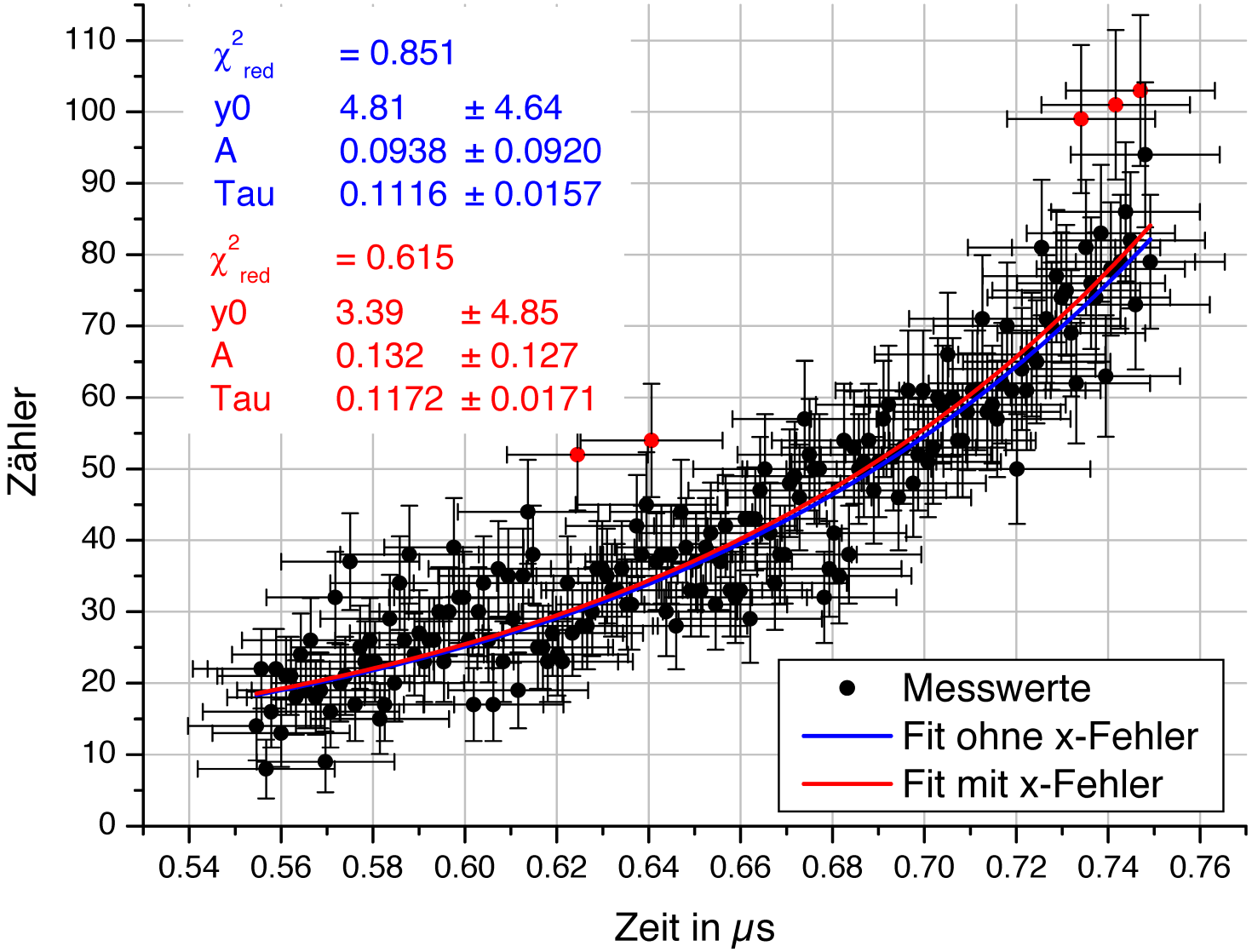
1. „normal“ Fitten mit y-Fehler
2. Fitten mit x,y-Fehler unter Verwendung der Parameter des ersten Fits

$$\Delta x, y = \sqrt{(\Delta y)^2 + \left(\frac{\partial y_{\mathrm{fit}}}{\partial x}\Delta x\right)^2}$$

Fehler folgt aus Fehlerfortpflanzung
(Einfluss des x-Fehlers auf y-Fehler)

Quelle: http://www.physik.uzh.ch/lectures/datenanalyse/ss04/Kap4.pdf
oder http://wwwhep.physik.uni-freiburg.de/fp/origin/Kap4.pdf

Messung der Zählrate



Bestimmung von Werten

Ergebnis:

χ^2_{red} = 0.62 bei D = 179

$\rightarrow P$ = 1.0

- χ^2_{red} wird kleiner, da x,y-Fehler > y-Fehler
- Fitparameter werden exakter
  $\rightarrow$ Fit mit x-Fehler ergibt ca. 5 % höheres τ
- Fehler der Fitparameter werden größer

Anwendungen für Fits

- Bestimmen von Werten
- Test, ob Messwerte einem physikal. Gesetz folgen

Test auf physik. Verlauf

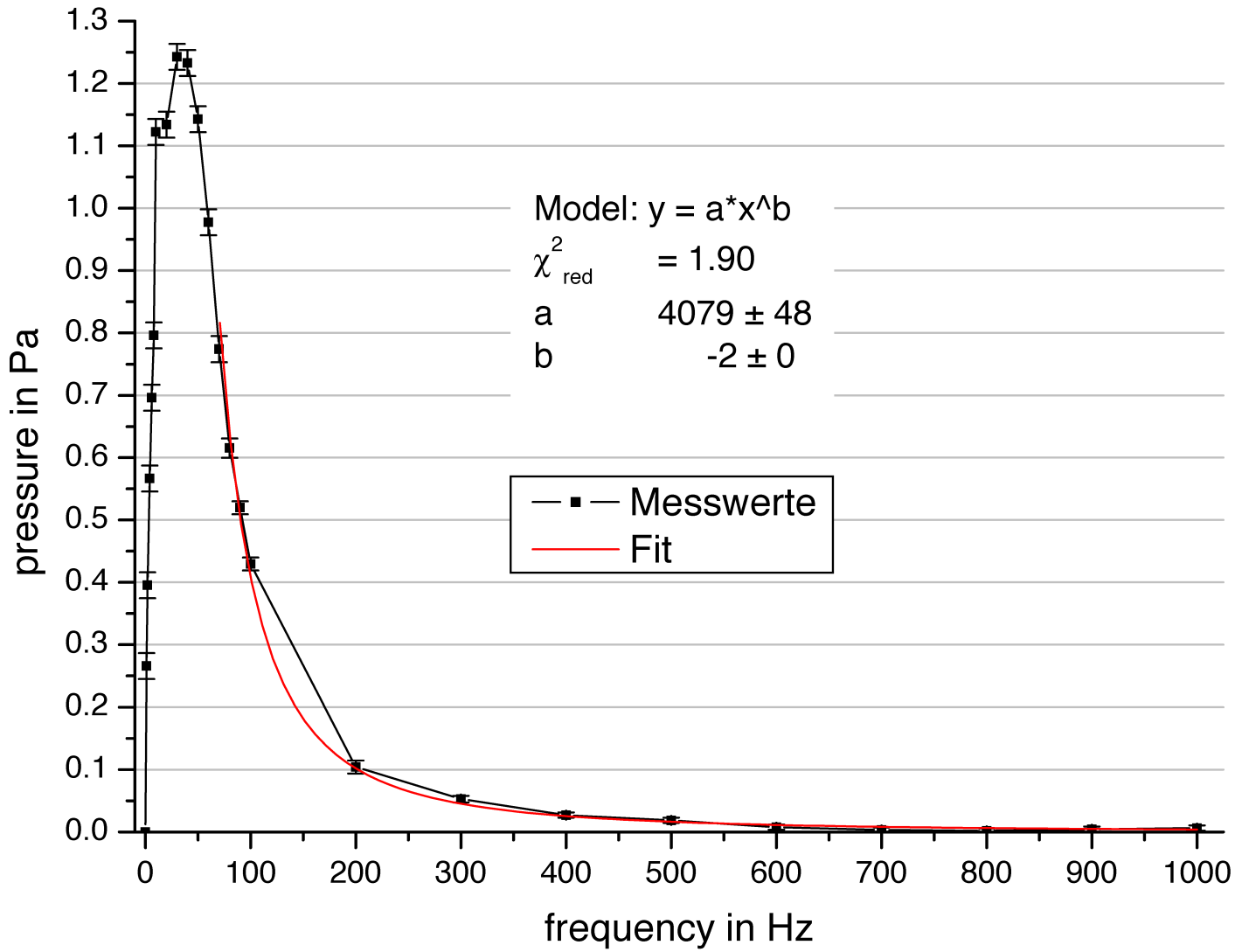
Anwendung:

Erzeugter Druck p in einer MHD-Pumpe sollte für große Anregungsfrequenzen quadratisch abnehmen:

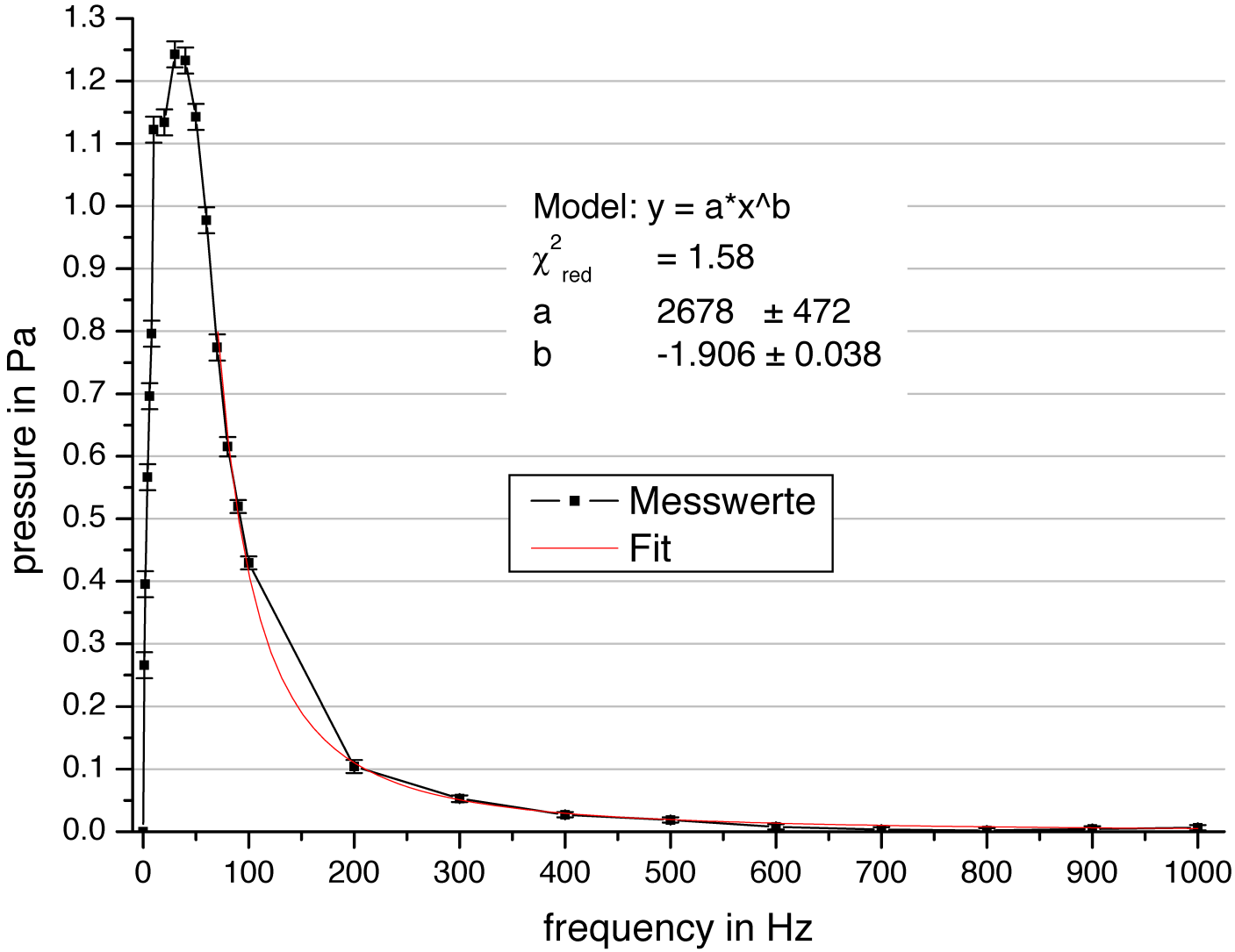
$$p = y_0 + 1/x^2$$

Messung ergibt:

Druckmessung der Pumpe



Druckmessung der Pumpe



Druckmessung der Pumpe

Ergebnis:

χ^2_{red} = 1.9 bei D = 11

$\rightarrow P$ = 0.03

Fit mit Variation des Exponenten:

χ^2_{red} = 1.58 bei D = 10

$\rightarrow P$ = 0.11

Verlauf kann nicht nachgewiesen werden
$\rightarrow$ mehr Messwerte benötigt

Anwendungen für Fits

- Bestimmen von Werten
- Test ob Messwerte einem physikal. Gesetz folgen
- Test wie Messwerte verteilt sind

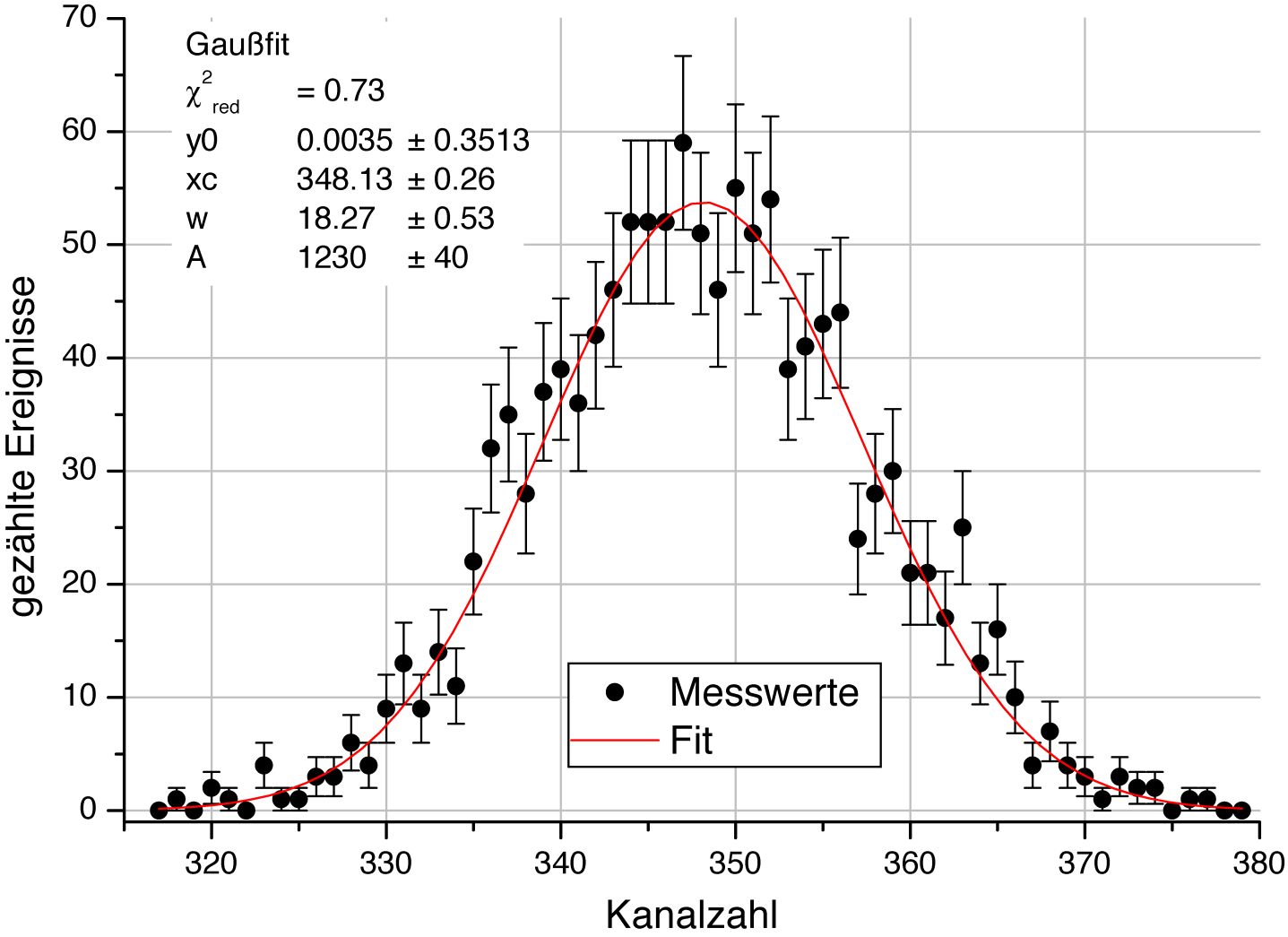
Test auf Verteilung

Anwendung:

Energieeichung eines Szintillators
Szintillator wird mit Gammastrahlung definierter Energie bestrahlt
Messwerte sollten Gauß-verteilt sein und gemessenes Maximum bei der definierten Energie liegen.

Energiemessung ergibt:

Messung der Energie



Messung der Energie

Ergebnis:

$X^2_{red} = 0.73$ bei $D = 59$

$\rightarrow P = 0.94$

Messwerte sind mit 94% Sicherheit Gauß-verteilt
Detektorkanal 348 entspricht der definierten Energie
Standardabweichung sind 18 Kanäle

Test auf Verteilung

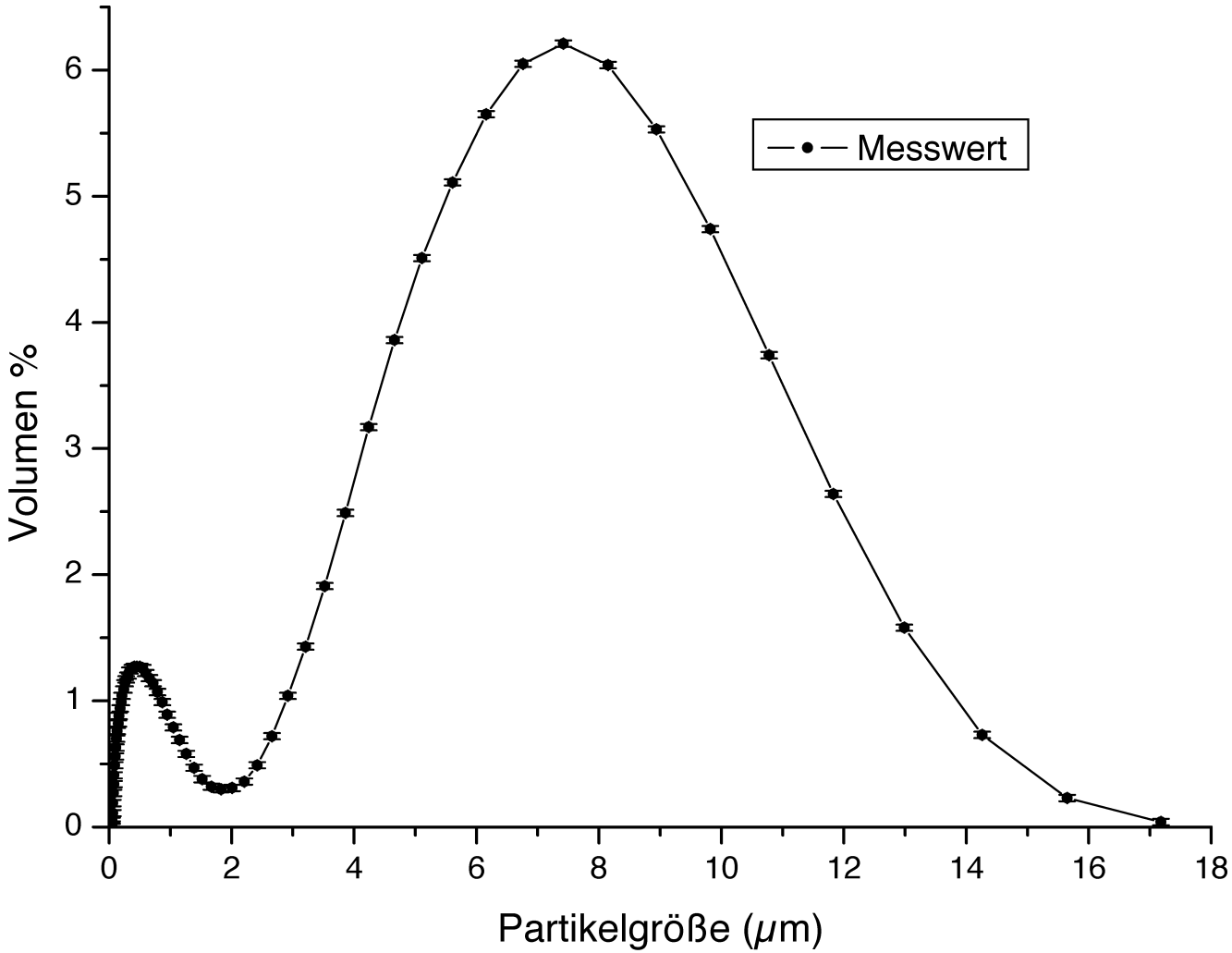
Anwendung:

Charge mit CaF_2-Pulver wurde geliefert

Hersteller gibt an, dass Korngröße des Pulvers Gauß-verteilt ist und die mittlere Größe bei 6,2 µm liegt

Messung der Korngröße ergibt:

Messung der Korngröße



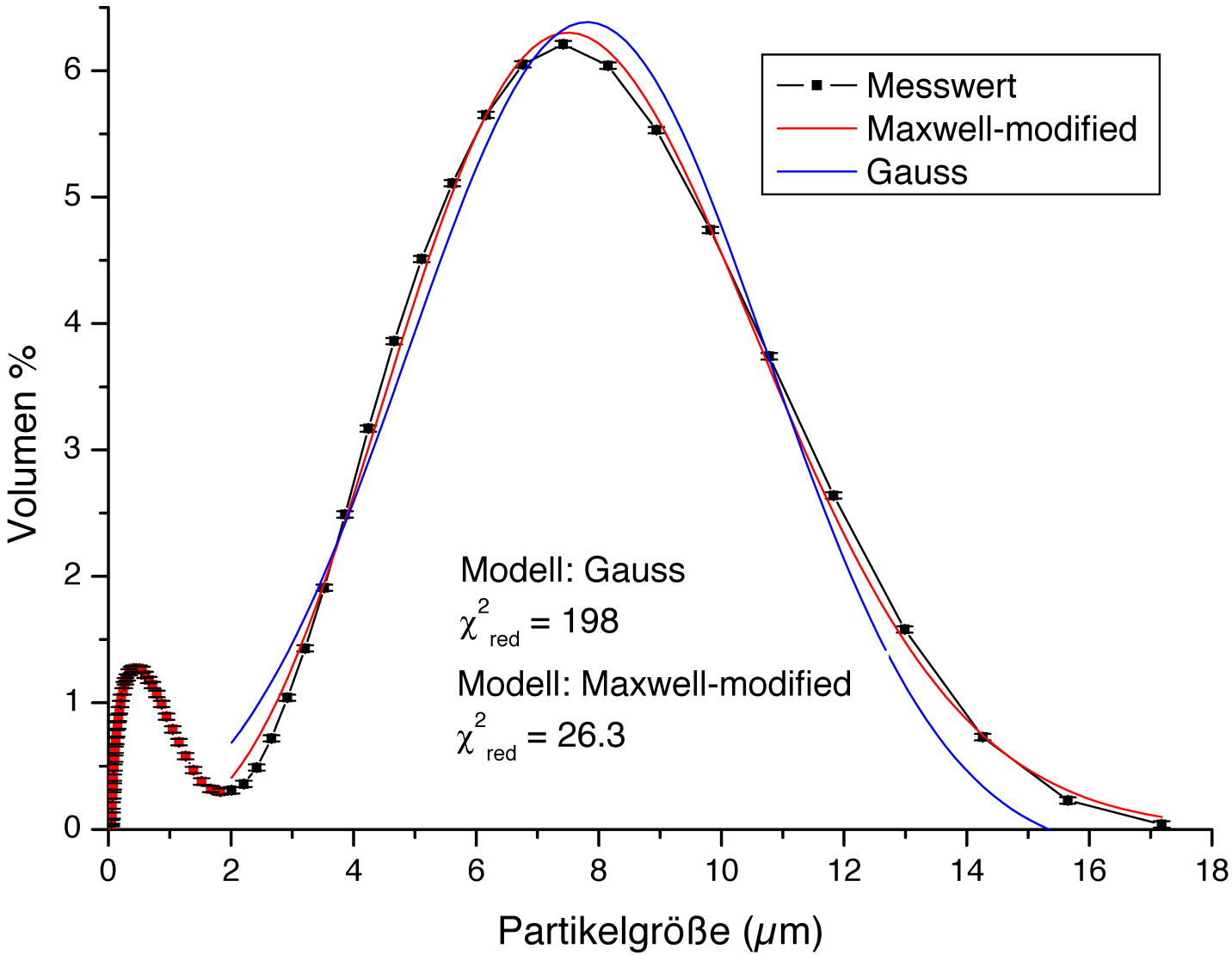
Messung der Korngröße

Ergebnis:

Charge ist nicht Gauß-verteilt
Man könnte alle Partikel < 2 µm herausfiltern, denn
Prozess erfordert eine definierte Korngrößenverteilung

Fit der Partikel > 2 µm ergibt:

Fit der Korngröße



Messung der Korngröße

Fitergebnis:
Auch Partikel > 2 µm nicht Gauß-verteilt

Fit mit modifizierter Maxwellverteilung

$$y = B\left(x^{C-1}A^{C/2}\exp\left(-0.5Ax^2\right)\right)$$

ist besser, aber beschreibt Verteilung auch nicht

Charge kann aufgrund nicht reproduzierbarer Verteilung nicht verwendet werden

Test auf Verteilung

Alternative zu Gaußfit und Bewertung von χ^2_{red}

ist χ^2-Test

χ^2-Test berücksichtigt aber keinen Messfehler

Anwendungen für Fits

- Bestimmen von Werten
- Test ob Messwerte einem physikal. Gesetz folgen
- Test wie Messwerte verteilt sind
- Interpolation von Prozessparametern

Interpolation von Prozessparametern

Anwendung:

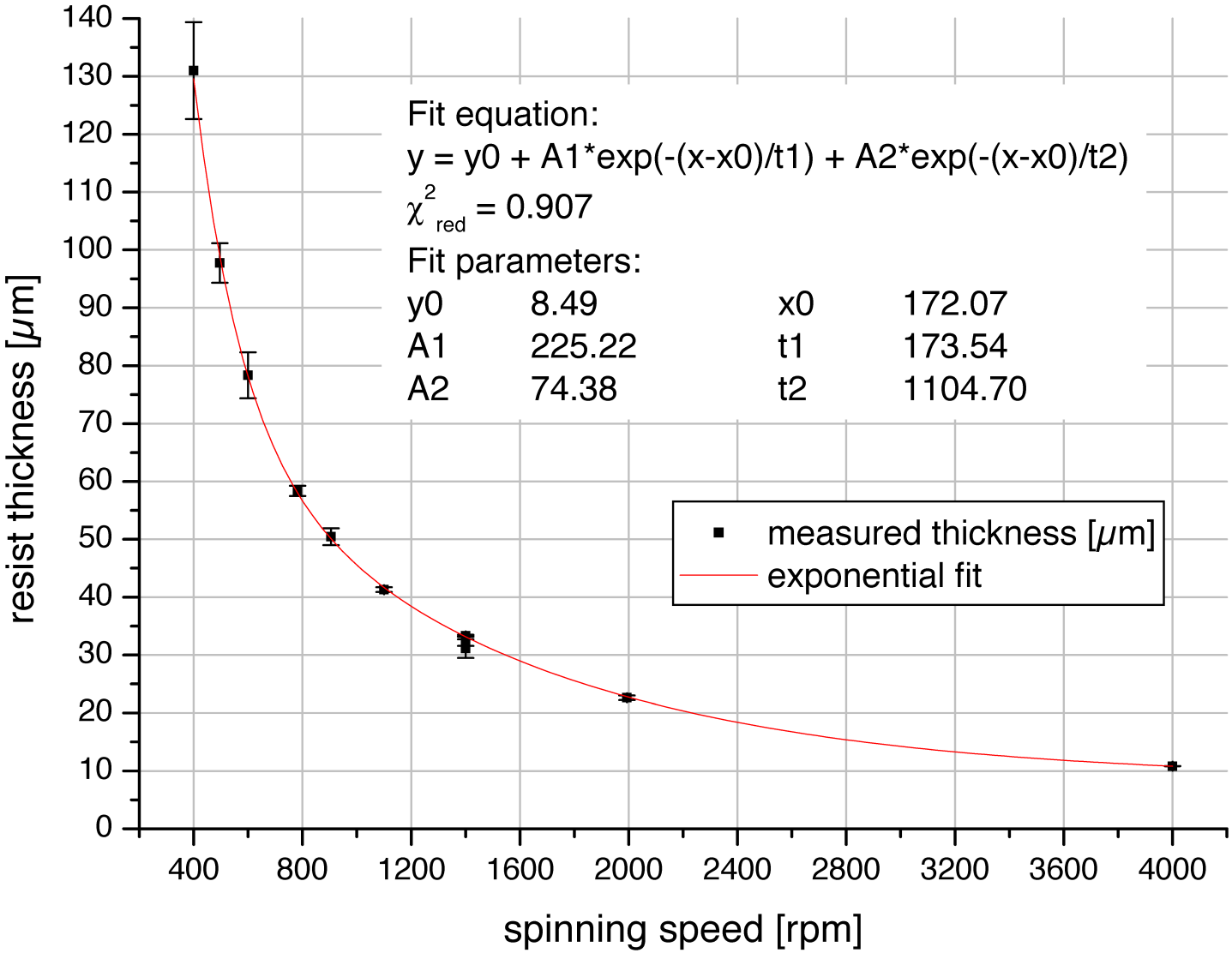
Lackkurve eines Fotolacks für die man ohne Weiteres keinen physikal. Zusammenhang angeben kann

Man möchte eine Funktion haben, die für jede Drehzahl die Lackdicke angibt

Messung der Lackkurve ergibt:

Least Square Fit

Lackkurve



Lackkurve

Ergebnis:

χ^2_{red} = 0.907 bei D = 3 (9 Messwerte, 6 Variablen)

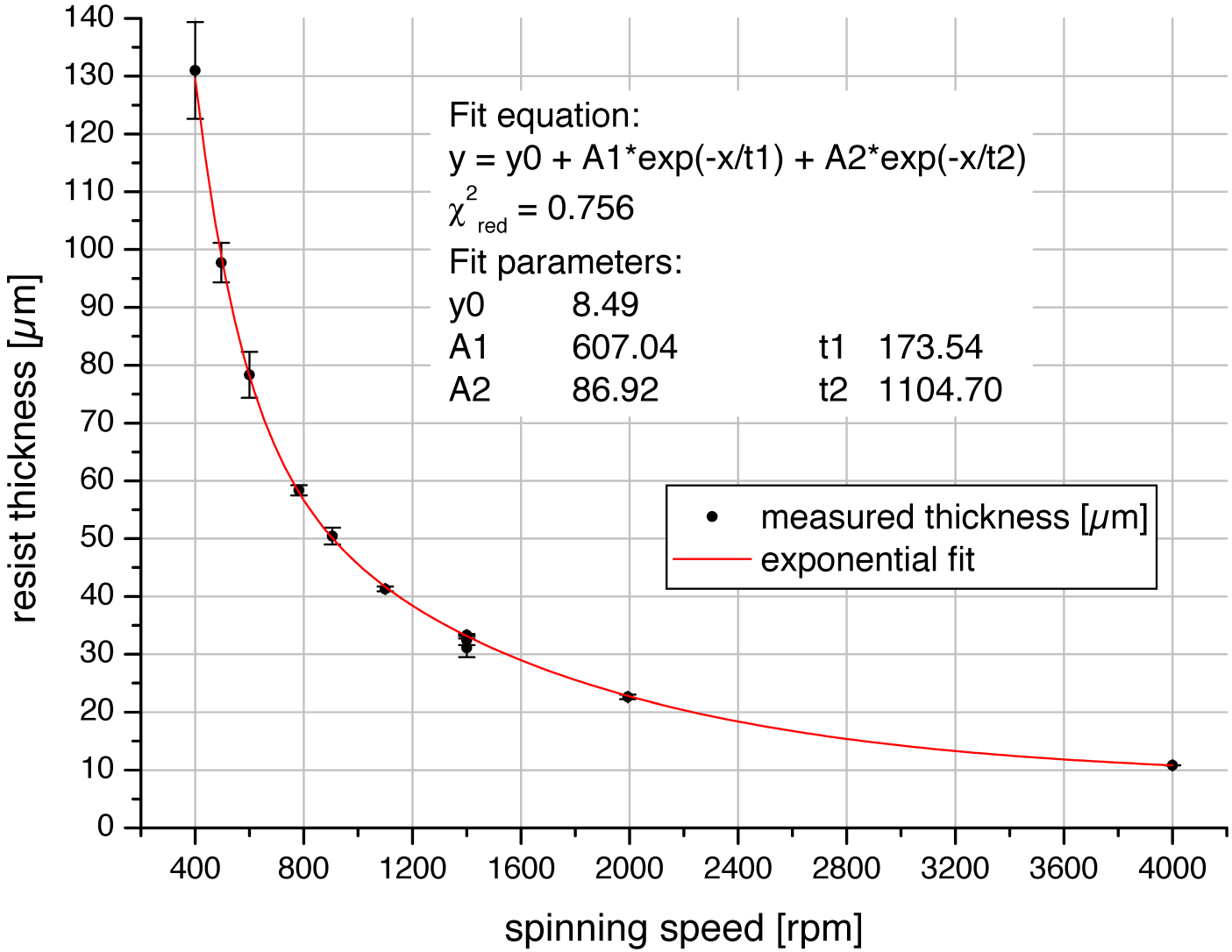
$\rightarrow P$ = 0.44

Um P zu erhöhen:

- Anzahl der Variablen reduzieren
  (besser jedoch: mehr Messwerte aufnehmen)
- Andere Fitfunktion nehmen, sofern das physikal. zulässig ist

Least Square Fit

Lackkurve



Lackkurve

Ergebnis:

$\chi^2_{red} = 0.756$ bei $D = 4 \rightarrow P = 0.55$

Andere Fitfunktionen oder Funktion mit 4 Variablen liefern deutlich schlechteres χ^2_{red}

Schlussfolgerung:
Für $P > 0.55$ müssen mehr Messwerte aufgenommen werden

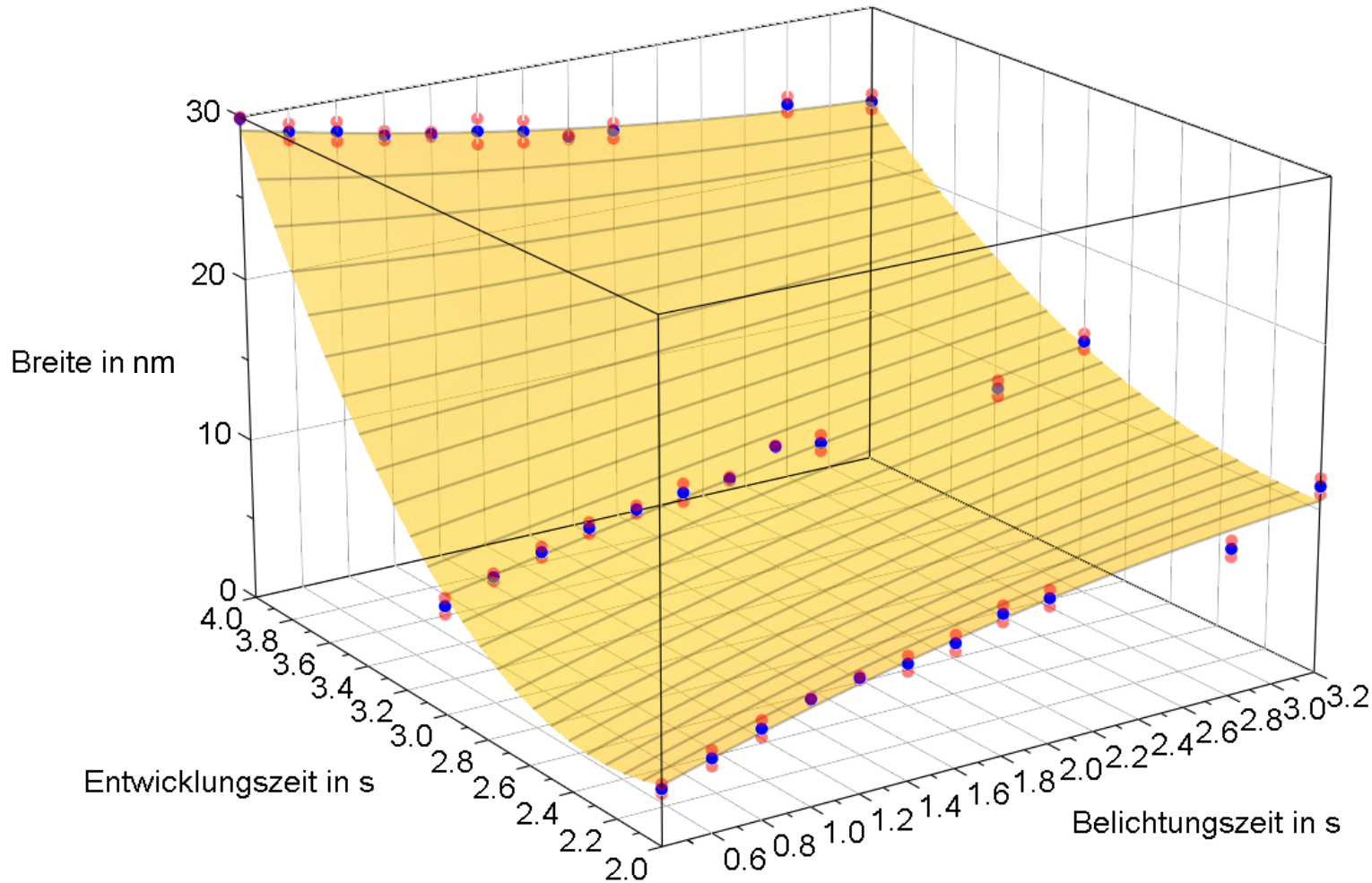
Interpolation von Prozessparametern - 3D

Anwendung:

Nanoimprintverfahren
Man möchte eine Funktion haben, die für jede Belichtungszeit UND jede Entwicklungszeit die Strukturdicke angibt

Messung der Dicke ergibt:

Interpolation von Prozessparametern - 3D



gelb: Fläche des 3D-Fits; blau: Messwerte; rot: „Fehlerbalken“

Interpolation von Prozessparametern - 3D

Ergebnis:

X^2_{red} = 2.90 bei D = 24 $\rightarrow$ P = 0.00

Für Fitfunktion

$$z = Ayx + By^2x + Cyx^2 + Dy^2x^2 + Ex + Fx^2 + Gy + Hy^2 + I$$

Interpolation von Prozessparametern - 3D

Ergebnis:

X^2_{red} = 2.90 bei D = 24 $\rightarrow$ P = 0.00

Für Fitfunktion

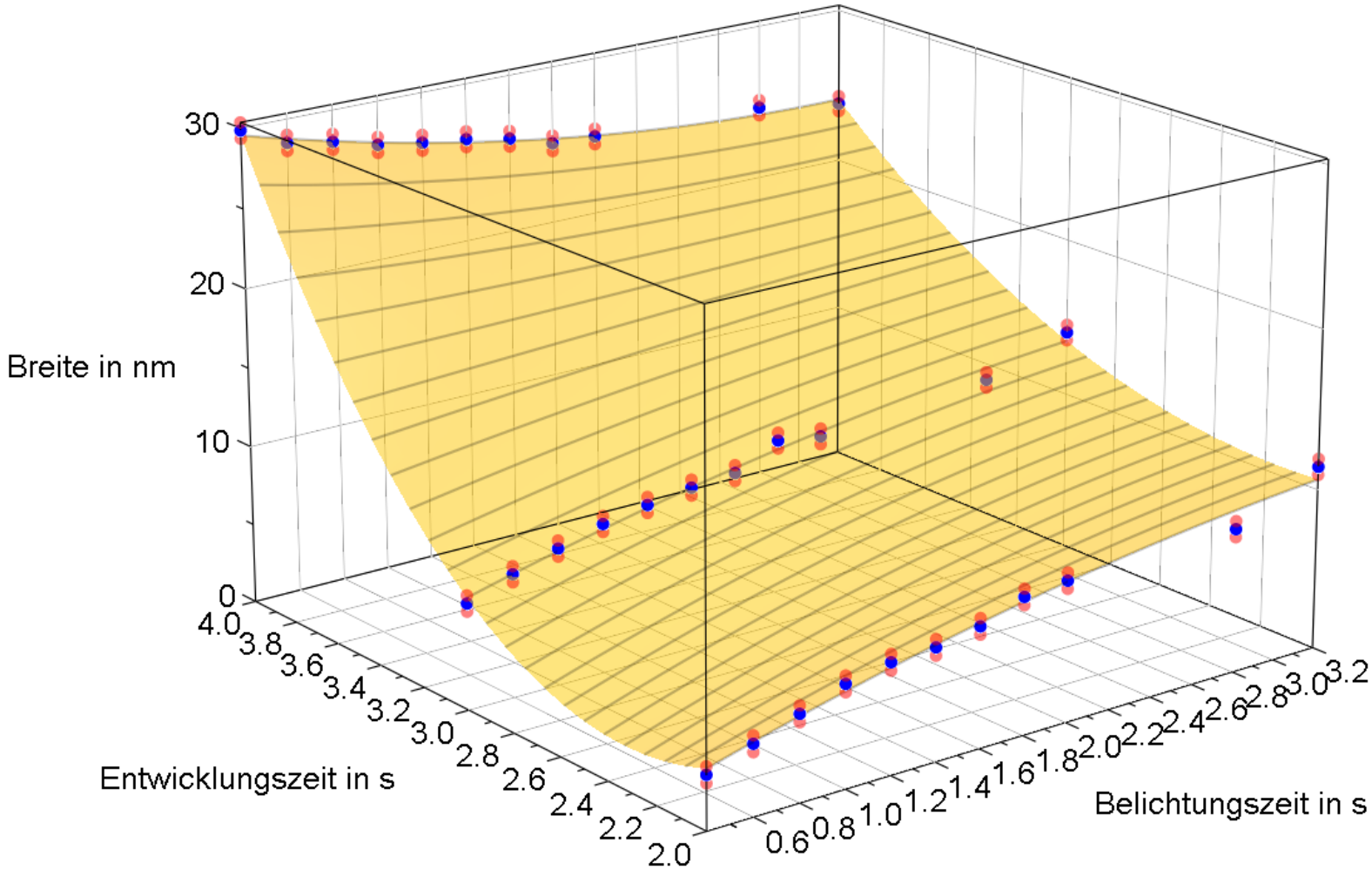
$$z = Ayx + By^2x + Cyx^2 + Dy^2x^2 + Ex + Fx^2 + Gy + Hy^2 + I$$

Aber:

- Messfehler sind sehr unterschiedlich
- Messpunkte fehlen

$\rightarrow$ X^2_{red} gibt Hinweis, dass „etwas nicht stimmt“, denn mit nur 3 Messpunkten für y müsste Polynom 2. Grades für Anpassung ausreichen

Interpolation von Prozessparametern - 3D



gelb: Fläche des 3D-Fits; blau: Messwerte; rot: „Fehlerbalken“

Interpolation von Prozessparametern - 3D

Ergebnis für konstanten Messfehler von 0.5 nm:

X^2_{red} = 0.82 bei D = 24 $\rightarrow$ P = 0.71

$\rightarrow$ Messfehler war falsch bestimmt

Kovarianzmatrix

Aus Kovarianzmatrix können Fehler der Fitparameter und deren Korrelation bestimmt werden

Definition: $$\mathcal{C}_{i,j}^{-1} = \frac{1}{2}\frac{\partial}{\partial p_i}\frac{\partial}{\partial p_j}\chi^2$$

Kovarianzmatrix: $$\mathcal{C}_{i,j} = \begin{pmatrix} \boxed{\sigma(p_i)^2} & \mathrm{cov}(p_i,\, p_j) \\ \mathrm{cov}(p_j,\, p_i) & \sigma(p_j)^2 \end{pmatrix}$$

Varianz der Fitvariablen

Korrelationskoeffizient: $$\rho = \frac{\mathrm{cov}(p_i,\, p_j)}{\sigma(p_i)\sigma(p_j)}$$

Kovarianzmatrix

Aus Kovarianzmatrix können Fehler der Fitparameter und deren Korrelation bestimmt werden

Korrelationskoeffizient: $$\rho = \frac{\mathrm{cov}(p_i,\, p_j)}{\sigma(p_i)\sigma(p_j)}$$

Hat man z.B. mit Funktion
$y = A(x - x_0) + y_0$
gefitter, stellt man fest, dass Korrelation von x_0 und y_0 = 1 ist, die Parameter also nicht unabhängig voneinander sind

Konfidenzbereich

Konfidenz = Vertrauen

Konfidenzlevel α ist Sicherheit mit der sich der korrekte Fitwert innerhalb des Konfidenzbereichs befindet

Konfidenzbereich

Konfidenzbereich mit Konfidenzlevel α:

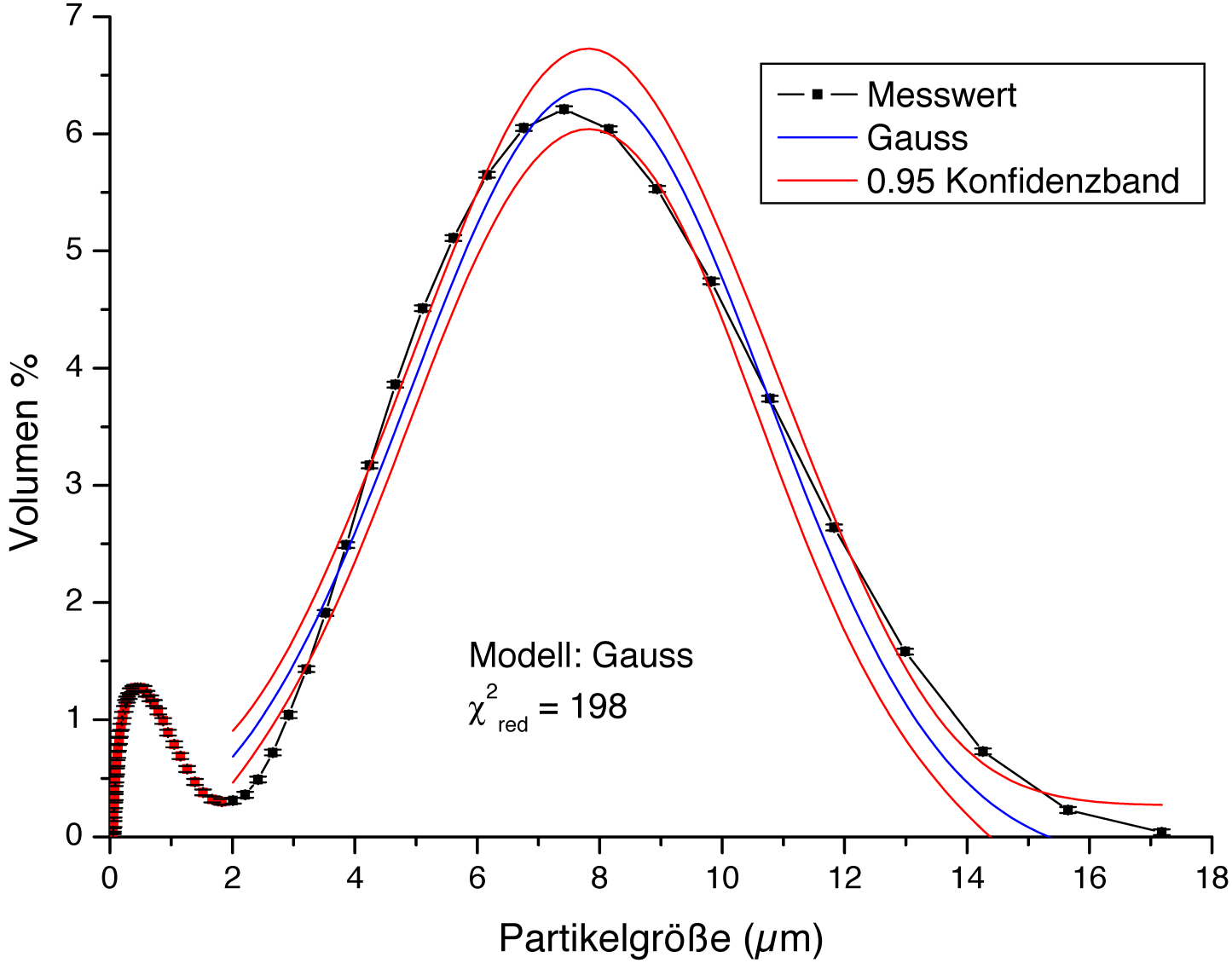
$$Y_{\mathrm{fit}} \pm \mathrm{t}\left(\frac{1-\alpha}{2};\, D\right)\sqrt{\chi^2_{\mathrm{red}}\left(1+\mathcal{J}^{-1}\mathcal{C}\mathcal{J}\right)}$$

$\mathrm{t}(\mathrm{A};\mathrm{B})$:= A-Quantil der t-Verteilung mit D Freiheitsgraden

C := Kovarianzmatrix

J := Jacobi-Matrix der Fitvariablen

Konfidenzbereich



Nützliche Links

Fitten mit Origin 7

http://wwwhep.physik.uni-freiburg.de/fp/origin/origin-einfuehrung.pdf

Fitten mit Origin 8

http://wwwhep.physik.uni-freiburg.de/fp/origin/origin8-einfuehrung.pdf

Danke für die Aufmerksamkeit