Let us consider the model output $h(x) \in \mathbb{R}^m$ with

$$h_i(x) := x^2 t_i + x \text{ for all } i = 1, \dots, m \tag{1}$$

for parameter $x \in \mathbb{R}$ and measurement points $t_1, \dots, t_m \in \mathbb{R}$.

Let us assume that we have a true parameter value

$$\bar{x} := -1 \tag{2}$$

and measurements

$$y \in \mathbb{R}^m \text{ as a realization of } \eta \sim \mathcal{N}(h(\bar{x}), I). \tag{3}$$

For the residual function F and its derivative J, the following holds.

$$F_i(x) := y_i - h_i(x) = y_i - (x^2 t_i + x) \tag{4}$$

$$J_i(x) := -(2xt_i + 1) \tag{5}$$

For the least-squares misfit function ϕ and its derivatives, the following holds.

$$\phi(x) := \frac{1}{2}\|F(x)\|_2^2 \tag{6}$$

$$\nabla\phi(x) = J(x)^T F(x) = -\sum_{i=1}^{m}(2xt_i + 1)(y_i - (x^2 t_i + x)) \tag{7}$$

$$\nabla^2\phi(x) = B(x) + E(x) \tag{8}$$

$$B(x) := J(x)^T J(x) = \sum_{i=1}^{m}(2xt_i + 1)^2 \tag{9}$$

$$E(x) := \sum_{i=1}^{m} F_i(x)\frac{\partial J_i(x)}{\partial x} = -2\sum_{i=1}^{m}(y_i - (x^2 t_i + x))t_i \tag{10}$$

Lets now assume that we have two measurements ($m := 2$) with

$$t := \begin{pmatrix} 0 \\ 2 \end{pmatrix} \text{ and } y := \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{11}$$

The minimizer of the least-squares misfit function with the measurements y is then

$$x_* := 0, \tag{12}$$

since the following applies.

$$\nabla\phi(x_*) = -\sum_{i=1}^{m} y_i = 0 \tag{13}$$

$$B(x_*) = m = 2 \tag{14}$$

$$E(x_*) = -2\sum_{i=1}^{m} y_i t_i = 4 \tag{15}$$

Hence, the estimated model trajectory is

$$h(x_*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{16}$$

Furthermore, it is a large residual problem, since

$$\varrho(B(x_*)^{-1}E(x_*)) = |B(x_*)^{-1}E(x_*)| = 2 > 1. \tag{17}$$

The perturbed measurements, reflected by the estimated model trajectory, are

$$\hat{y} := \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{18}$$

In addition, it holds,

$$h(\bar{x}) = h(-1) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \hat{y} \tag{19}$$

This means, $\bar{x}$ is the minimizer of the least-squares misfit function with the measurements $\hat{y}$.

Why should the minimizer of the misfit function with the reflected measurements not be used as an estimate of the true parameter value? In this example, the minimizer even coincides with the true parameters.