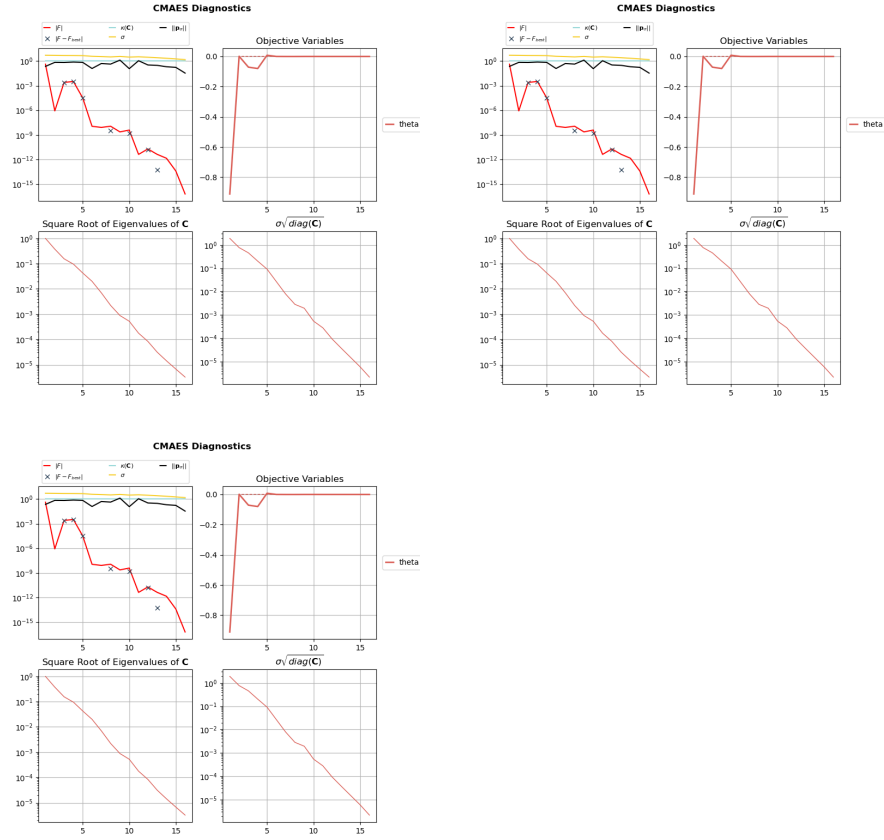


# HPCSE2: Exercise 2

Kalman Joseph Szenes

March 26, 2022

## 1 Task 1



(a) Example 3

Figure 1: Plots showing Korali output for Example 1, 2 and 3 of Task 1.

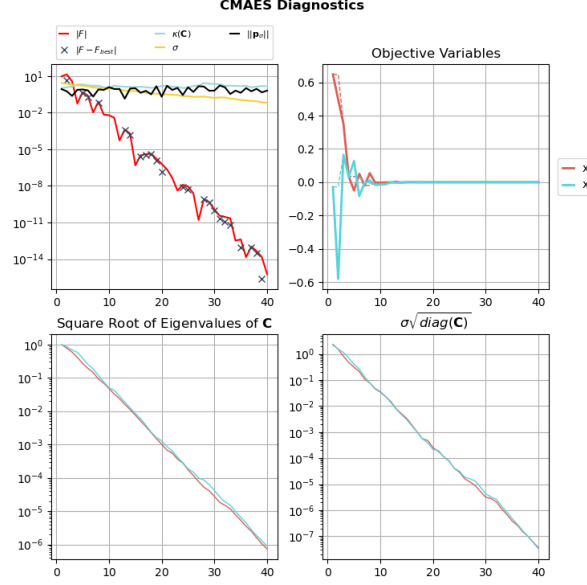


Figure 2: Plot showing Korali output for Task 2.

## 2 Task 2

Korali finds the following minimum at  $(x_{min}, y_{min}) = (-4.219e - 09, +2.051e - 09)$  with  $f(x_{min}, y_{min}) = -5.721571e - 16$ . This is reaching the value 0.0 with machine precision. This solution is trivial as all terms in  $f$  are squared and hence the minimum is found at the origin with value 0.

When the optimization was run with different number of samples, we observed negative correlation with the number of generation. Hence, if we increase the number of samples, the number of generations decreases.

## 3 Task 3

Bayesian inference can be used for parameter approximation technique and it relies on Bayes' Theorem:

$$P(\theta|data) = \frac{P(data|\theta)P(\theta)}{P(data)} \quad (1)$$

In this equation,  $\theta$  groups together all the parameters that we would like to approximate.

In Bayesian inference, our prior probability distribution of the parameters ( $P(\theta)$ ) is updated and improved as we obtain more data. The prior represents

our best guess for the probability distribution of the parameters before observing the data. In the agnostic case, we simply assign to it a Uniform distribution in a reasonable range (i.e.  $P(\theta) \sim U[\theta_{min}, \theta_{max}]$ ). The prior is multiplied by a likelihood function ( $P(data|\theta)$ ) which represents the probability of obtaining the observed data given the parameters. We assign to it the parametrised model with which we would like to describe the data. For this problem, as we assume that there is an error in the measured data, we will model the likelihood using a normal distribution with a mean given by our assumed model ( $P(data|\theta) \sim N(k_1x + k_3x^3 + k_5x^5, \sigma^2)$ ). Finally, to ensure that our posterior is well behaved, we need to normalise the numerator by the appropriate factor ( $P(data) = \int_{\theta} P(data|\theta)P(\theta)d\theta$ ) which is called the evidence. In Bayesian inference, this factor is not important, as we want to find the maximum a posteriori with respect to the parameters. Since the evidence does not depend on the parameters and is hence simply a constant, it will not affect the maximum.

In this concrete case, equation 1 can be written:

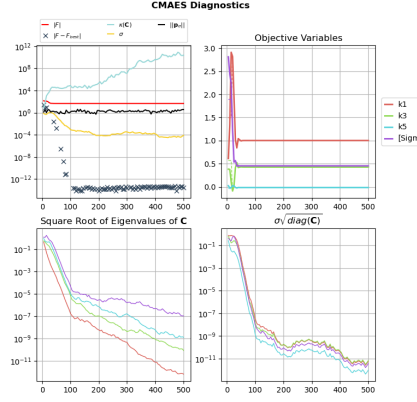
$$P(\theta|data) \sim N(k_1x + k_3x^3 + k_5x^5, \sigma^2) * U[\theta_{min}, \theta_{max}] \quad (2)$$

We chose as range  $[-5, 5]$  for the parameters and  $[0, 5]$  for the standard deviation. We can find the maximum by taking the derivative with respect to the parameters and setting it to zero. We would then need to make sure that the stationary point obtained is maximum not a minimum.

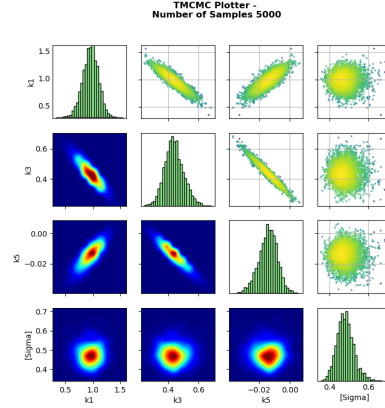
As illustrated in Figure 3 (a), the maximum likelihood exhibits smooth convergence for all the parameters

As illustrated in Figure 3 (b),  $k_1$  and  $k_5$  are negatively correlated with  $k_3$  and  $k_5$  is positively correlated with  $k_3$ . The standard deviation does not seem to be correlated with the other parameters.

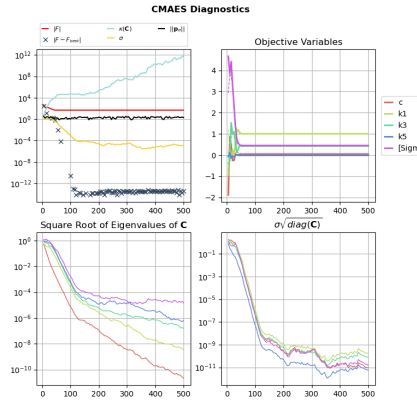
As illustrated in Figure 3 (c), the maximum likelihood converges smoothly with the added offset as well. However, as illustrated in Figure 3 (d), adding the offset does not seem to significantly decrease the correlation between the other parameters. It is hence preferable to use a simpler model (without offset).



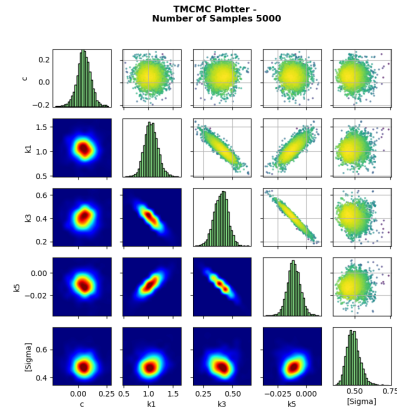
(a) CMA-ES optimization.



(b) TMCMC optimization.



(c) CMA-ES optimization with offset.



(d) TMCMC optimization with offset.

Figure 3: Plots for Task 3.