L1 nonstationary adaptive subtraction

Scott Keating

Oct 16 2015

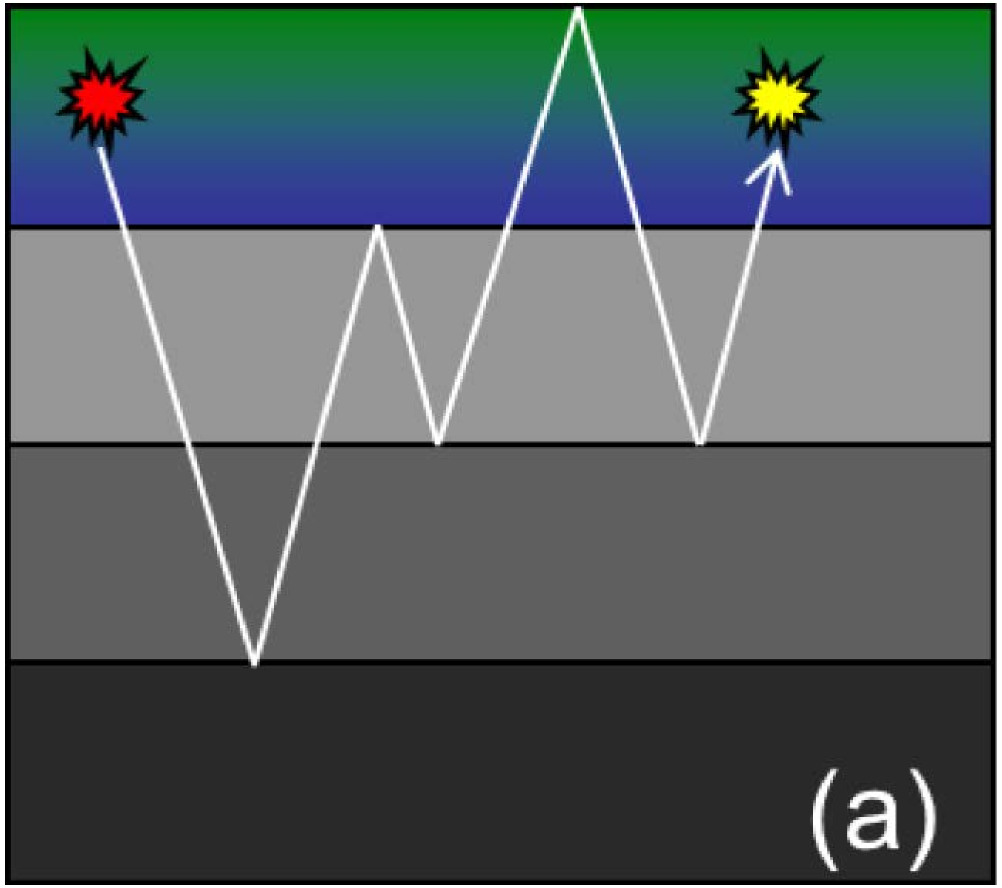
Multiples

- Multiples in seismic data are often undesirable, harming both processing and interpretation.
- There exist computationally cheap multiple removal methods, such as predictive deconvolution and filtering based multiple removal.
- These can fail when confronted with complex geology.
- For complicated internal multiples, inverse scattering multiple prediction can be a preferable method for multiple removal.

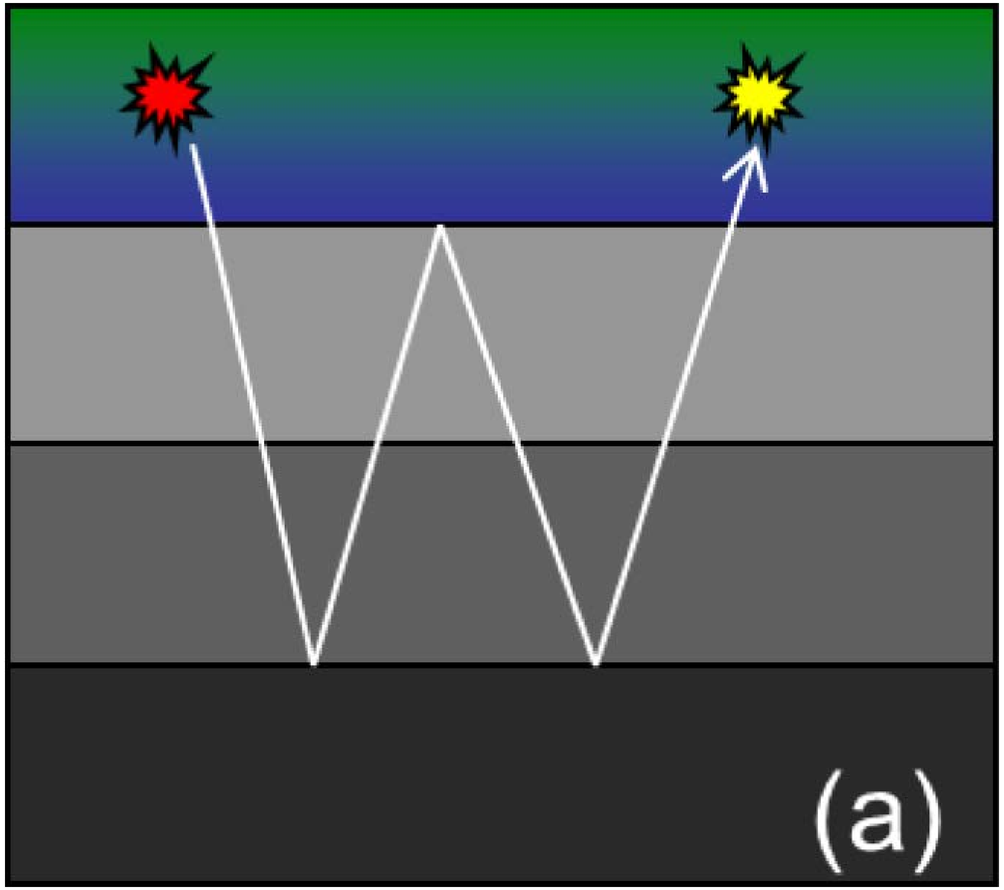
Types of Multiples

- There are two major types of multiples
- Free surface multiples experience at least one downward reflection at the free surface
- Internal multiples are multiples that do not have any reflection at the free surface
- Internal multiples are typically of greater importance in land data, as the near surface scatters away most of the energy associated with free surface multiples

Free Surface Multiple



Internal Multiple



Figures from Weglein and Dragoset, 2005

Inverse Scattering Multiple Prediction

- Inverse scattering multiple prediction generates an internal multiple prediction based on the seismic data alone, and is capable of dealing with very complex geologies.
- CREWES is working on several multidimensional implementations of inverse scattering in a variety of novel domains, such as Jian's work on plane wave domain prediction and Kris' work on space-time domain prediction.

Inverse Scattering Multiple Prediction

- In scattering theory, an observed wavefield, G_s , can expressed in a Born series, an infinite series in terms of reference wavefield, G_0 , and a perturbation from the reference model, V. (Weglein et. Al., 2003)

$$G_s = G_0VG_0 + G_0VG_0VG_0 + G_0VG_0VG_0VG_0 + \ldots$$

- In inverse scattering, we rearrange this equation to give a perturbation, V, as an infinite series in terms of reference wavefield G_0 and observed data D,

$$V = V_1 + V_2 + V_3 + \ldots$$

where V_n is the term which is nth order in D

Inverse Scattering Multiple Prediction

- By equating equal orders in D, we obtain a recursive relation, where

$$D = G_0 V_1 G_0$$

$$G_0 V_2 G_0 = -G_0 V_1 G_0 V_1 G_0$$

and so on.

- Inverse scattering multiple prediction works by identifying the subset of this series which contributes to the generation of internal multiples.

Inverse Scattering Multiple Prediction

- Inverse scattering multiple prediction is in practice generated using a truncation of an infinite series, and so is not exact.
- Additionally, theoretically necessary preprocessing steps such as deconvolution and deghosting are often neglected, introducing additional errors in both amplitude and phase (Pan, 2015).

Adaptive Subtraction

- In order to remove these predicted multiples, we need to match them to the observed multiples and subtract them.
- This matching is typically done by applying a filter to the predicted multiple.
- As we do not know the correct multiple, we need to determine some way of choosing the filter we apply.

Least squares adaptive subtraction

- One of the simplest methods of adaptive subtraction is to choose the filter which gives the multiple prediction that minimizes the total energy in the data. (Verschuur et. al. 1992)
- This hinges on the assumption that multiples and primaries do not overlap in the data, and so eliminating the multiples yields the minimum energy.
- Minimizing the energy is equivalent to minimizing the L2 norm, given by

$$\sum_{i=0}^{t} r_i^2$$

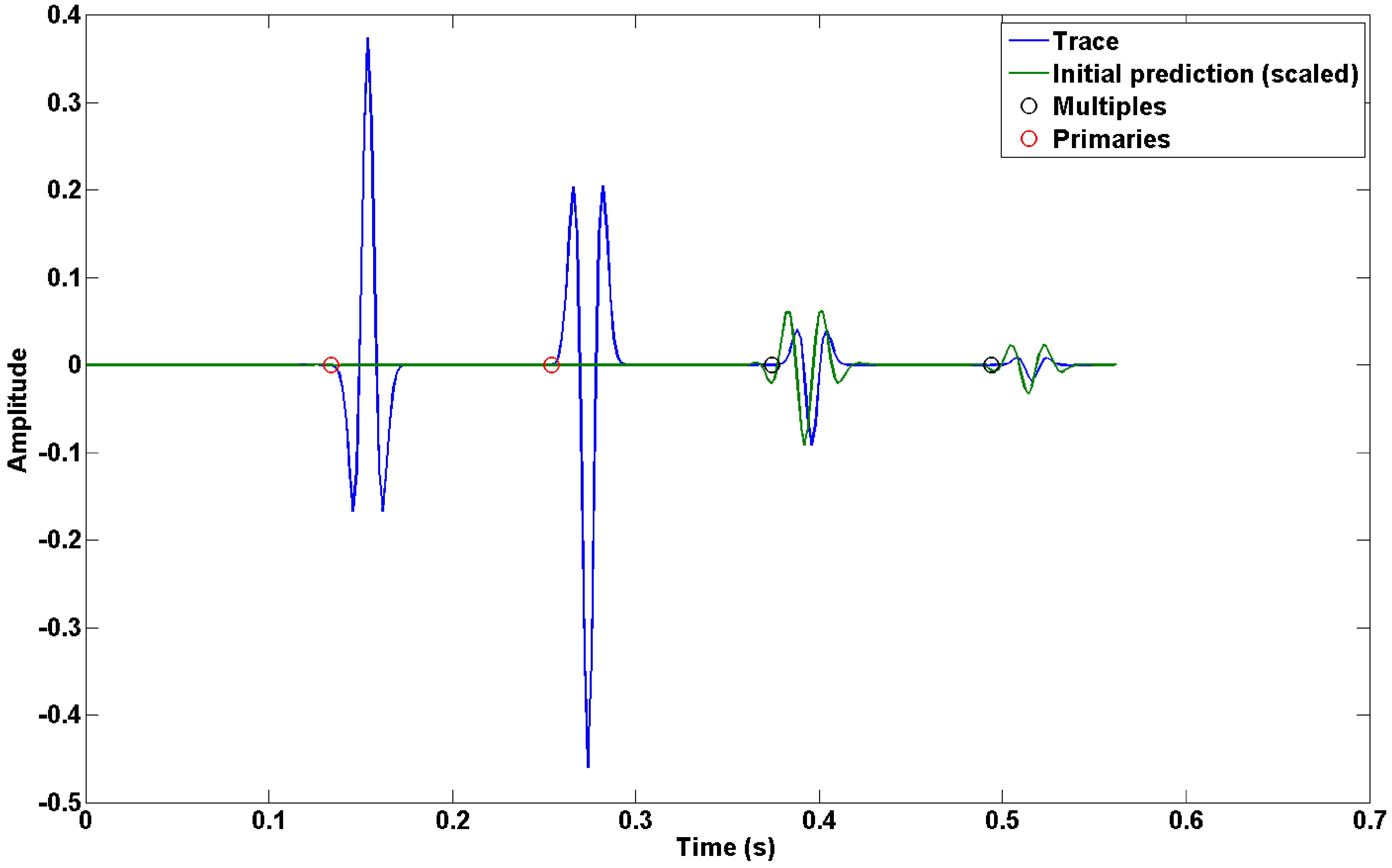
where r_i is the data after subtraction

Least squares adaptive subtraction

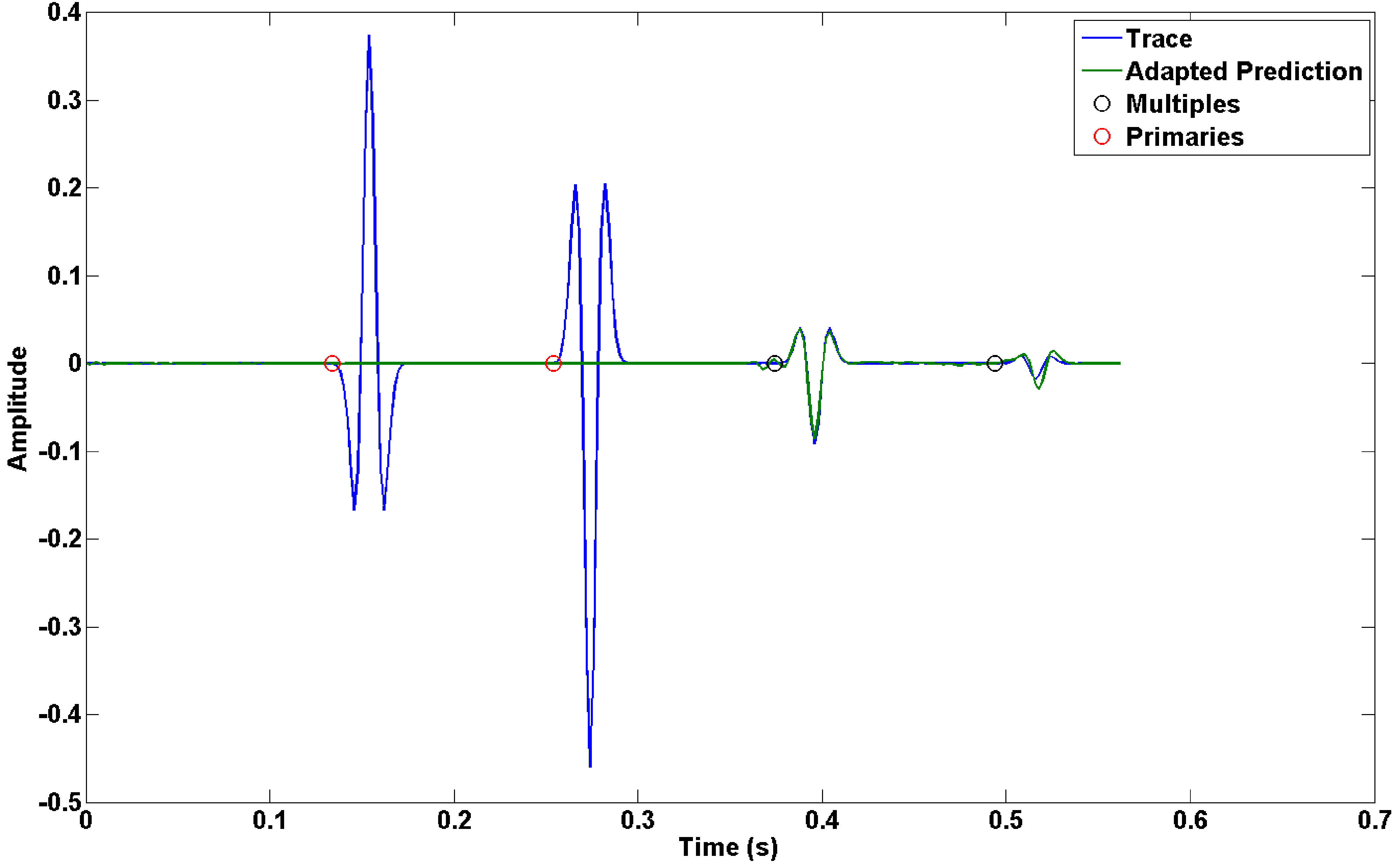
- To find the filter, we wish to solve $Mf = d$, where M is the matrix representing convolution with the predicted multiple trace, d is the data trace, and f is the filter

- We do this by least squares:

$$f = (M^T M)^{-1} M^T d$$



Trace
Initial prediction (scaled)
Multiples
Primaries
Amplitude
Time (s)
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

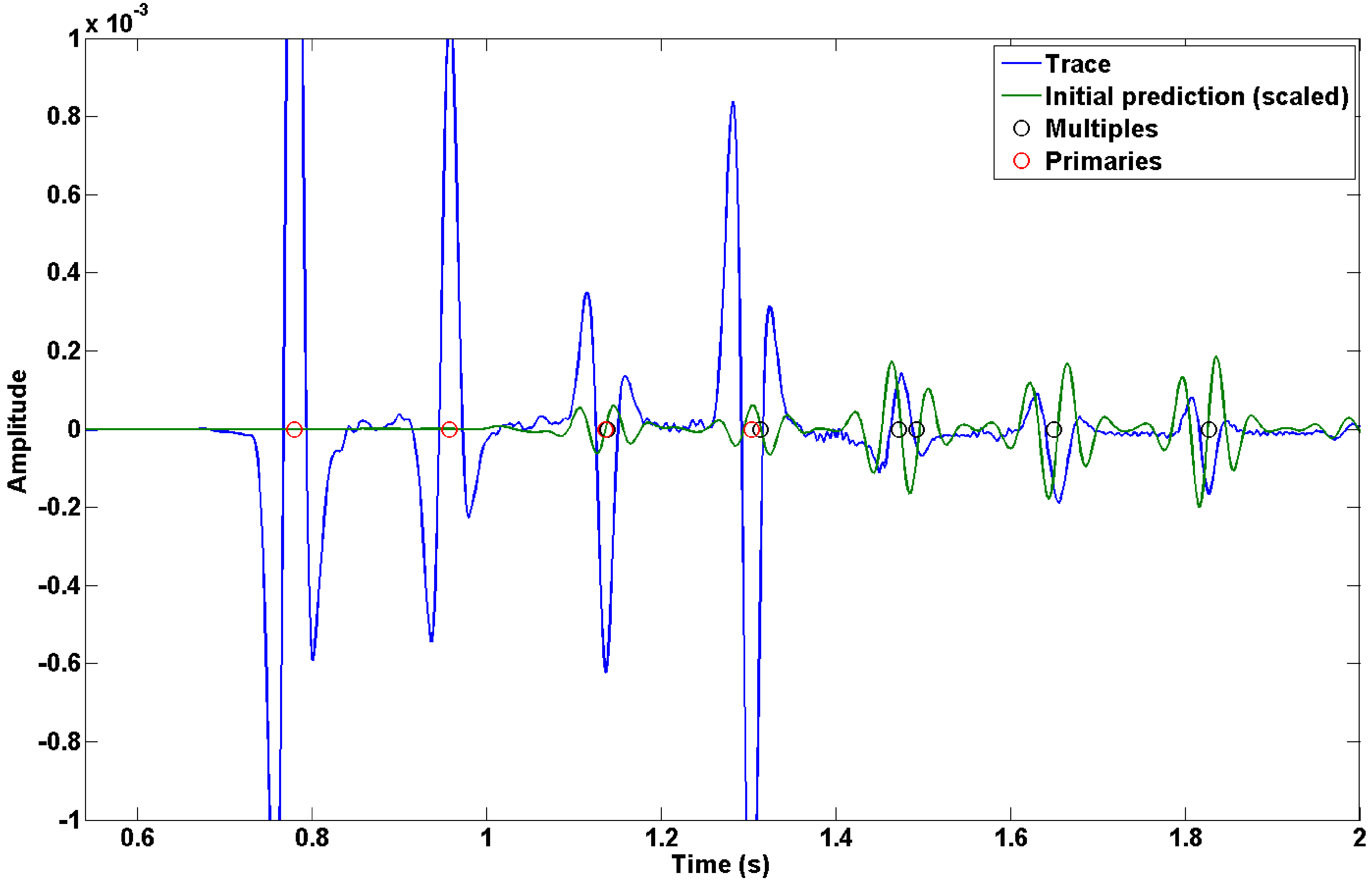

Trace
Adapted Prediction
Multiples
Primaries
Amplitude
Time (s)
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7

Least squares adaptive subtraction

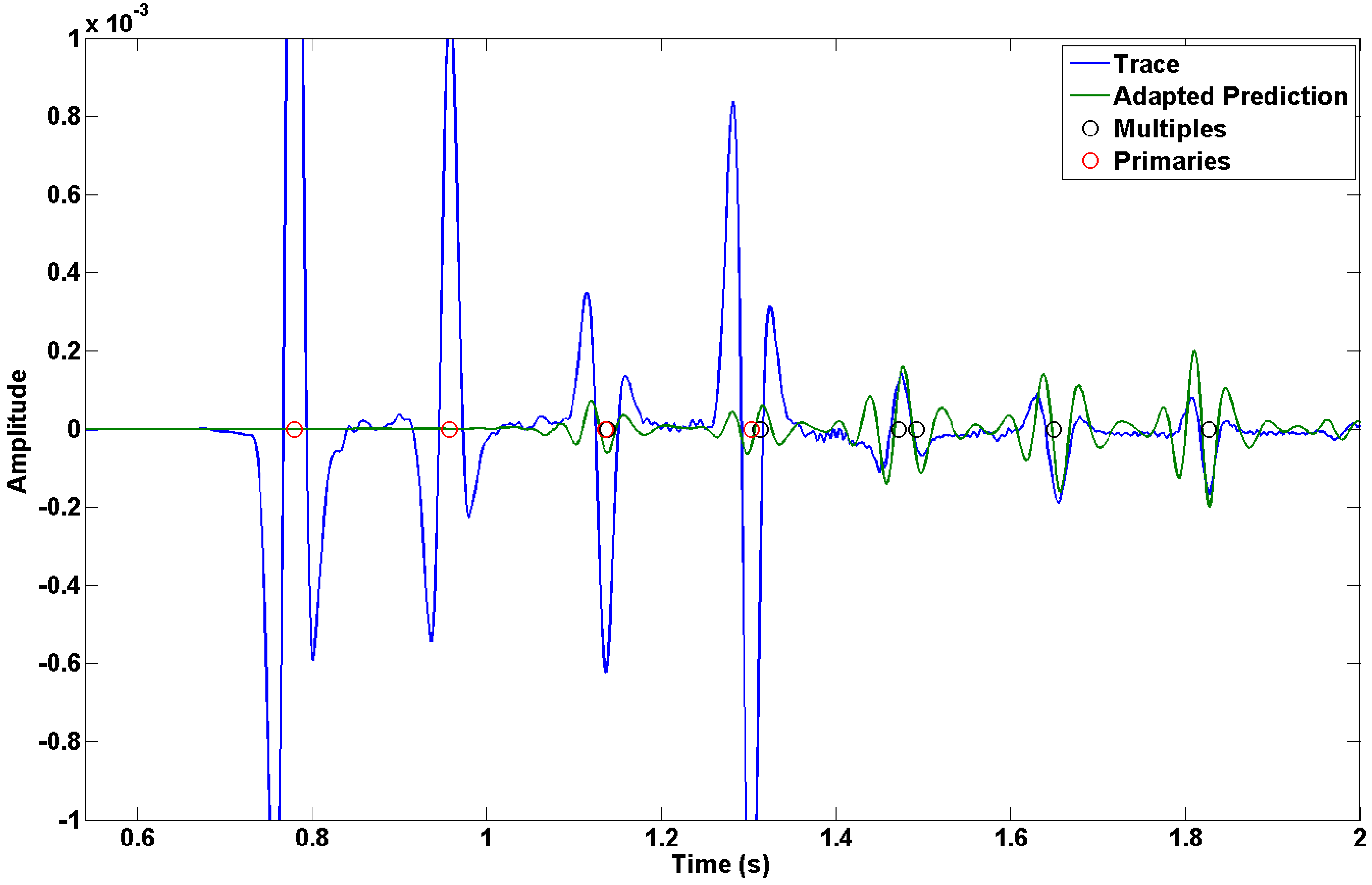
- Unfortunately, multiples and primaries often overlap to some extent.
- Given that primaries typically have greater energy than multiples, least squares subtraction will prioritize the removal of primaries in the case of primary-multiple overlap.
- This can lead to poor multiple removal, and worse, removal of signal from the primaries .

Modeling multiples

- Synthetic shot records were generated using the *afd_shotrec* function from the CREWES toolbox.
- For initial conditions, an analytic solution of the wavefield at some time t and later time t+dt were used.
- This allowed for a more accurate representation of the wavefield as compared to point source initial conditions



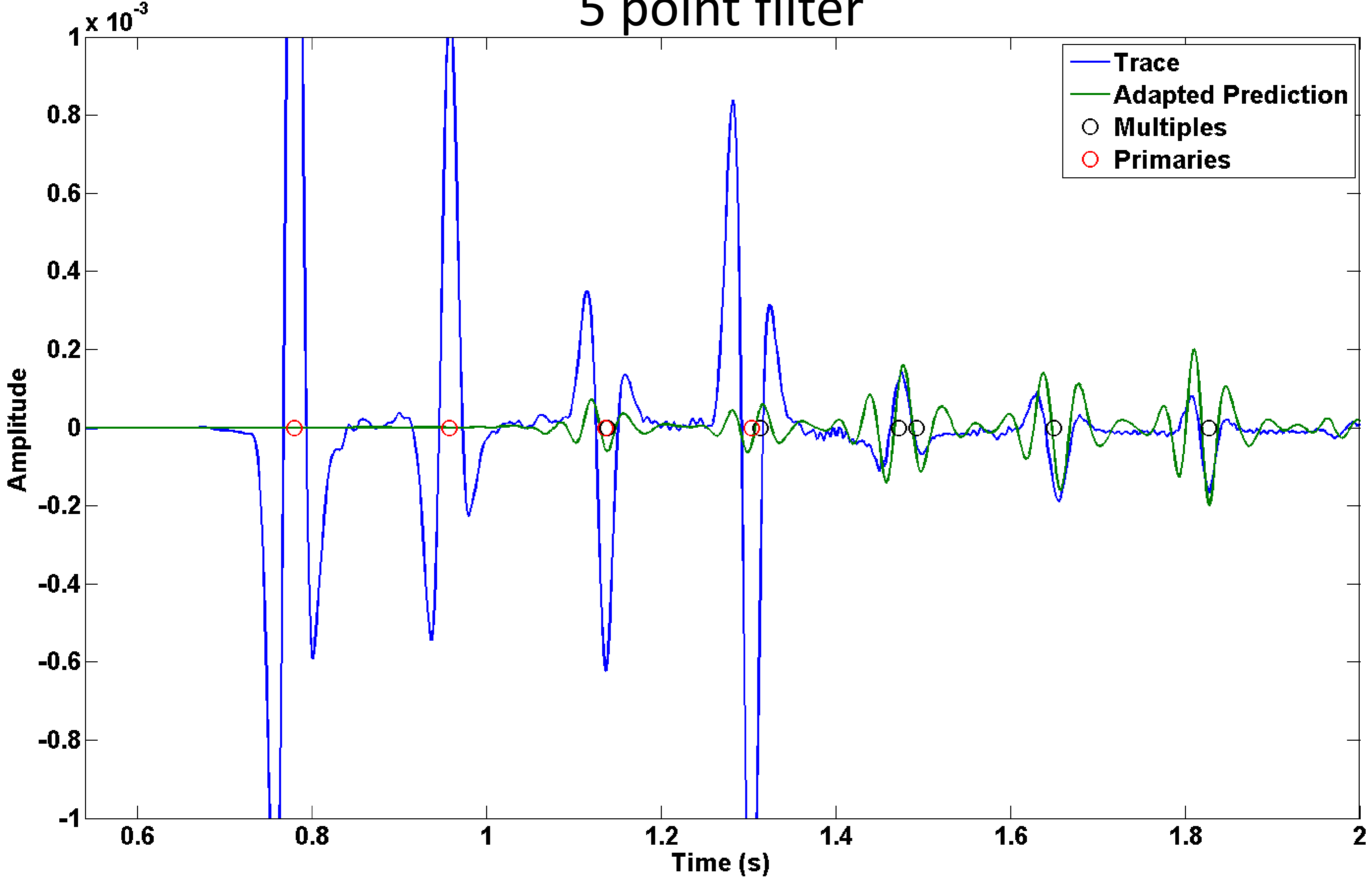
x 10^-3
Amplitude
Time (s)
Trace
Initial prediction (scaled)
Multiples
Primaries
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0.6
0.8
1
1.2
1.4
1.6
1.8
2


x 10^-3
Amplitude
Time (s)
Trace
Adapted Prediction
Multiples
Primaries
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0.6
0.8
1
1.2
1.4
1.6
1.8
2

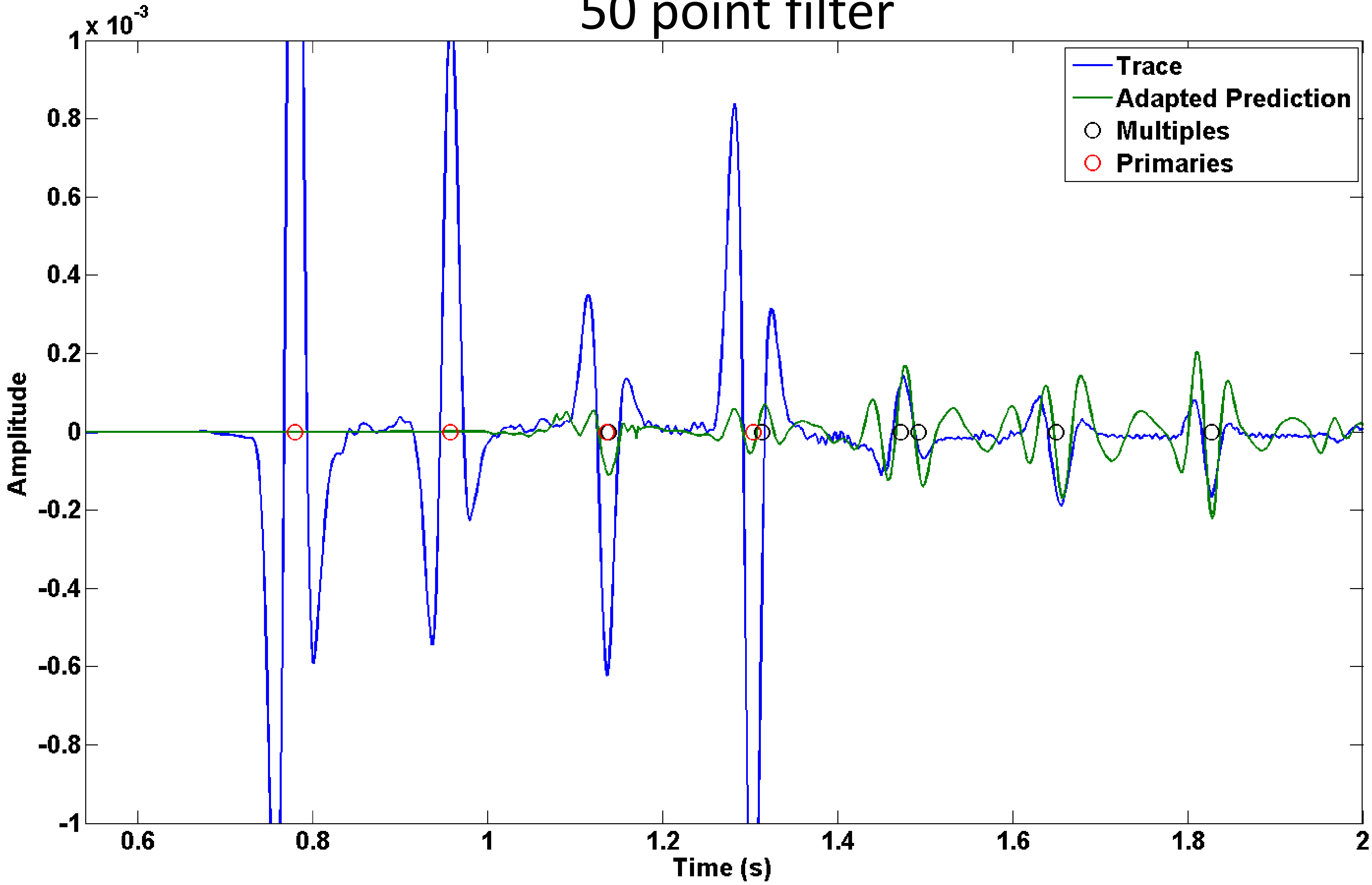
Least squares adaptive subtraction

- An important parameter in adaptive subtraction is the filter length.
- With a sufficiently long filter, the predicted multiples can be made to match any signal.
- A longer filter can be advantageous, as it increases the level of matching, and allows for more complete removal of multiples
- Longer filters also increase the chance that primary data will be removed

5 point filter



50 point filter



L1 Norm

- The major problem in the least squares adaptive subtraction method is the tendency for predicted multiples to be matched to primaries.
- This largely arises due to the large amplitudes of the primaries.
- If the same amount of signal is subtracted from a large signal and a smaller signal, the subtraction will lower the L2 norm more in the case where the large signal is reduced.
- This means that the energy minimizing filter often does its best to match primaries over a small time period, at cost to the matching to multiples over a longer period.

L1 Norm

- An alternative to the L2 norm for adaptive subtraction is the L1 norm (Guitton and Verschuur, 2004).
- When a constant amount of signal is reduced from a large signal or a small signal, the L1 norm reduces by the same amount.
- Consequently, high amplitude primaries are of dramatically less relevance in L1 norm minimization.
- The L1 norm is given by

$$\sum_{i=0}^{t} |r_i|$$

L1 Norm

- The L1 norm minimizing filter is found by solving the following least-squares equation

$$f = (M^T W M)^{-1} M^T W d$$

where W is a diagonal matrix whose elements W_{ii} are related to the residual at time *i* by $w_{ii} = r_i^{-1}$, where

$$r = Mf - d$$

- Unfortunately, this expression is singular where r_i is zero.

L1/L2 Norm

- Bube and Langan (1997) propose an L1/L2 hybrid norm, which in the limit of very small residuals behaves as an L2 norm, in the limit of very large residuals behaves as an L1 norm, and transitions smoothly between them.
- As the L2 norm is well behaved as the residual approaches zero, the hybrid norm is non-singular everywhere.
- The expression for this hybrid norm is again

$$f = (M^T W M)^{-1} M^T W d$$

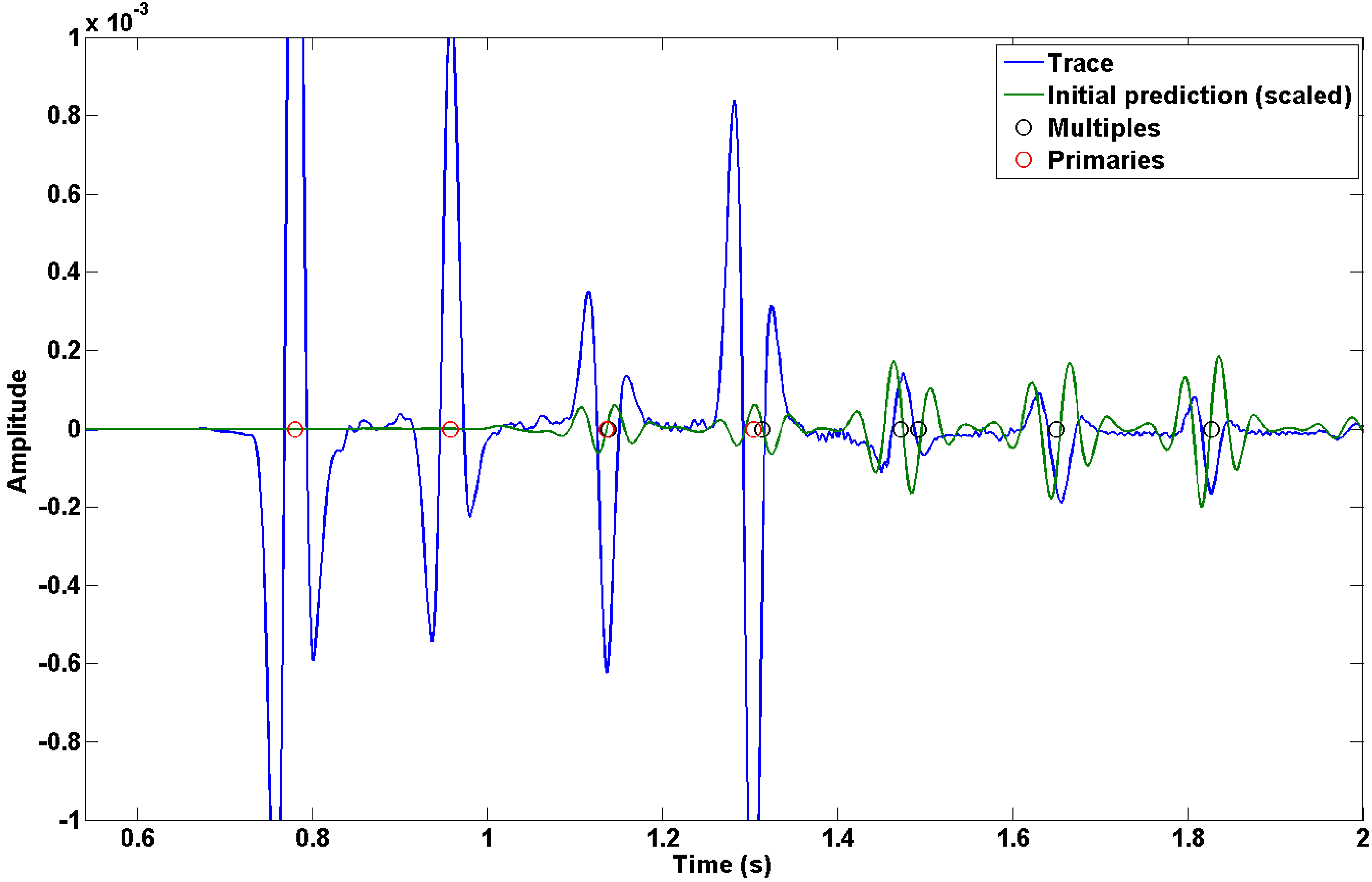
but with

$$w_{ii} = \left(\frac{1}{1+(\frac{r}{\sigma})^2} \right)^{\frac{1}{2}}$$

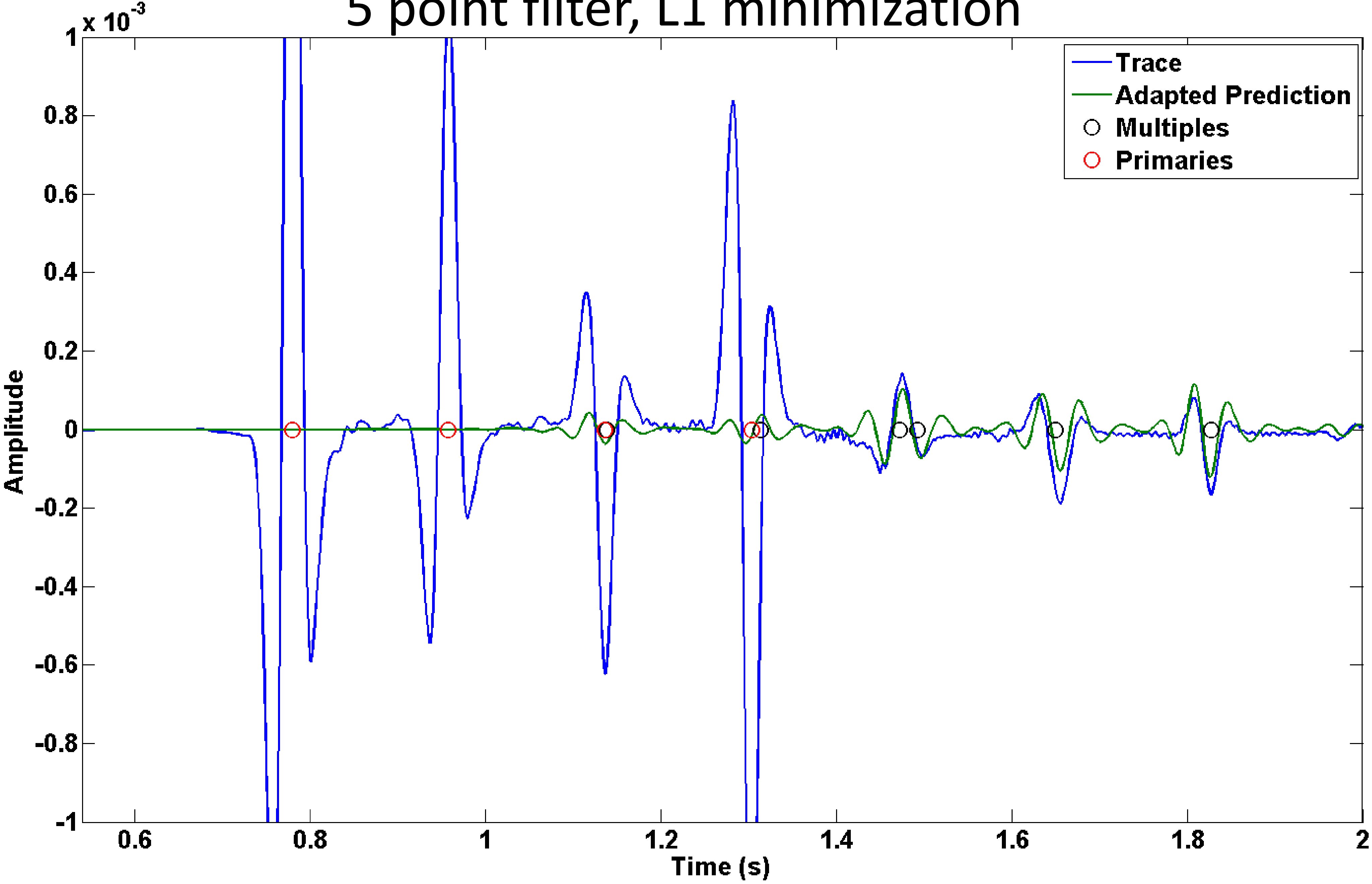
where r are the residuals and sigma is a chosen factor

L1/L2 Norm

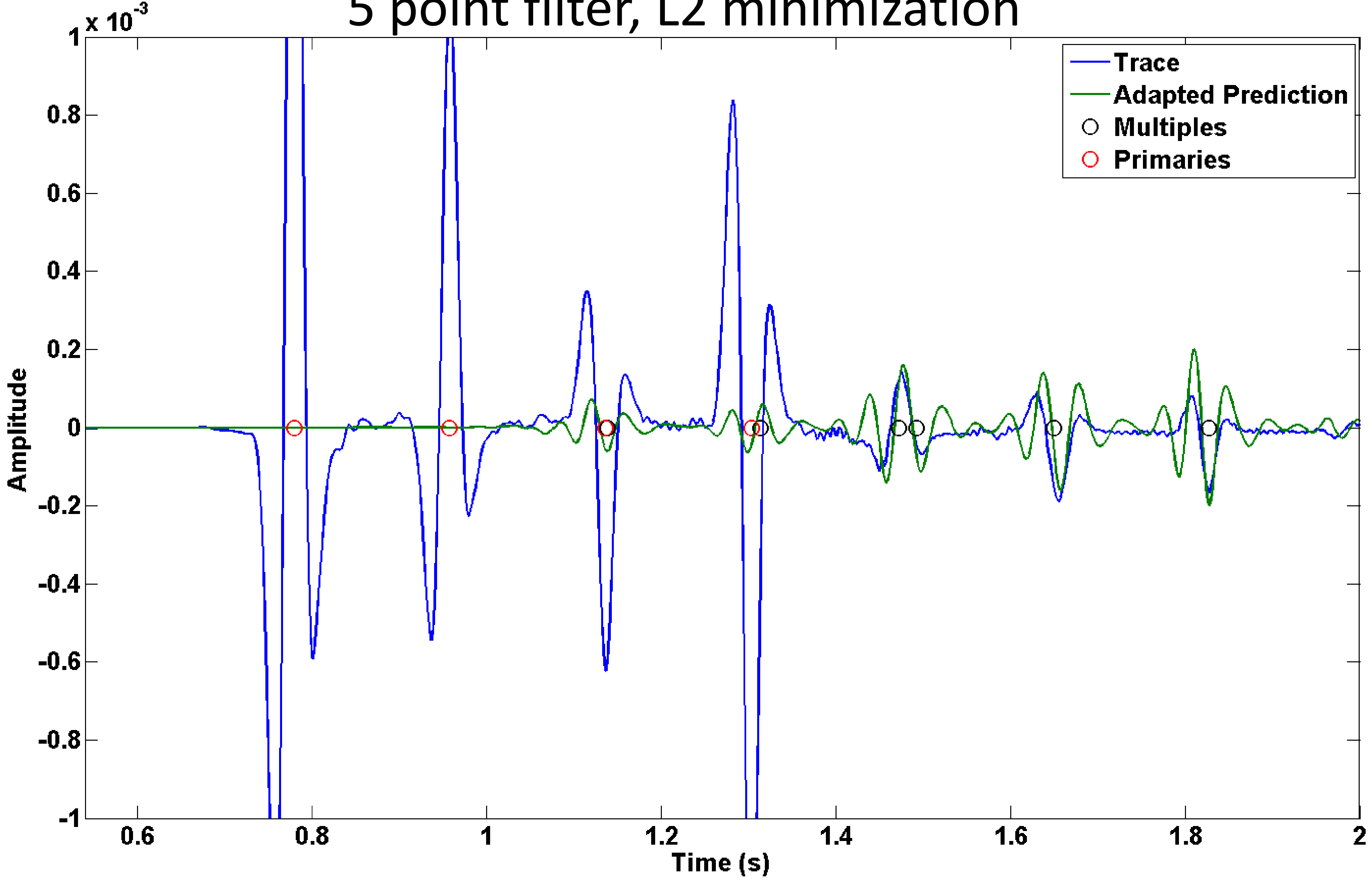
- In effect this equation minimizes J, where $J = \sqrt{1 + (\frac{r}{\sigma})^2} - 1$
- Small sigma will closely emulate the L1 norm, while large sigma will approximate L2.
- The factor sigma must be decided on by the user.
- The expression for the L1/L2 norm is nonlinear.
- It can be solved iteratively.


x 10^-3
Amplitude
Time (s)
Trace
Initial prediction (scaled)
Multiples
Primaries
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0.6
0.8
1
1.2
1.4
1.6
1.8
2

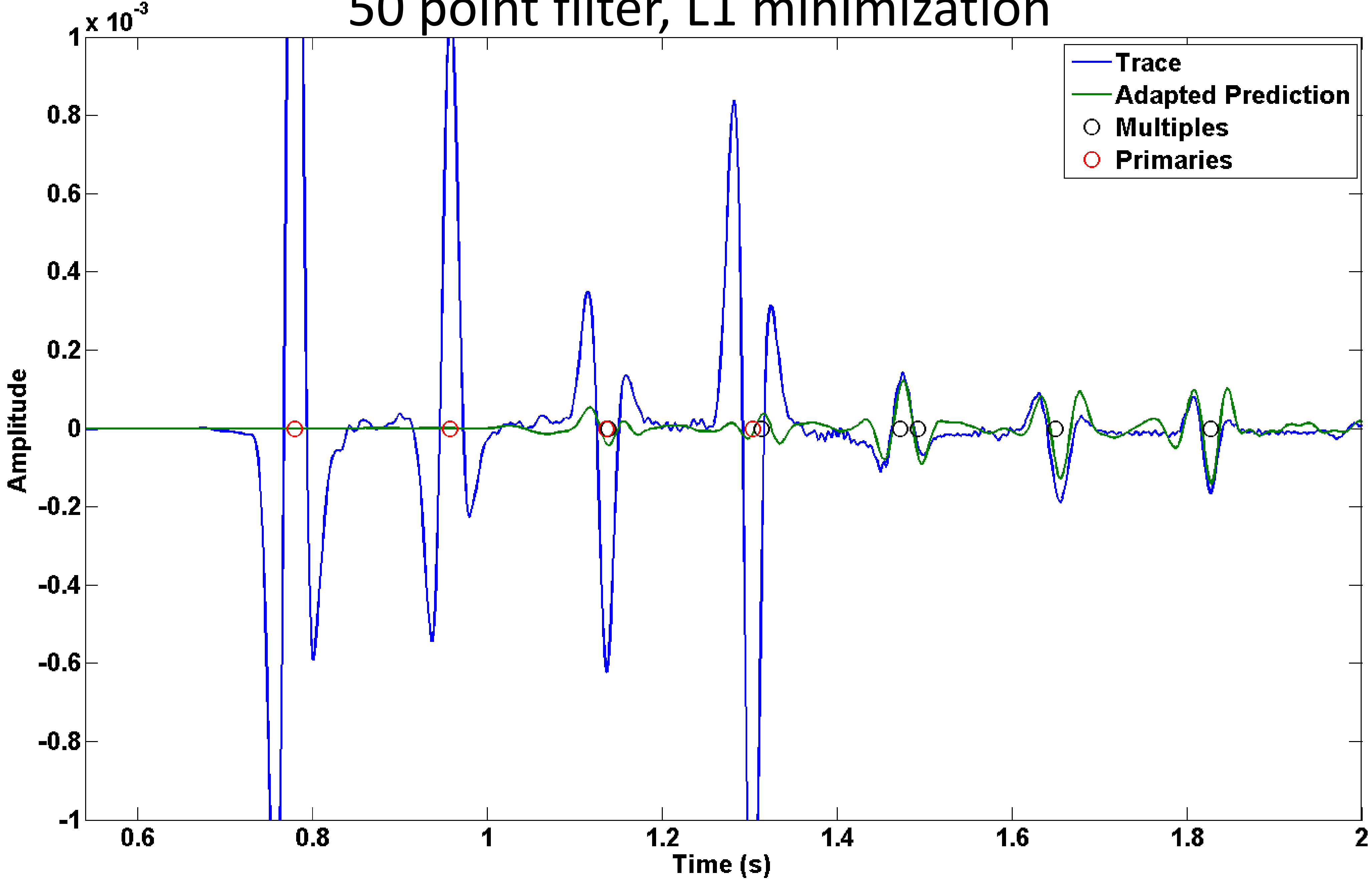
5 point filter, L1 minimization



5 point filter, L2 minimization



50 point filter, L1 minimization

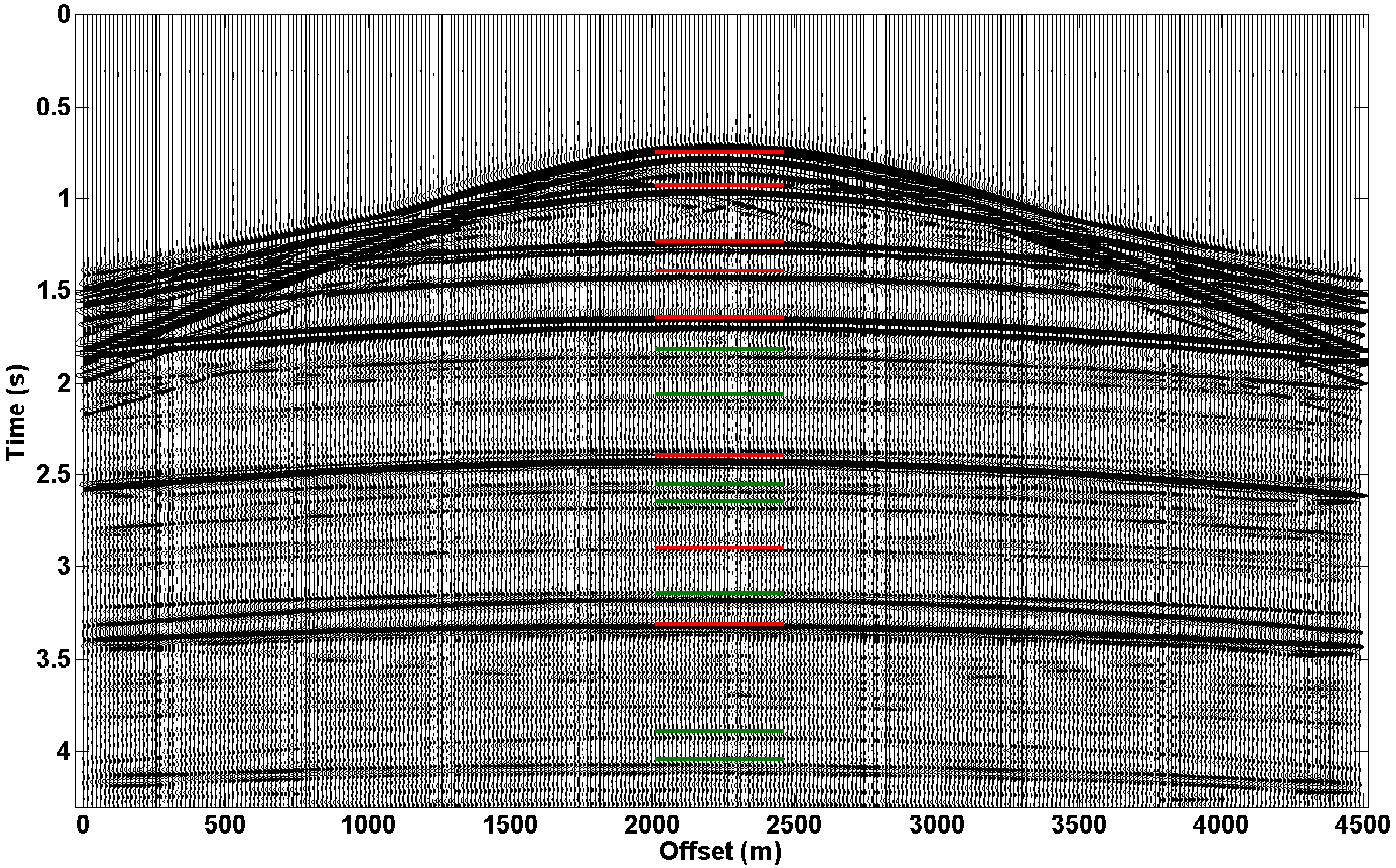


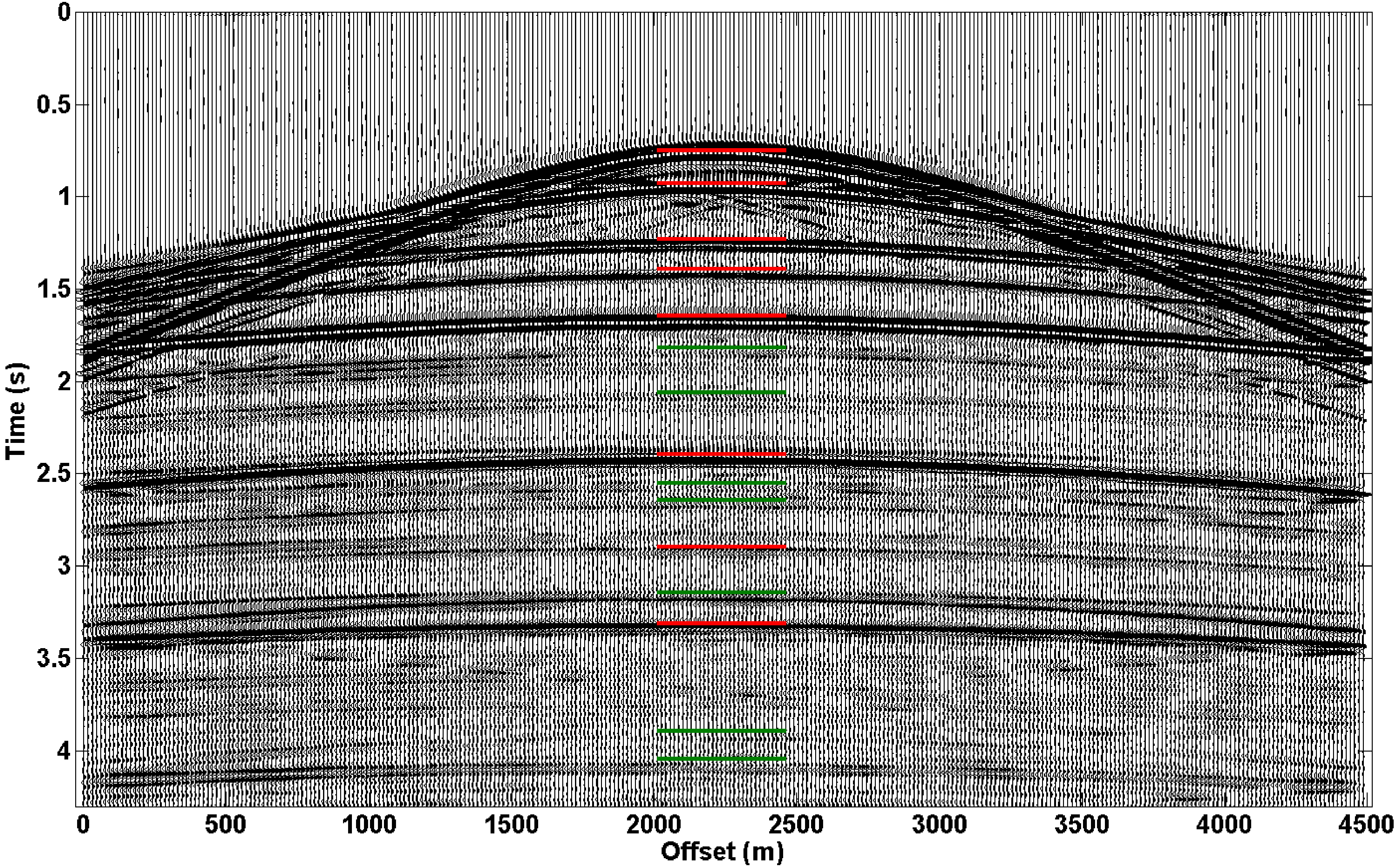
Nonstationarity

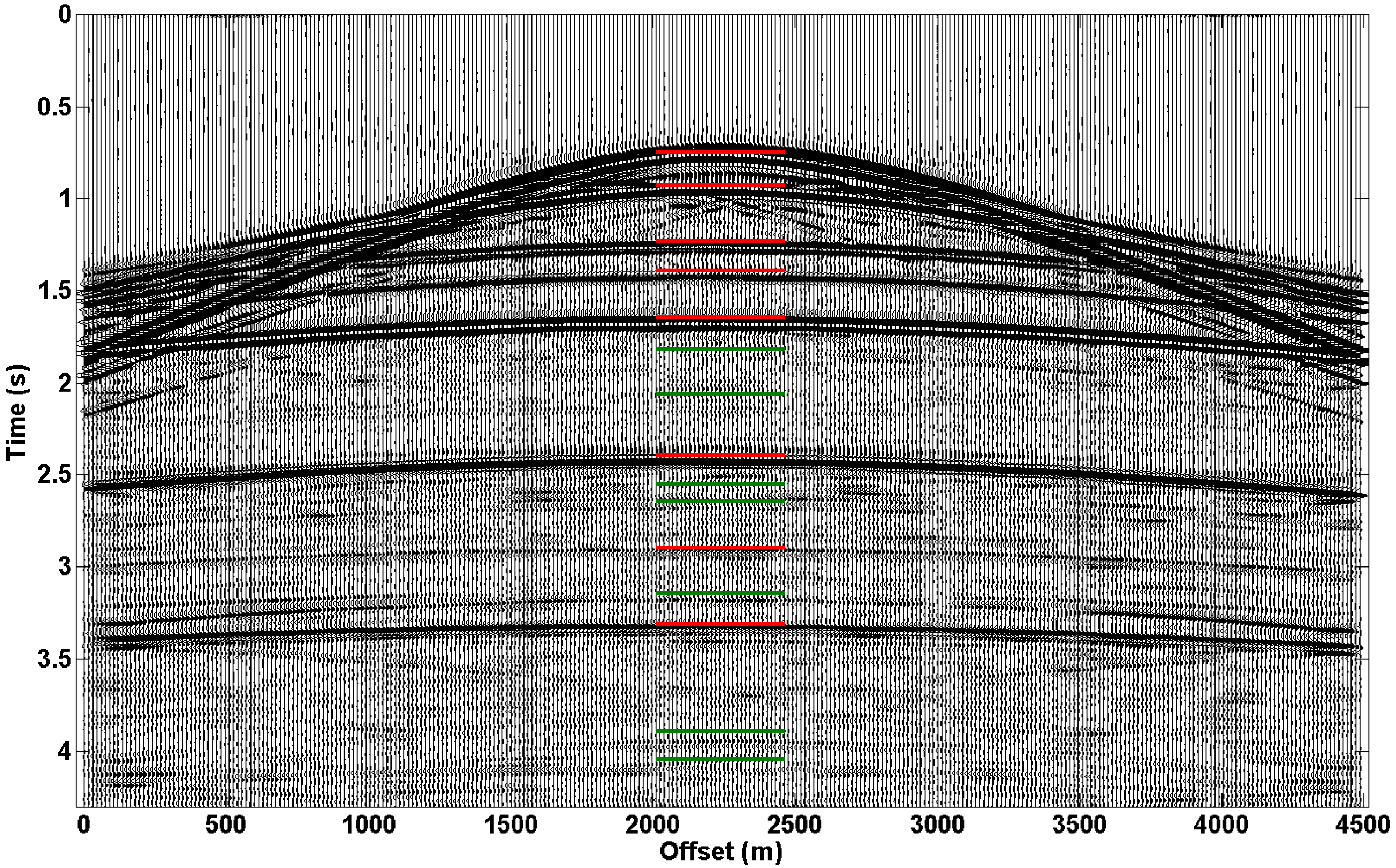
- Seismic data is itself often nonstationary.
- Additionally, approximations may be made in multiple prediction whose validity varies in space or time.
- These factors can lead to a multiple prediction which is not related to the true multiples by a single, stationary filter.
- In this case, a time and/or space variant filter is necessary.

Nonstationarity

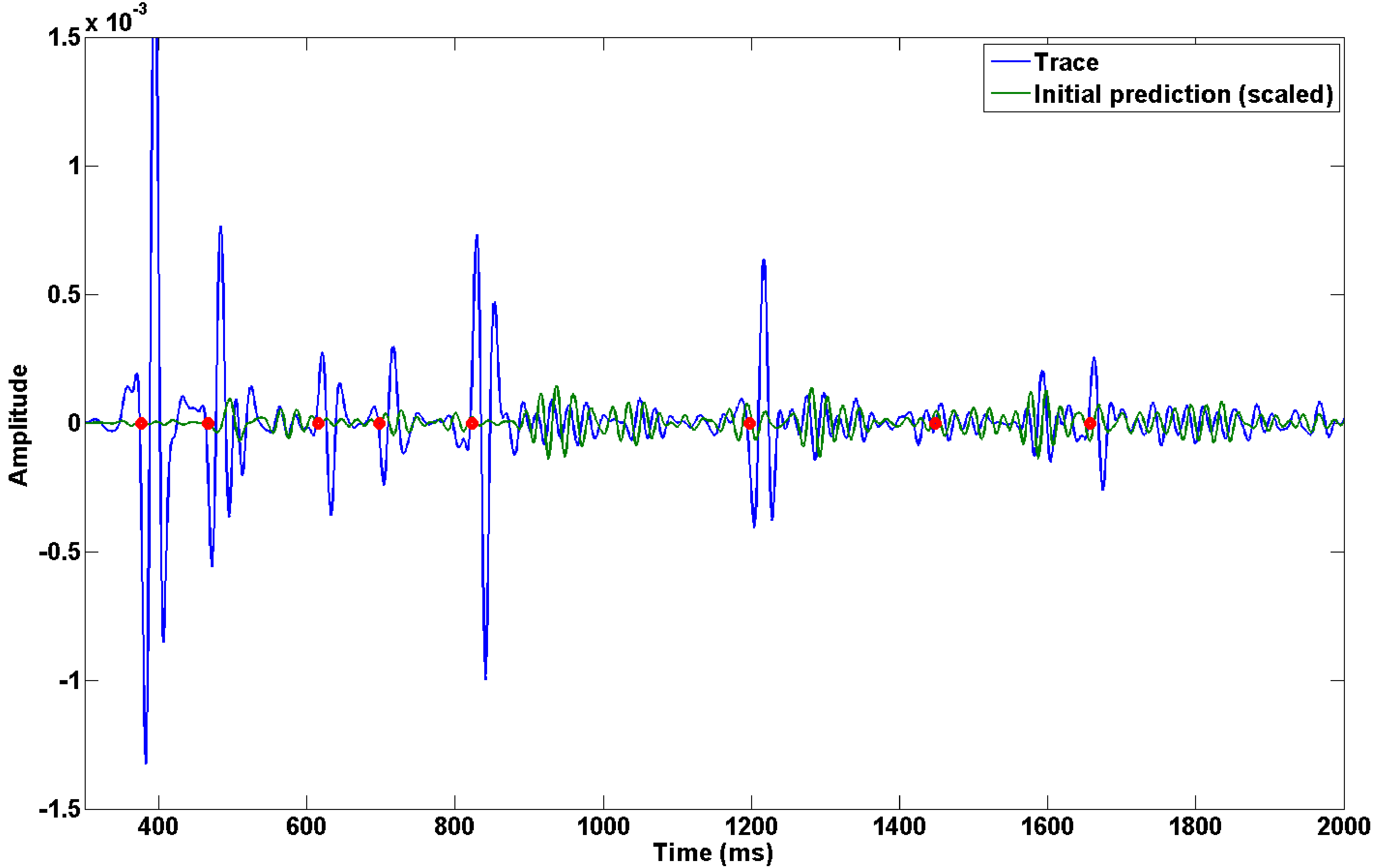
- A nonstationary filter can be created by windowing the data about each point in succession, and calculating the filter which works best in each window.
- This creates a different filter for each point.
- The window size controls how quickly the filter is allowed to vary.
- A Gaussian-shaped window was found to give good results.


Time (s)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
500
1000
1500
2000
2500
3000
3500
4000
4500
Offset (m)

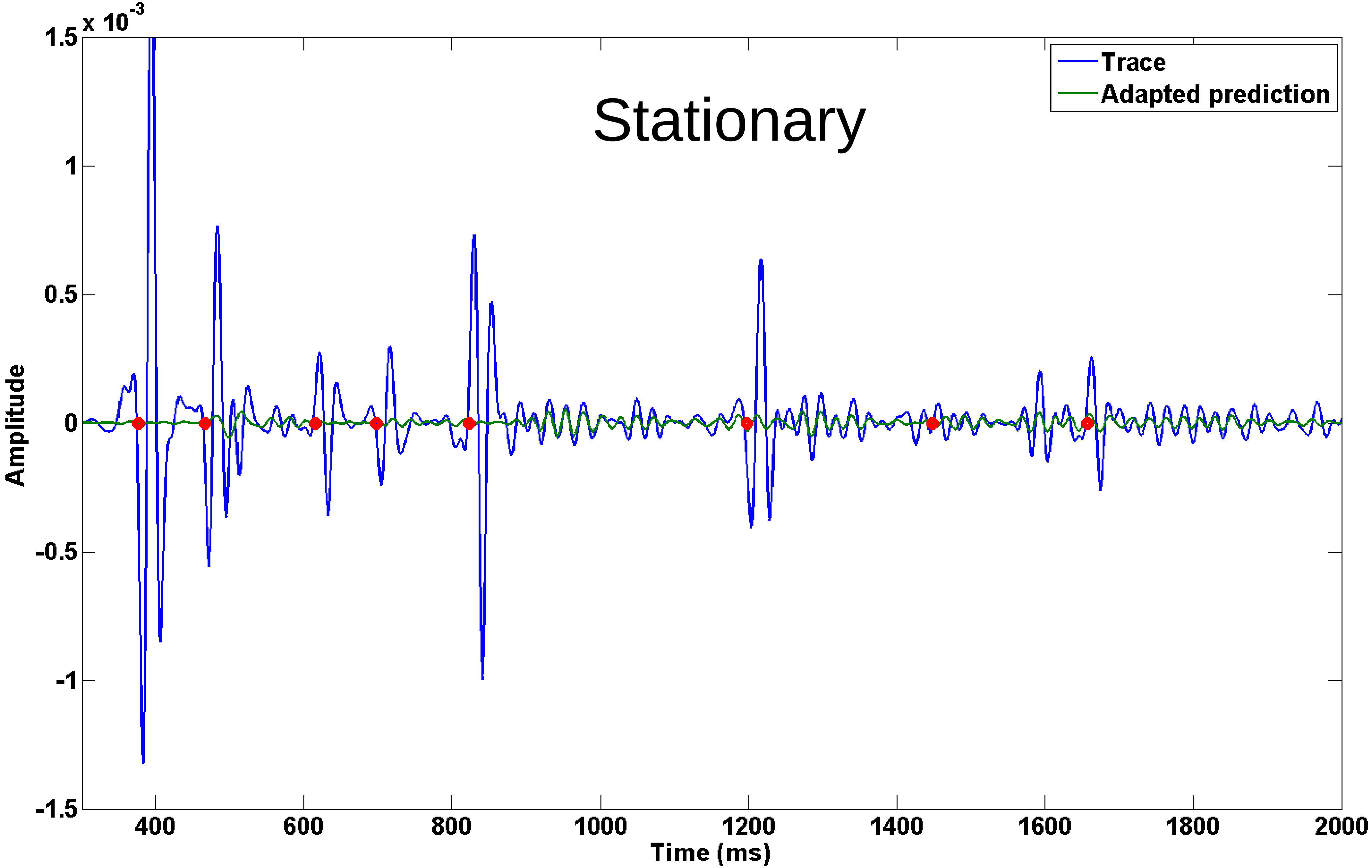

Time (s)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
500
1000
1500
2000
2500
3000
3500
4000
4500
Offset (m)



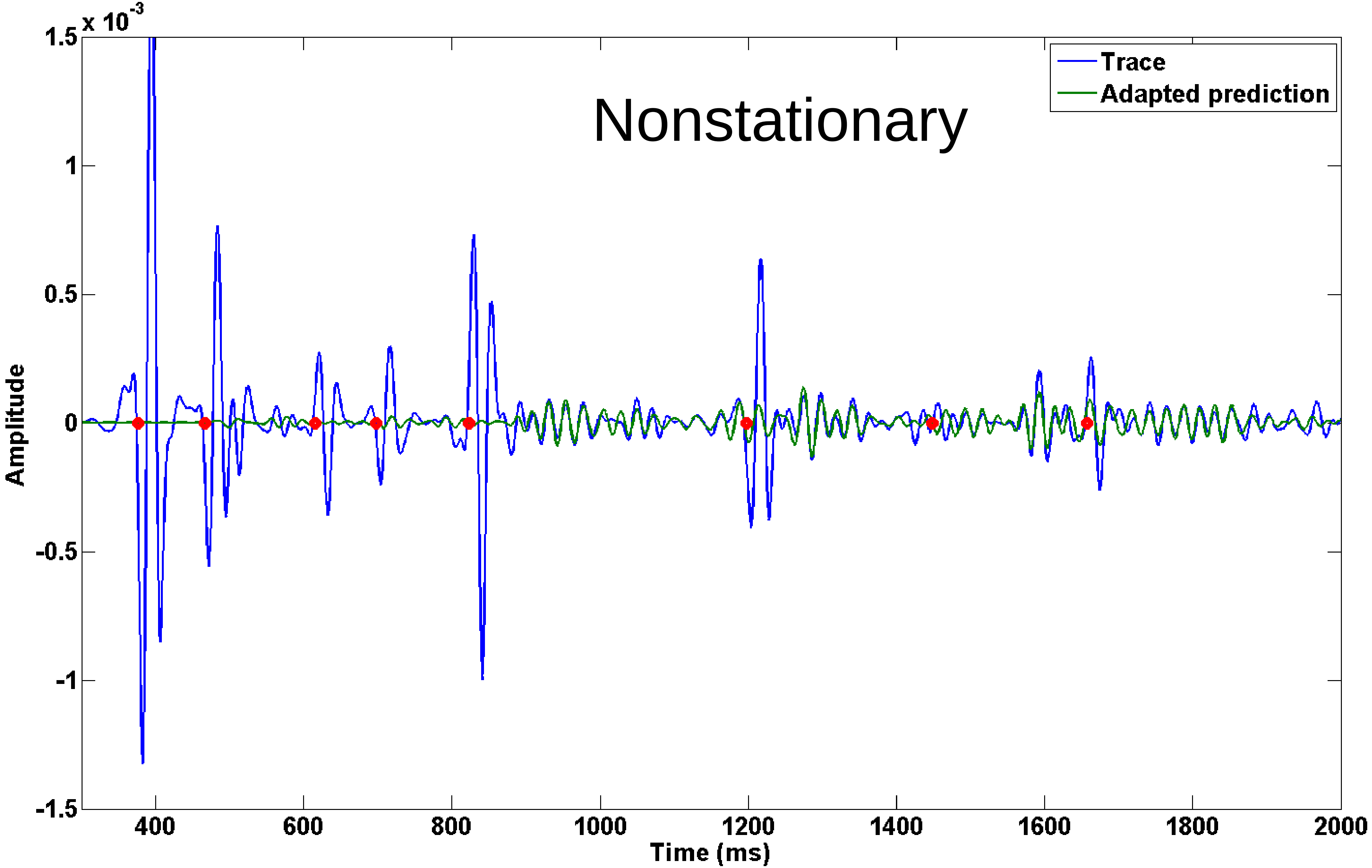
Time (s)
0
0.5
1
1.5
2
2.5
3
3.5
4
0
500
1000
1500
2000
2500
3000
3500
4000
4500
Offset (m)



x 10^-3
1.5
1
0.5
0
-0.5
-1
-1.5
Amplitude
400
600
800
1000
1200
1400
1600
1800
2000
Time (ms)
Trace
Initial prediction (scaled)



Stationary
Trace
Adapted prediction
x 10^-3
1.5
1
0.5
0
-0.5
-1
-1.5
Amplitude
400
600
800
1000
1200
1400
1600
1800
2000
Time (ms)



x 10^-3
1.5
1
0.5
0
-0.5
-1
-1.5
Amplitude
400
600
800
1000
1200
1400
1600
1800
2000
Time (ms)
Trace
Adapted prediction
Nonstationary

Conclusions

- Inverse scattering multiple predictions are inexact in practice, and need to be modified before they can be subtracted from the data.
- This modification can be done by convolving the data with a filter.
- An L1 minimizing, nonstationary filter provides a means of achieving a reliable and robust adaptive subtraction

Future Work

- Integrate adaptive subtraction with prediction implementations in tau-pg-ps, kg-t and xg-t domains
- Apply multiple prediction and adaptive subtraction to field data

References

- Bube, K., and Langan, R., 1997, Hybrid L1/L2 minimization with applications to tomography, Geophysics, **62**, 1183-1195
- Guitton, J., and Verschuur, D., 2004, Adaptive subtraction of multiples using the L1-norm, Geophysical Prospecting, **52**, 27-38
- Weglein, A., et al, 2003, Inverse scattering series and seismic exploration, Inverse Problems, **19**, R27–R83.
- Pan, P., 2015, 1.5D Internal Multiple Prediction: an Application on Synthetic Data, Physical Modelling Data and Land Data Synthetics.
- Sun, J., 2014,1.5D Internal multiple prediction in the plane wave domain, CREWES Annual Report, **26**.

Acknowledgments

My supervisor, Kris Innnanen

Crewes staff and students

Andrew Mills, for his help with the presentation

Jian and Penny, for their help with and codes on inverse scattering multiple prediction

CREWES

NSERC-CRD (CRDPJ 379744-08)