Introduction

Our mission is to reconstruct the last 30 seconds of a ECG signal given some other complete ECG signals,

This is done using two algorithms:
RLS and **Adam**

RLS (Recursive Least Squares)

Notation:

λ - Forgetting factor
θ - Filter parameters
p - Filter length
P - Covariance matrix of parameters
K - Kalman gain
h - observations
x - input signals

1. Uses the Weighted Squared Error as “error function” with the aim to minimize it.
2. The initial guess is important.
3. Updates the filter coefficients for every iteration.

$$\hat{x}_T[n] = \sum_{k=0}^{M-1} a_k[n]\, x_1[n-k] + \sum_{k=0}^{N-1} b_k[n]\, x_2[n-k]$$

$$\hat{x}[n] = h[n]^T \theta[n-1]$$

$$p = M + N$$

$$h \in \mathbb{R}^{px1}$$

$$P \in \mathbb{R}^{pxp}$$

$$K \in \mathbb{R}^{px1}$$

$$\lambda \in \mathbb{R}$$

Notation:

μ, α - Step size
β - Decay rate
h - Filter parameters
m - Momentum
v - RMSprop **(2nd momentum)**

Adam

Adam is a algorithm that combines the features of two gradient descent algorithms: **Momentum** and **RMSprop**

Momentum

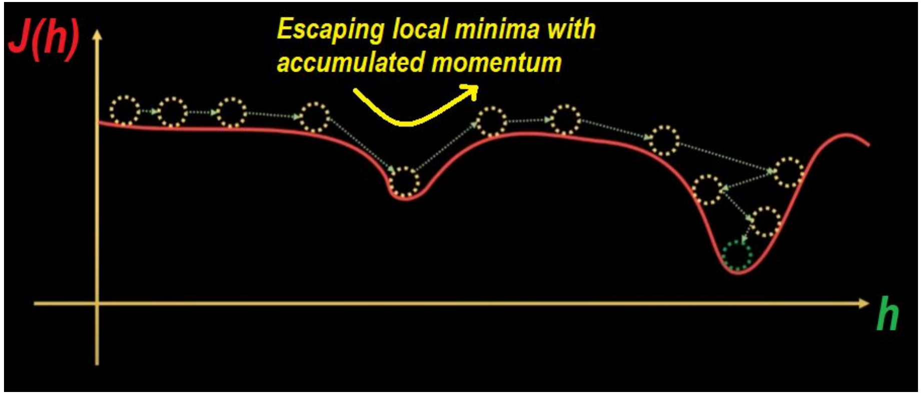
$$m_n = \beta m_{n-1} - \mu g_n$$

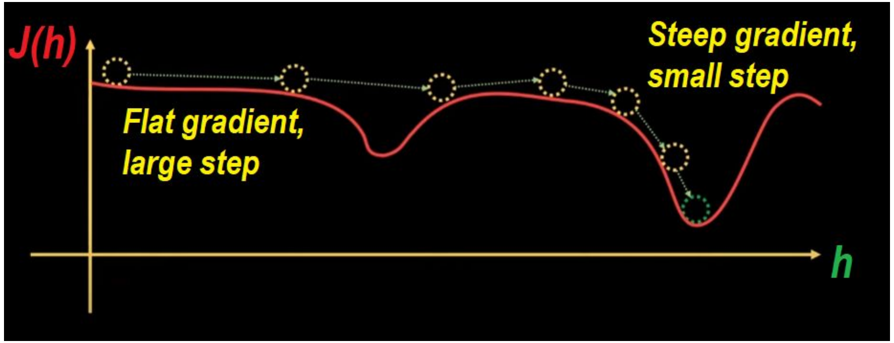
$$h[n] = h[n-1] + m_n$$

RMSprop

$$v_n = \beta v_{n-1} + (1-\beta) g_n^2$$

$$h[n] = h[n-1] - \frac{\mu}{\epsilon + \sqrt{v_n}} g_n$$





Pictures from: https://youtu.be/NE88eqLngkg?si=AXPjzhEFQgcr3aSU

Adam

Adam is a algorithm that combines the features of two gradient descent algorithms: **Momentum** and **RMSprop**

Notation:

μ, α - Step size
β - Decay rate
(ε - Numerical stability)
h - Filter parameters
p - Filter length
m - Momentum
m_hat - Momentum unbiased
v - RMSprop **(2nd momentum)**
v_hat - RMSprop unbiased

Momentum

$$m_n = \beta m_{n-1} - \mu g_n$$

$$h[n] = h[n-1] + m_n$$

RMSprop

$$v_n = \beta v_{n-1} + (1-\beta) g_n^2$$

$$h[n] = h[n-1] - \frac{\mu}{\epsilon + \sqrt{v_n}} g_n$$

Adam

$$
\begin{aligned}
&\textbf{for } n = 1 \textit{ to } N \textbf{ do}\\
&\quad \boldsymbol{m}_n = \beta_1 \boldsymbol{m}_{n-1} + (1-\beta_1)\boldsymbol{g}_n\\
&\quad \boldsymbol{v}_n = \beta_2 \boldsymbol{v}_{n-1} + (1-\beta_2)\boldsymbol{g}_n^2\\
&\quad \hat{\boldsymbol{m}}_n = \boldsymbol{m}_n / (1-\beta_1^n)\\
&\quad \hat{\boldsymbol{v}}_n = \boldsymbol{v}_n / (1-\beta_2^n)\\
&\quad \boldsymbol{h}[n] = \boldsymbol{h}[n-1] - \alpha\, \hat{\boldsymbol{m}}_n / (\sqrt{\hat{\boldsymbol{v}}_n} + \epsilon)\\
&\textbf{end}
\end{aligned}
$$

$h \in \mathbb{R}^{px1}$ $\quad v \in \mathbb{R}^{px1}$

$m \in \mathbb{R}^{px1}$ $\quad \hat{v} \in \mathbb{R}^{px1}$

$\hat{m} \in \mathbb{R}^{px1}$ $\quad g \in \mathbb{R}^{px1}$

m_hat and **v_hat** compensates for Adams initial **bias towards zero,** originating from vectors **h**, **m** and **v** being initialised with zeros

Adam

Notation:

μ, α - Step size
β - Decay rate
(ε - Numerical stability)
h - Filter parameters
p - Filter length
m - Momentum
m_hat - Momentum unbiased
v - RMSprop **(2nd momentum)**
v_hat - RMSprop unbiased
y - observations
x - input signals

For Adam optimizer we assume that the estimated signal can be written as a linear combinations of **x_1 (ECG V)** and **x_2 (ECG AVR**)

$$\hat{x}_T[n] = \sum_{k=0}^{M-1} a_k[n]\, x_1[n-k] + \sum_{k=0}^{N-1} b_k[n]\, x_2[n-k]$$

$$\hat{x}[n] = h[n-1]^T y[n]$$

$$h[n] = \begin{bmatrix} a_0[n] \\ a_1[n] \\ \dots \\ a_{M-1}[n] \\ b_0[n] \\ b_1[n] \\ \dots \\ b_{N-1}[n] \end{bmatrix} \qquad y[n] = \begin{bmatrix} x_1[n] \\ x_1[n-1] \\ \dots \\ x_1[n-(M-1)] \\ x_2[n] \\ 2_2[n-1] \\ \dots \\ x_2[n-(N-1)] \end{bmatrix}$$

Initialise the filter:

- Filter length, how many parameters we want for the inputs (M and N)
- Select the Adam parameters: α, β1, β2, ε
- Initialize the vectors h, m, v with zeros
- Define a cost function f(h(n-1))

> Run the Adam algorithm

Adam algorithm

$$\begin{aligned}
&\textbf{for } n = 1 \textit{ to } N \textbf{ do} \\
&\quad \boldsymbol{m}_n = \beta_1 \boldsymbol{m}_{n-1} + (1-\beta_1)\boldsymbol{g}_n \\
&\quad \boldsymbol{v}_n = \beta_2 \boldsymbol{v}_{n-1} + (1-\beta_2)\boldsymbol{g}_n^2 \\
&\quad \hat{\boldsymbol{m}}_n = \boldsymbol{m}_n/(1-\beta_1^n) \\
&\quad \hat{\boldsymbol{v}}_n = \boldsymbol{v}_n/(1-\beta_2^n) \\
&\quad \boldsymbol{h}[n] = \boldsymbol{h}[n-1] - \alpha\, \hat{\boldsymbol{m}}_n/(\sqrt{\hat{\boldsymbol{v}}_n} + \epsilon) \\
&\textbf{end}
\end{aligned}$$

$v \in \mathbb{R}^{px1}$ $\qquad p = M + N$

$\hat{v} \in \mathbb{R}^{px1}$ $\qquad h \in \mathbb{R}^{px1}$

$g \in \mathbb{R}^{px1}$ $\qquad m \in \mathbb{R}^{px1}$

$\alpha, \epsilon, \beta_1, \beta_2 \in \mathbb{R}$ $\qquad \hat{m} \in \mathbb{R}^{px1}$

RLS Parameters

- Lambda = 1 (forgetting factor)
- N = 100 (steps back used for signal A)
- M = 100 (steps back used for signal B)
  - p = N + M (filter length)

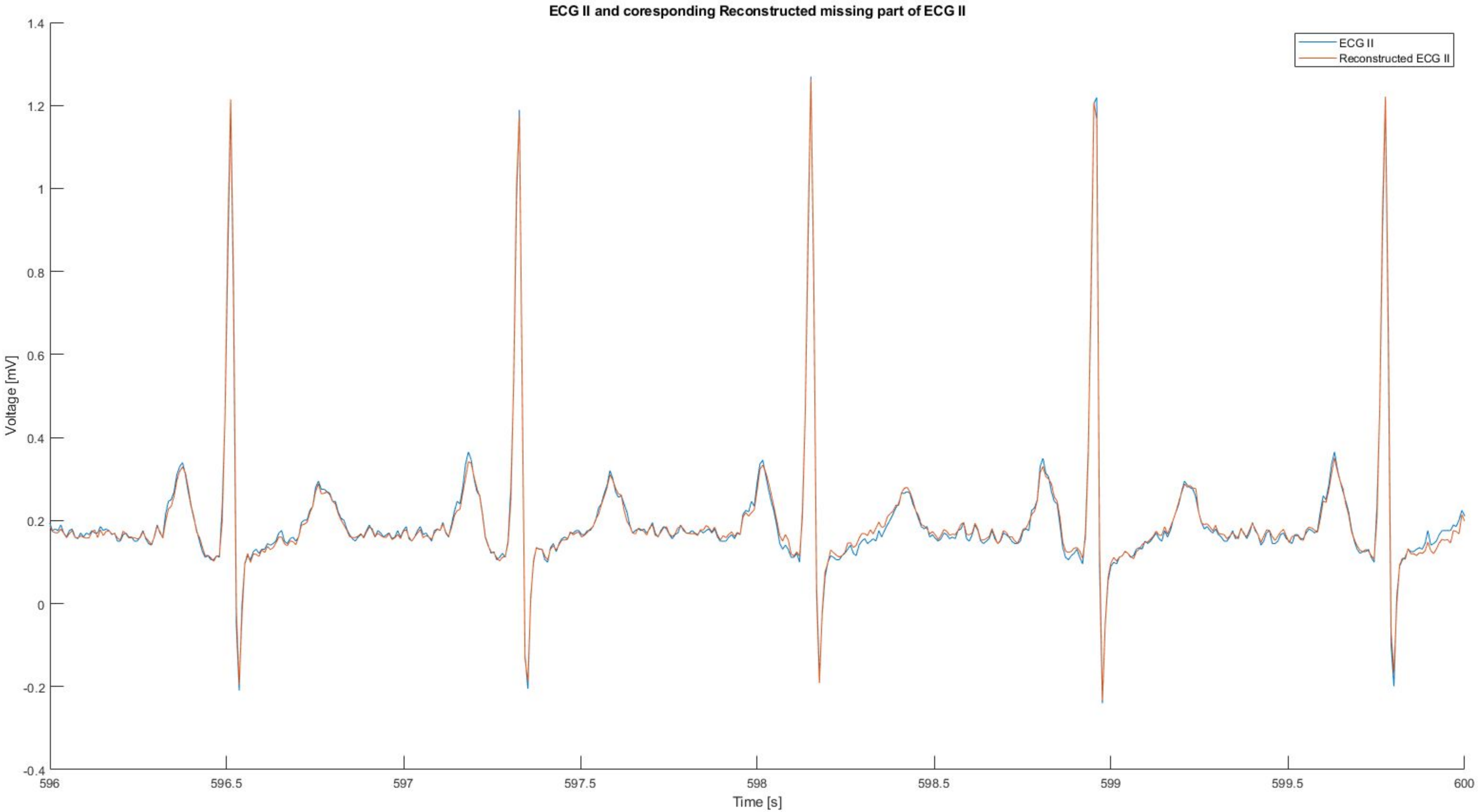
Adam Parameters

- alfa = 0.001 (step size)
- beta1 = 0.9 (decay rates)
- beta2 = 0.999 (decay rates)
- epsilon = 1e-8 (constant for numerical stability)

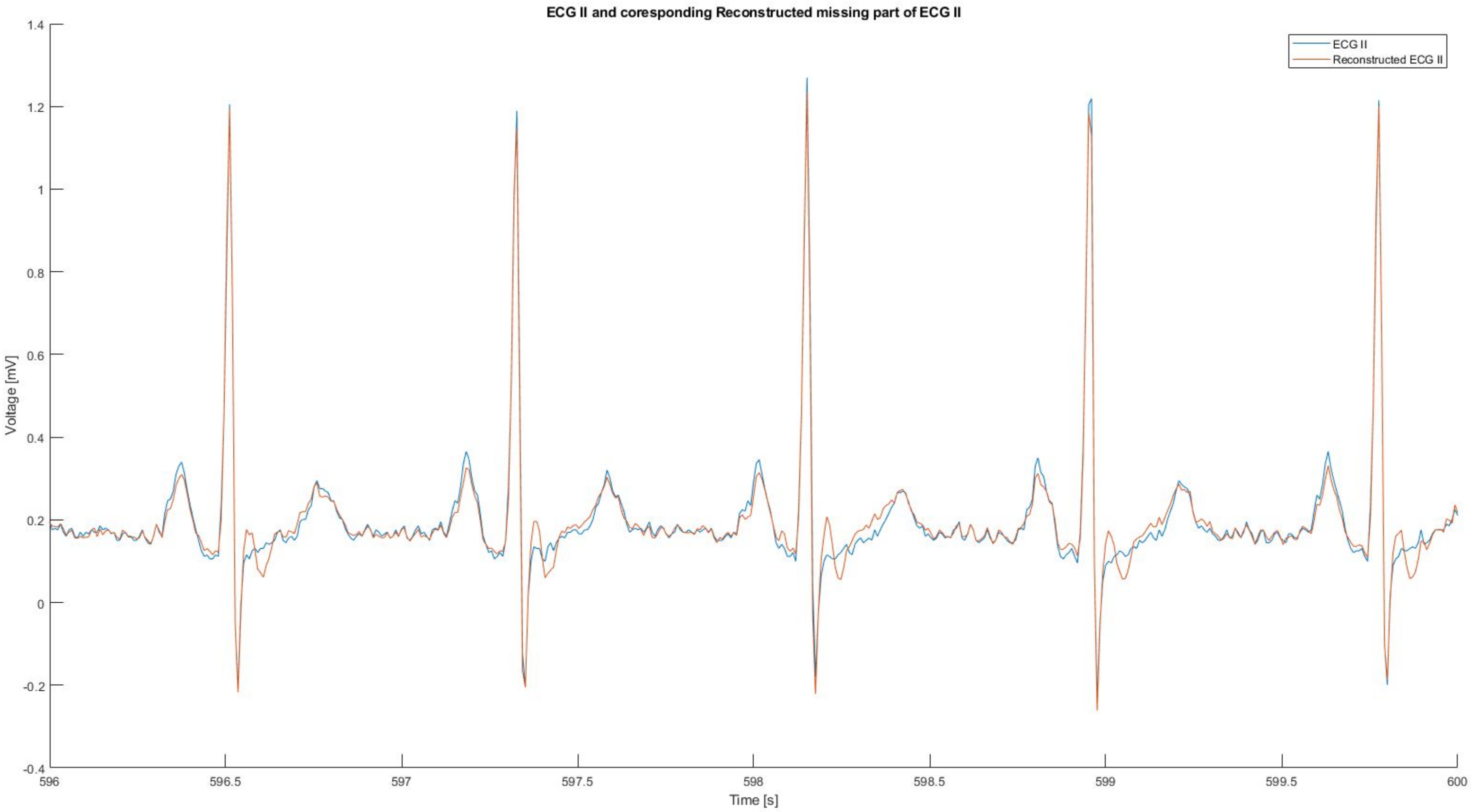
Default settings from the Adam paper

- N = 15 (steps back used for signal A)
- M = 15 (steps back used for signal B)
  - p = N + M (filter length)

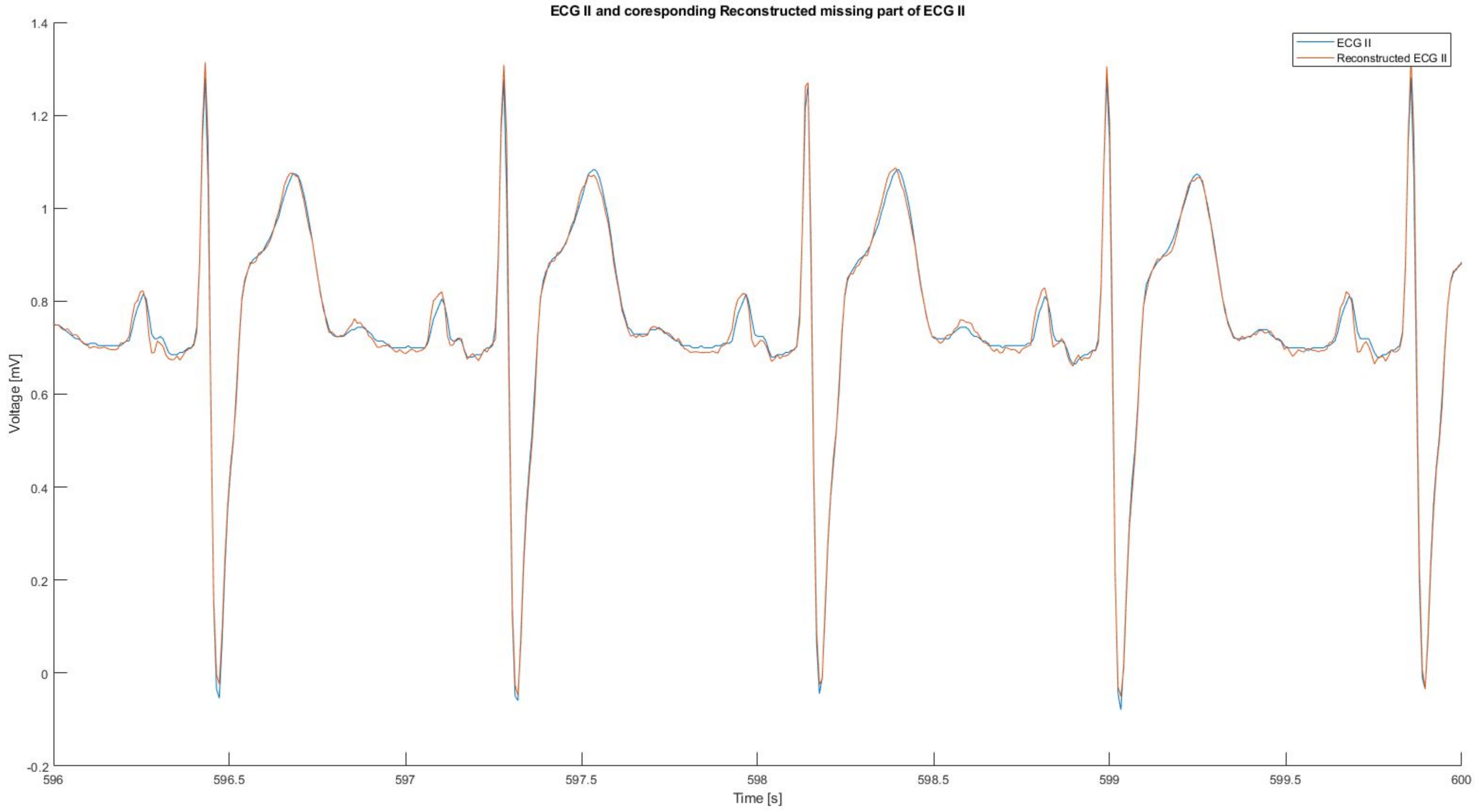
Patient 2 (RLS)



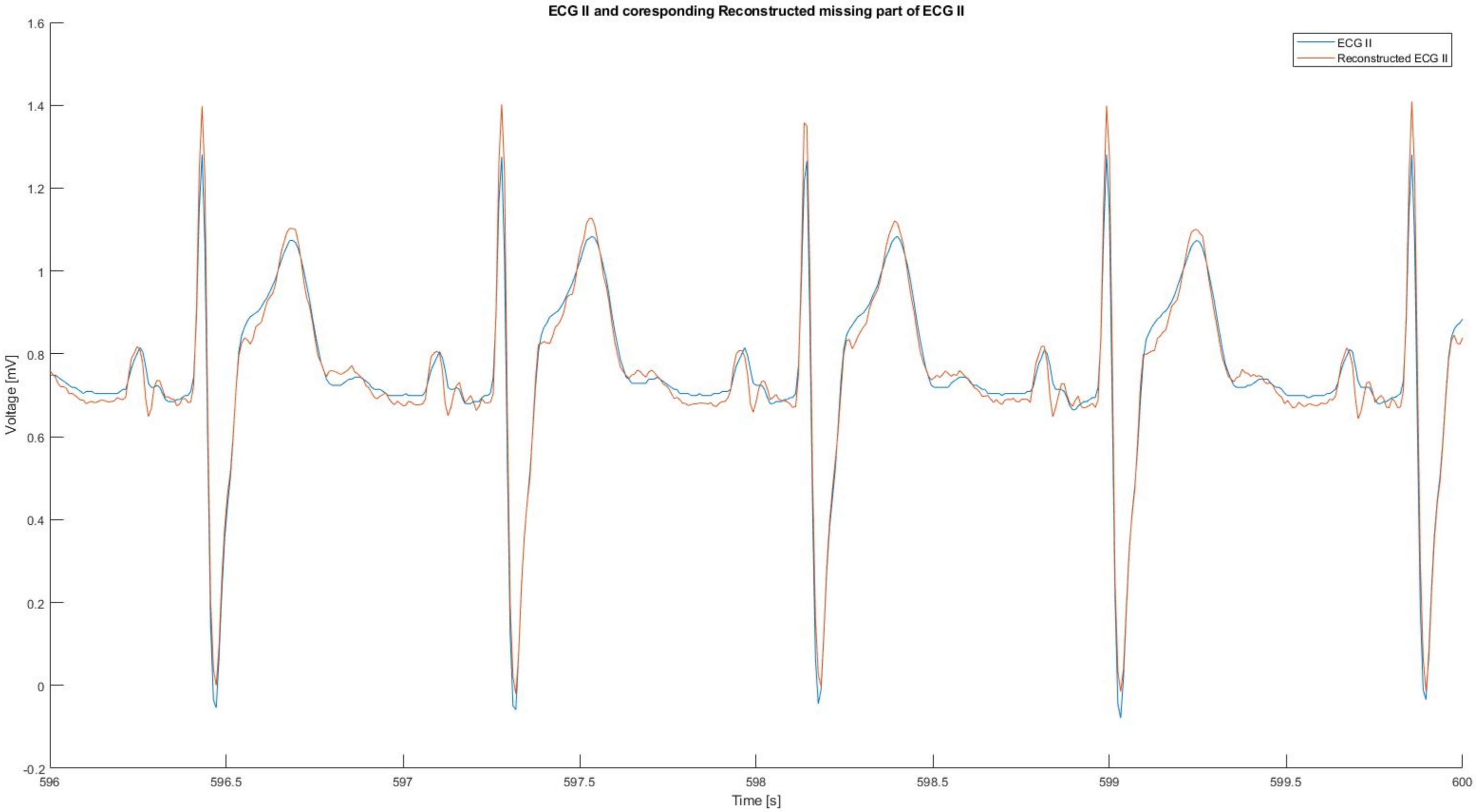
Patient 2 (Adam)



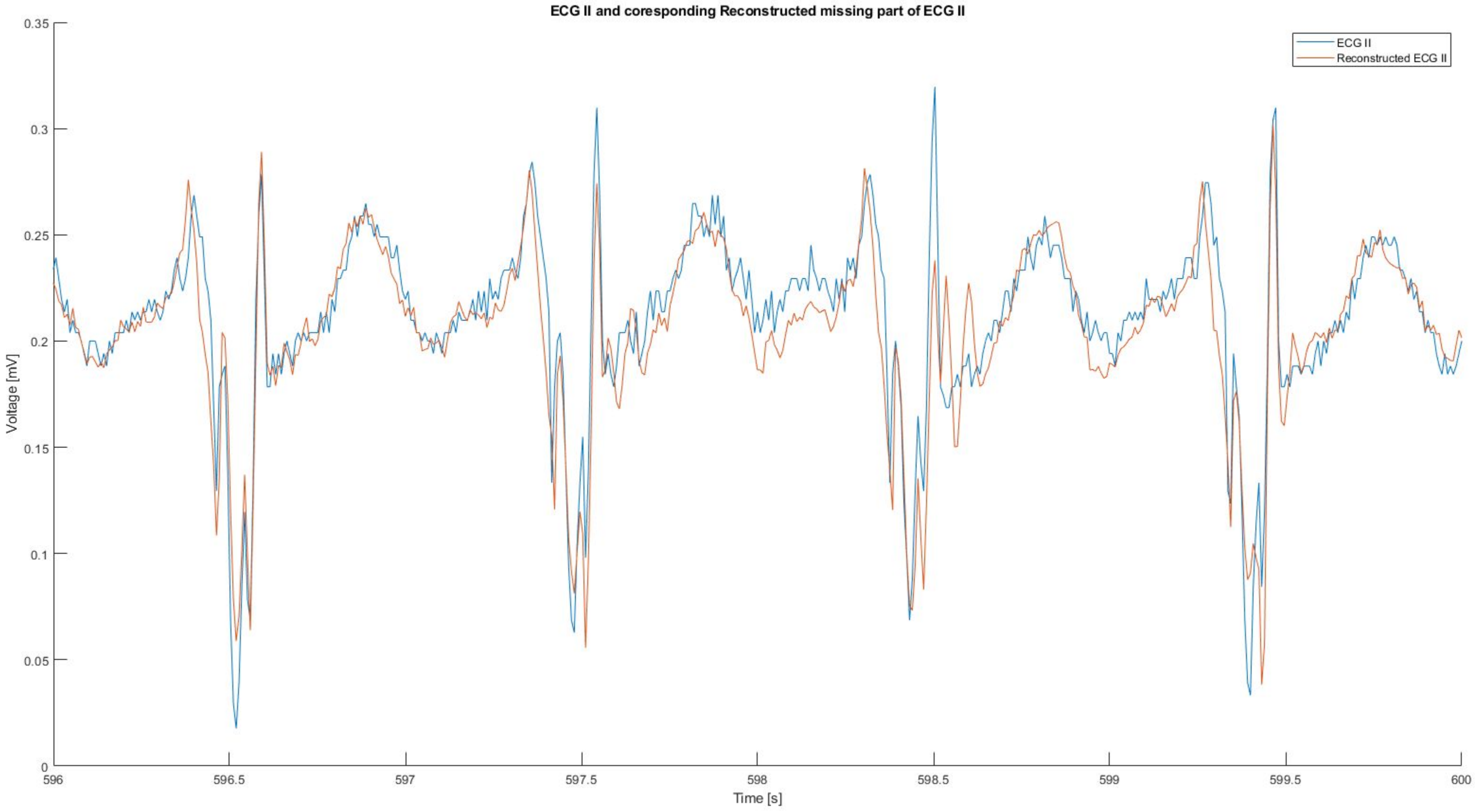
Patient 3 (RLS)



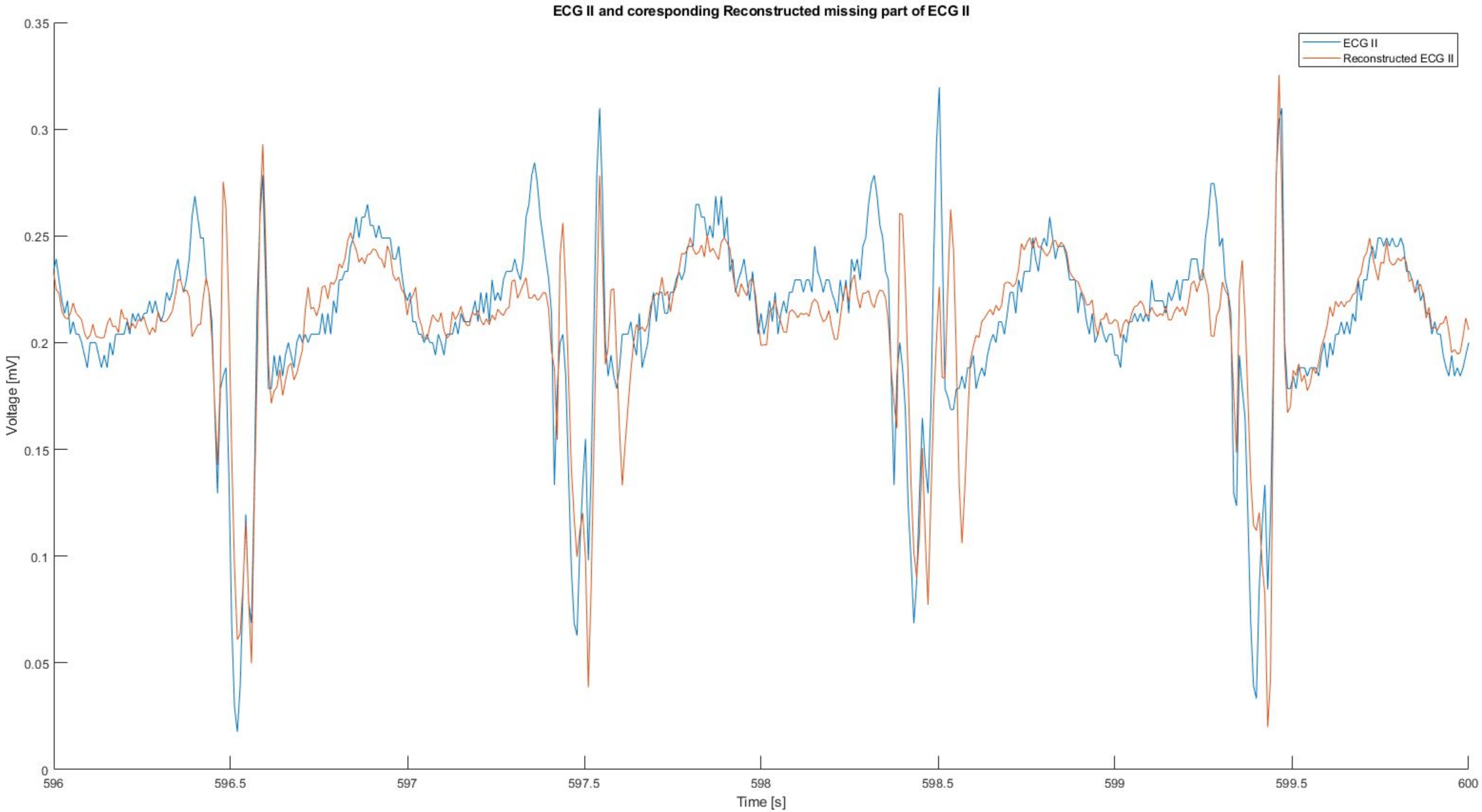
Patient 3 (Adam)



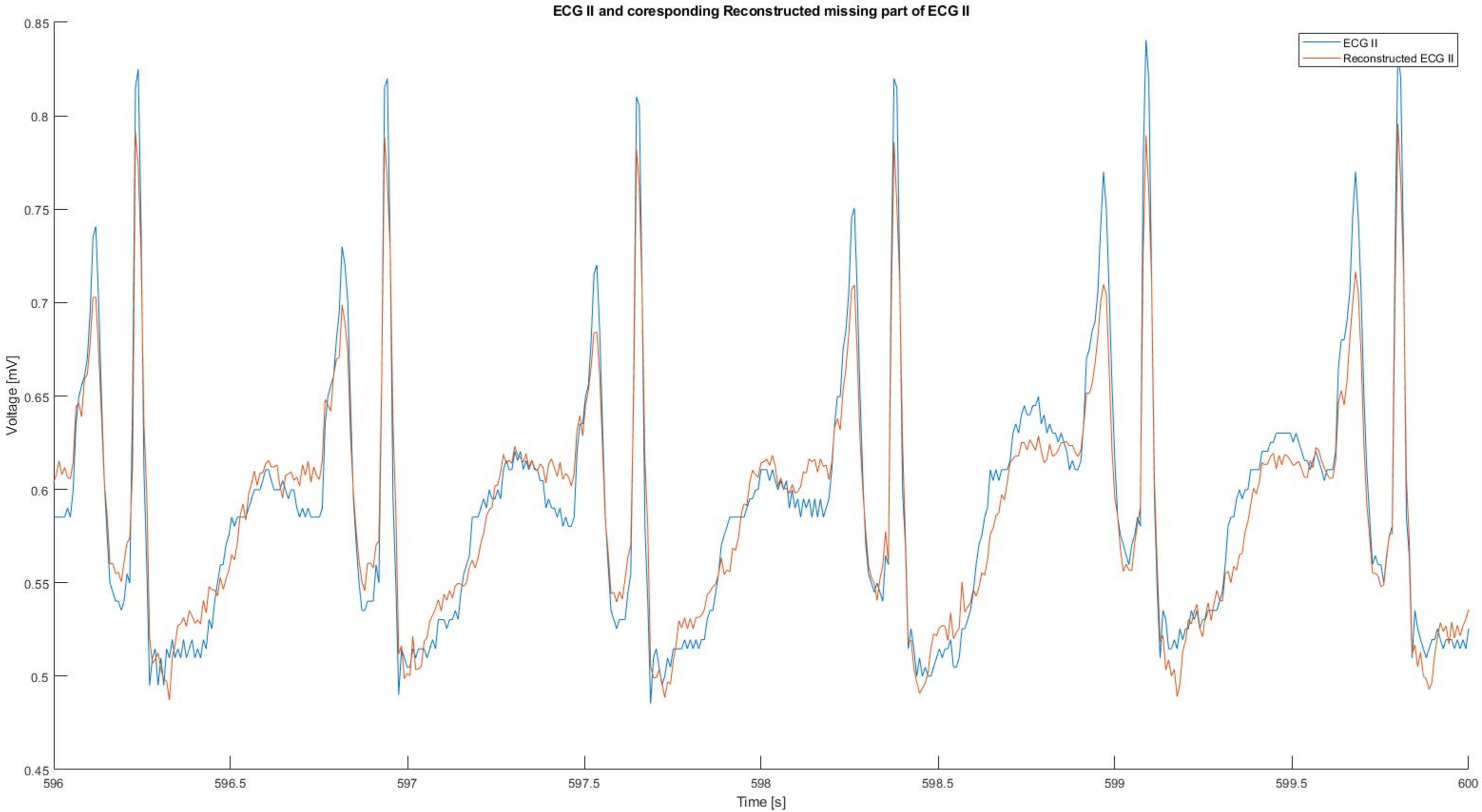
Patient 4 (RLS)



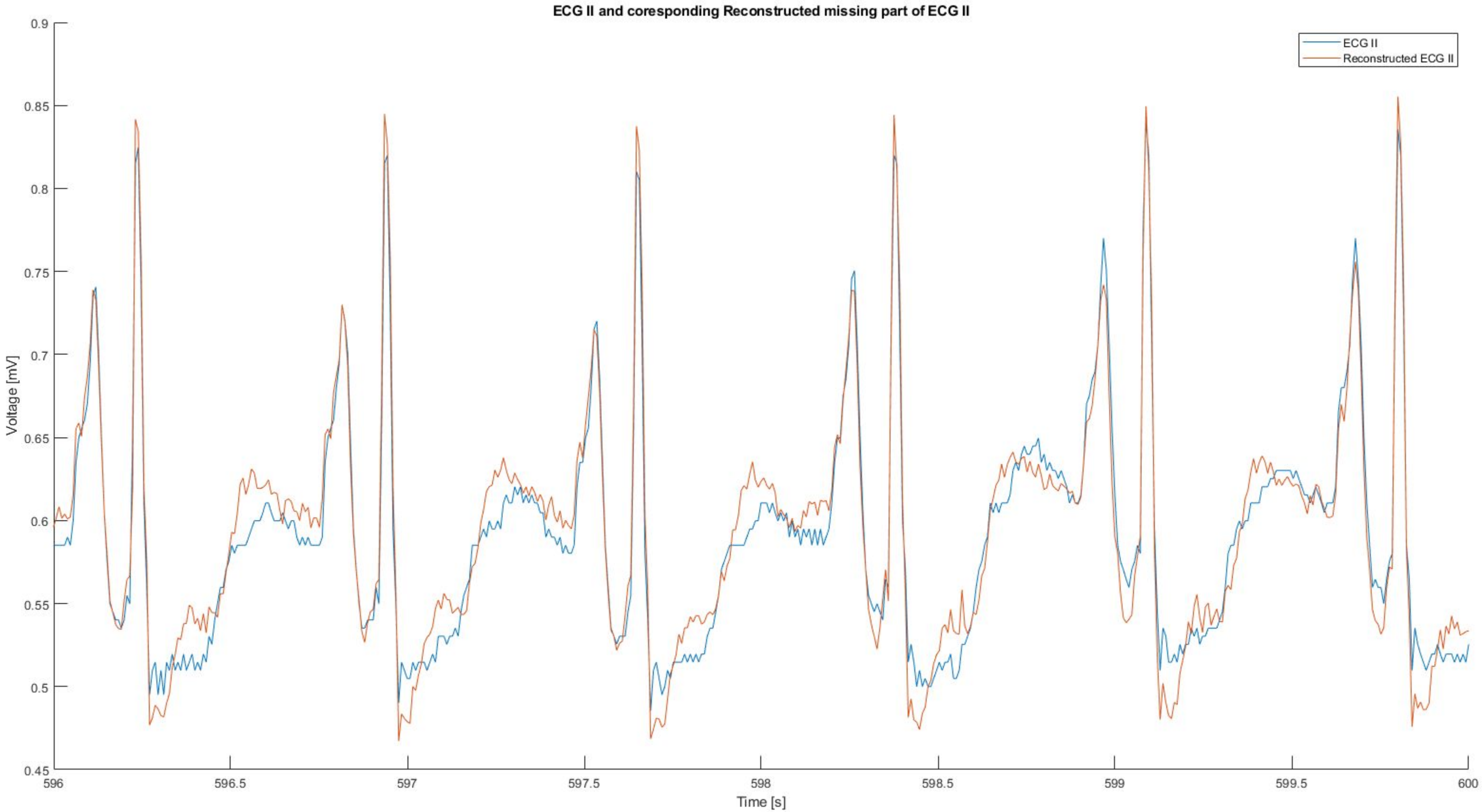
Patient 4 (Adam)



Patient 5 (RLS)



Patient 5 (Adam)



Patient 2-5 discussion

- We expected RLS to work well
- Graphs for patient 2 & 3 shows that the reconstructed signal is very accurate
- Reconstructed signal for patient 4 is not as good, follows pattern but not amplitudes of the real signal, possible reasons:
  - Signal used for reconstruction not being correlated to target signal
  - Parameters not being well enough optimized
- Adam performs well but worse than RLS
- Adam had considerably lower computation time
- Adam had better performance with patient 5

Q1 & Q2 for RLS & LMS for all patients

RLS

Patient	Q1	Q2
1	0.97124	0.98558
2	0.99519	0.9976
3	0.99361	0.99682
4	0.88209	0.93953
5	0.87335	0.95315
6	0.96125	0.98045
7	0.97053	0.98534
8	0.98798	0.99402

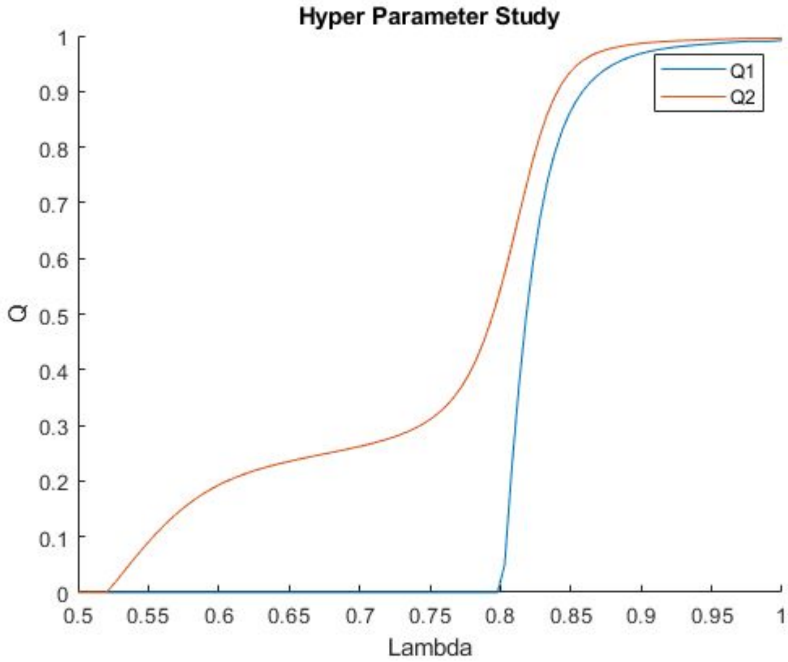
Adam

Patient	Q1	Q2
1	0.94283	0.97458
2	0.97996	0.98999
3	0.97011	0.98524
4	0.73159	0.85539
5	0.91537	0.95941
6	0.91183	0.95516
7	0.96923	0.98454
8	0.98379	0.99195

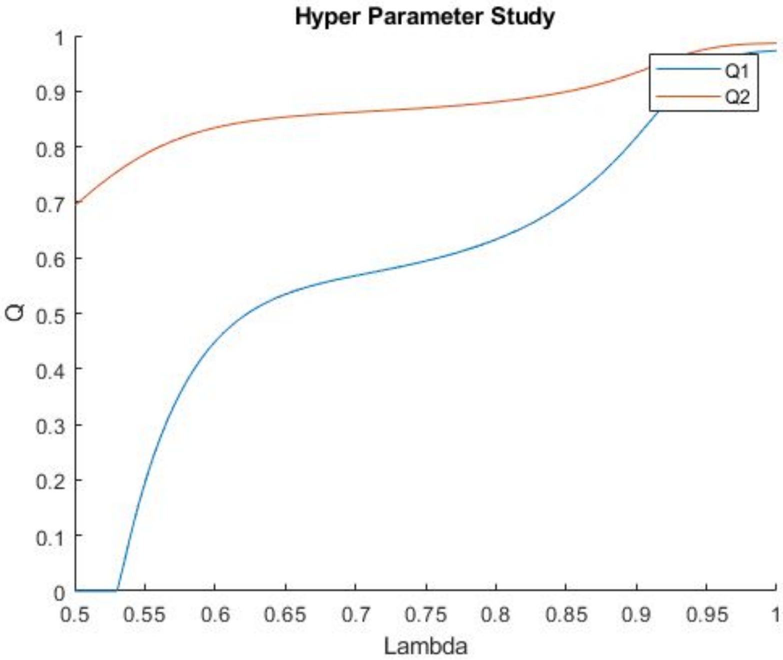
- Only with patient 5 Adam had a better performance.

RLS
Hyper Parameter Study: Forgetting Factor, lambda

Patient 2

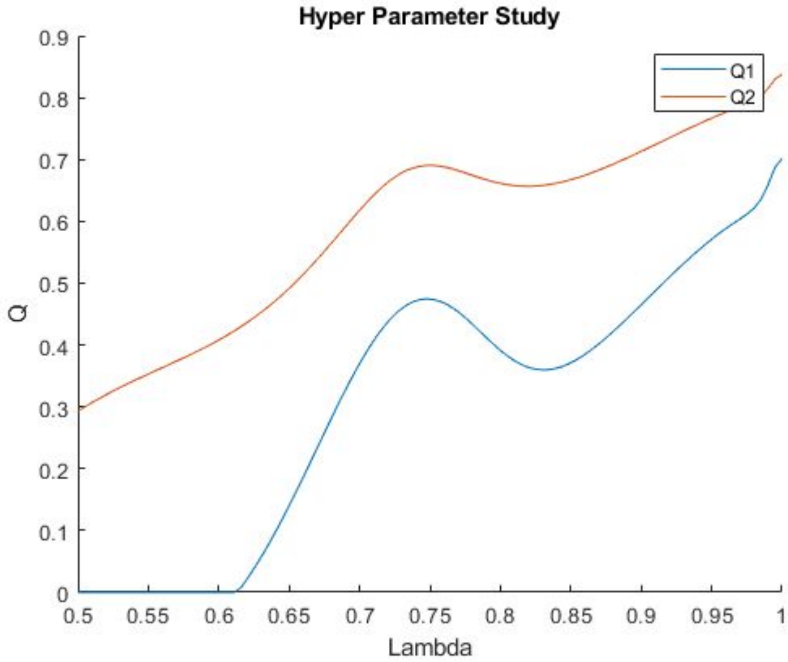


Patient 3

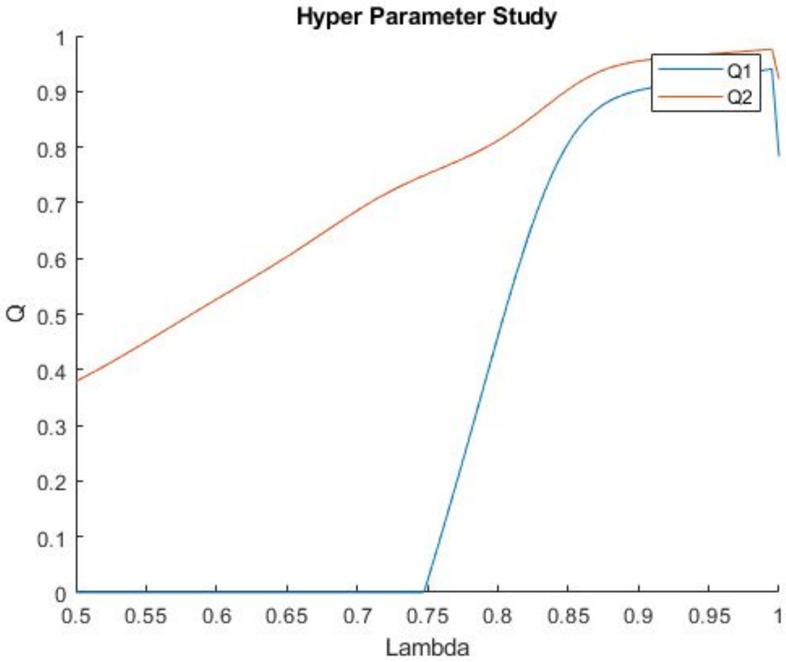


RLS
Hyper Parameter Study: Forgetting Factor, lambda

Patient 4



Patient 5

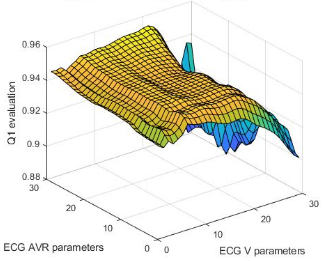


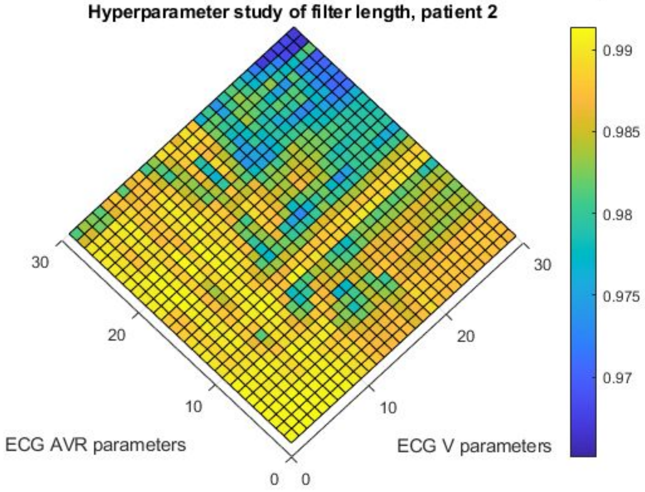
RLS
Hyper Parameter Study: Forgetting Factor, lambda

- Expected lambda values closer to one to work better
    - Plots for patient 2-5 show just that
- We need to take old data into account, hence higher lambda (forgetting factor)

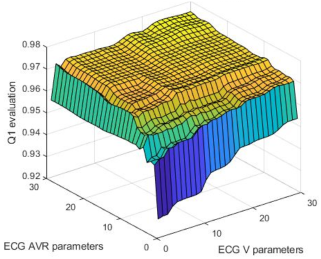
Adam
Hyper Parameter Study: Filter Length

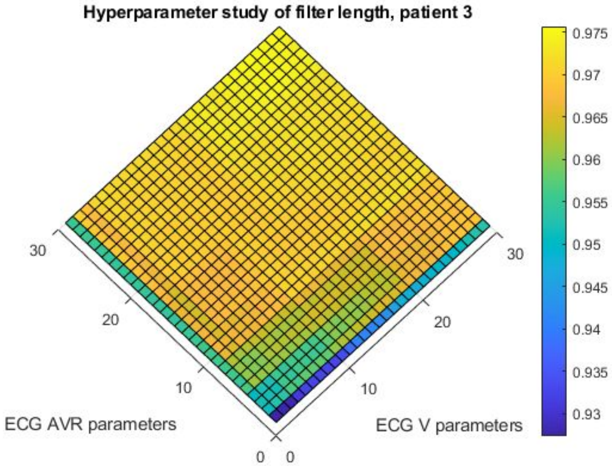
Patient 2, Q1





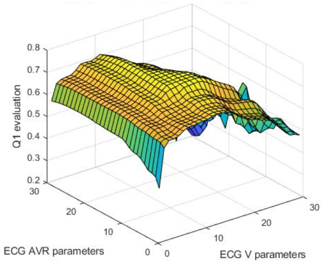
Patient 3, Q1

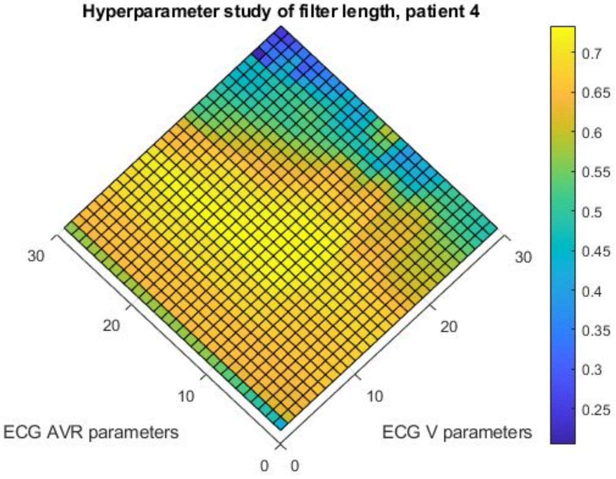




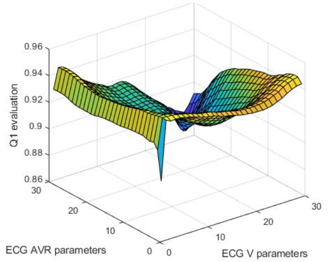
Adam
Hyper Parameter Study: Filter Length

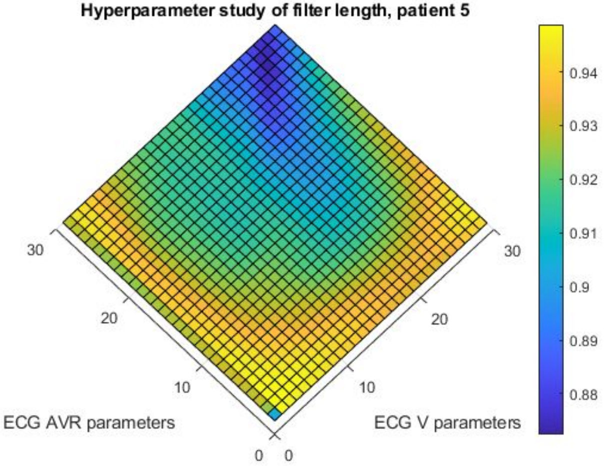
Patient 4, Q1





Patient 5, Q1





Adam
Hyper Parameter Study: Filter Length

- Small filter lengths bad performance ($N = M \approx 1$)
- Large filter length performance decrease ($N = M \approx 30$)
  - $N = M \approx 30$ had considerably worse performance
- For small filter lengths ECG V seems to be more important
- Filter lengths around $N = M \approx [10, 15]$ seems to work the best
- Safest to choose small filter lengths

Discussion

- Both methods accomplish the task, however RLS does it better
- We would choose RLS instead of Adam
    - We can handle the higher computational power
    - We did not have to do anything in real time (for this task)
    - When talking about ECG, we have to put weight on accurate results

Conclusion

- Both methods work well for the given task
- RLS would have been chosen over LMS with Adam optimizer
- In the future we would have optimized all parameters