Evelyne Meier
FEL09, Liverpool, August 2009

Electron beam stabilization test results using a neural network hybrid controller at the Australian Synchrotron and Linac Coherent Light Source

Overview

1. Motivation
2. Background on Neural Networks
3. Control Scheme
4. Australian Synchrotron Linac studies
5. LCLS studies
6. Application to the FERMI Linac

1. Motivation

- **Project Background**

  - Design of a control system for longitudinal parameter stabilization of the new FERMI@Elettra FEL Linac.

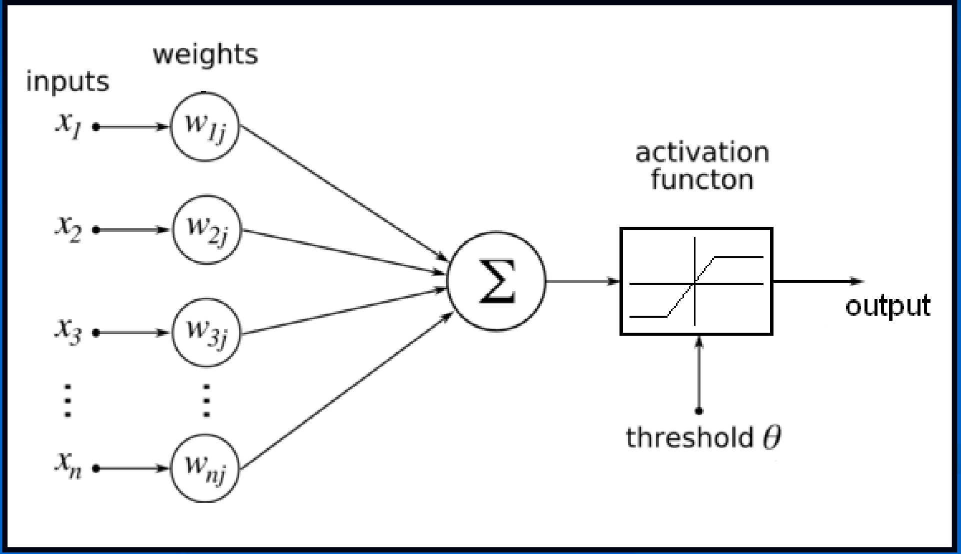
- **Goals**

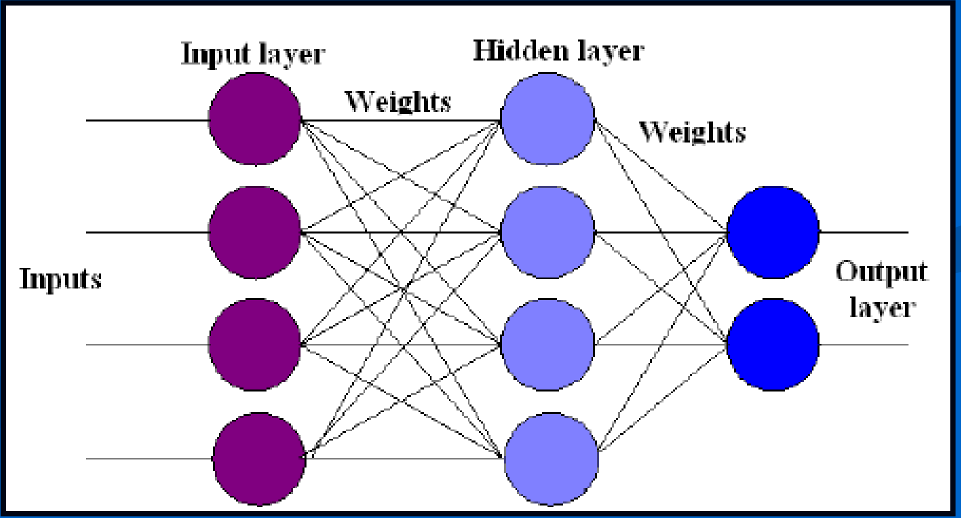
  - Perform energy and bunch length control at different stages of the Linac.
  - Eliminate high frequency jitter in a feed forward way.
  - Ensure the system robustness by coupling the feed forward control to a feedback.
  - Control system with large bandwidth, self adaptive to changes in jitter conditions.

2. Background on Neural Networks

Neural Network description

- Inputs from different neurons (for example klystron phase and voltage).
- The network weights contain the knowledge of the system.
- The activation function tells the artificial neuron when and with what strength to fire.
- The training algorithm adjusts the weights to minimize the error between the network output and the desired output.





2. Background on Neural Networks

- $\vec{x_j}$ = input vector for unit j (x_{ji} = ith input to the jth unit)
- $\vec{w_j}$ = weight vector for unit j (w_{ji} = weight on x_{ji})
- $z_j = \vec{w_j} \cdot \vec{x_j}$, the weighted sum of inputs for unit j
- y_o = network output
- $\hat{y}_o$ = targeted output

The weights are updated: $\boxed{w_{ji} \leftarrow w_{ji} + \Delta w_{ji}}$

Using the error:

$$E = \frac{1}{2} \sum_{o \in Outputs} (y_o - \hat{y}_o)^2$$

Learning rate

The weight change can be written as:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = \eta \delta_j x_{ji}$$

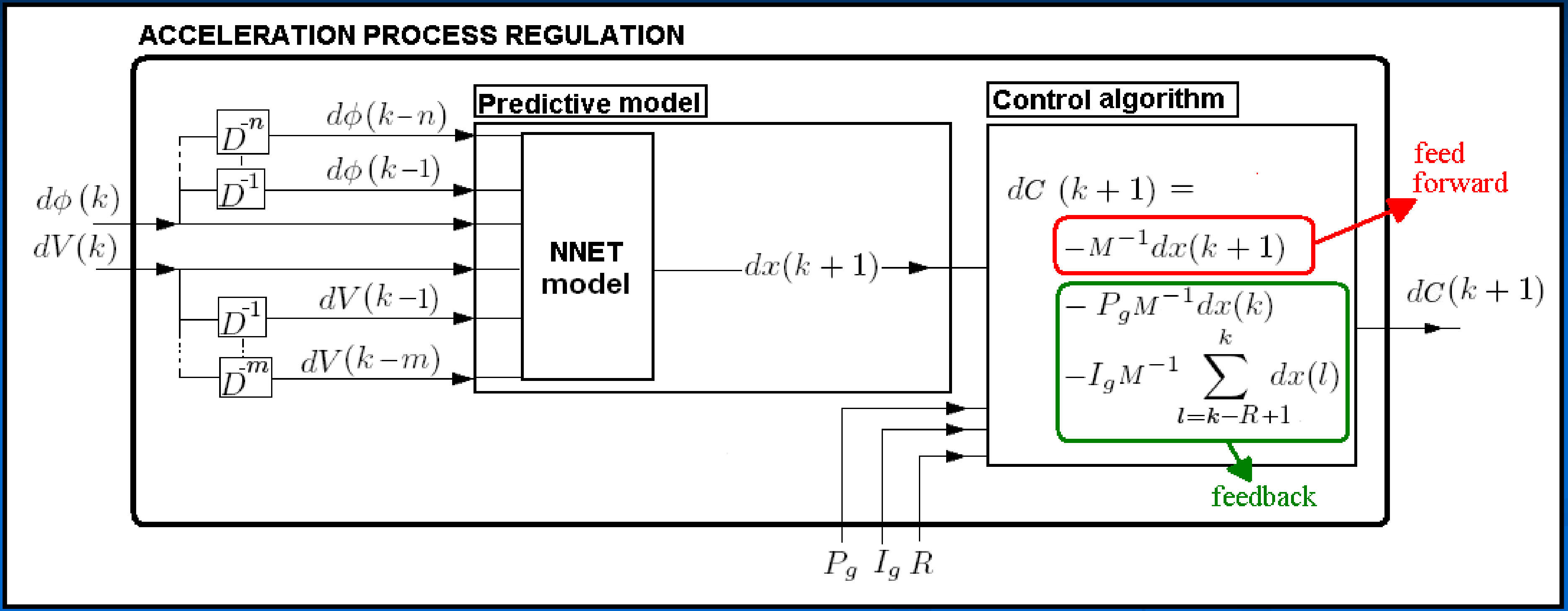
With, at the output layer:

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial z_j} \frac{\partial z_j}{\partial w_{ij}} = \frac{\partial E}{\partial z_j} x_{ij}$$

And for the other layers:

$$\frac{\partial E}{\partial w_{ij}} = \sum_{k \notin Output} \frac{\partial E}{\partial z_k} \frac{\partial z_k}{\partial f_j(z_k)} \frac{\partial f_j(z_k)}{\partial z_j} \frac{\partial z_j}{\partial w_{ij}} = \sum_{k \notin Output} \frac{\partial E}{\partial z_k} \frac{\partial z_k}{\partial f_j(z_k)} \frac{\partial f_j(z_k)}{\partial z_j} x_{ij}$$

3. Control Scheme

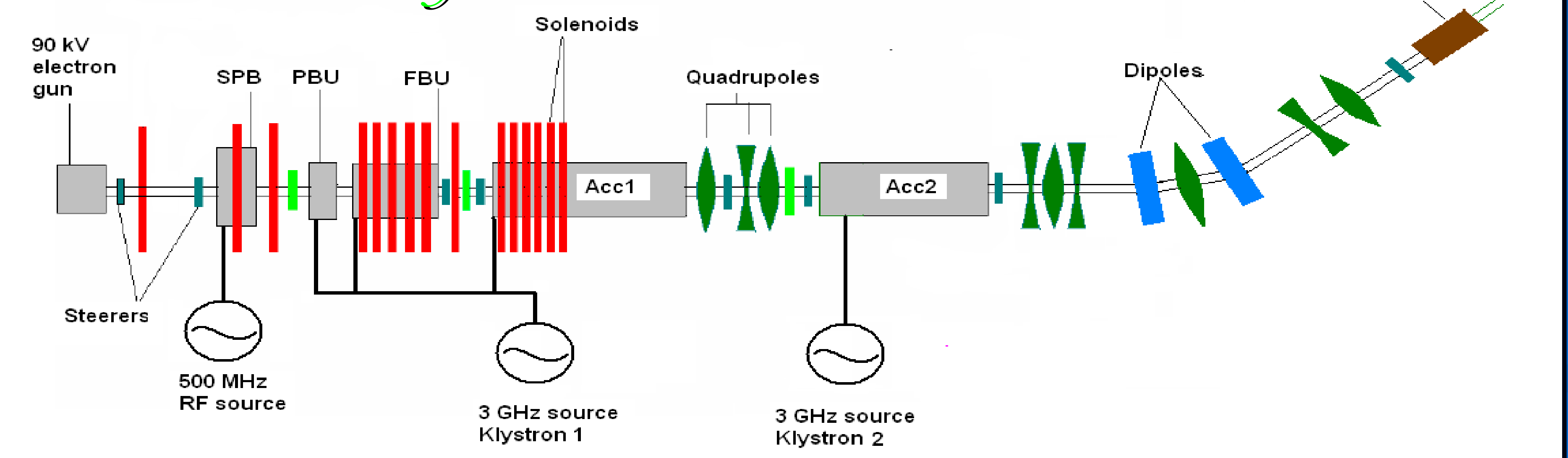


- Predictive model (left block): The Neural Network (NNET) receives lagged values of the perturbed klystron phase and voltage. It gives a prediction of the next pulse position deviation *dx(k+1)* used to compute the feed forward correction.

- Control algorithm (right block): The algorithm is composed of a feed forward (first term) augmented by PI control terms (second and third terms) to compute the correction to apply :*dC(k+1)* .

4. Australian Synchrotron Linac Studies

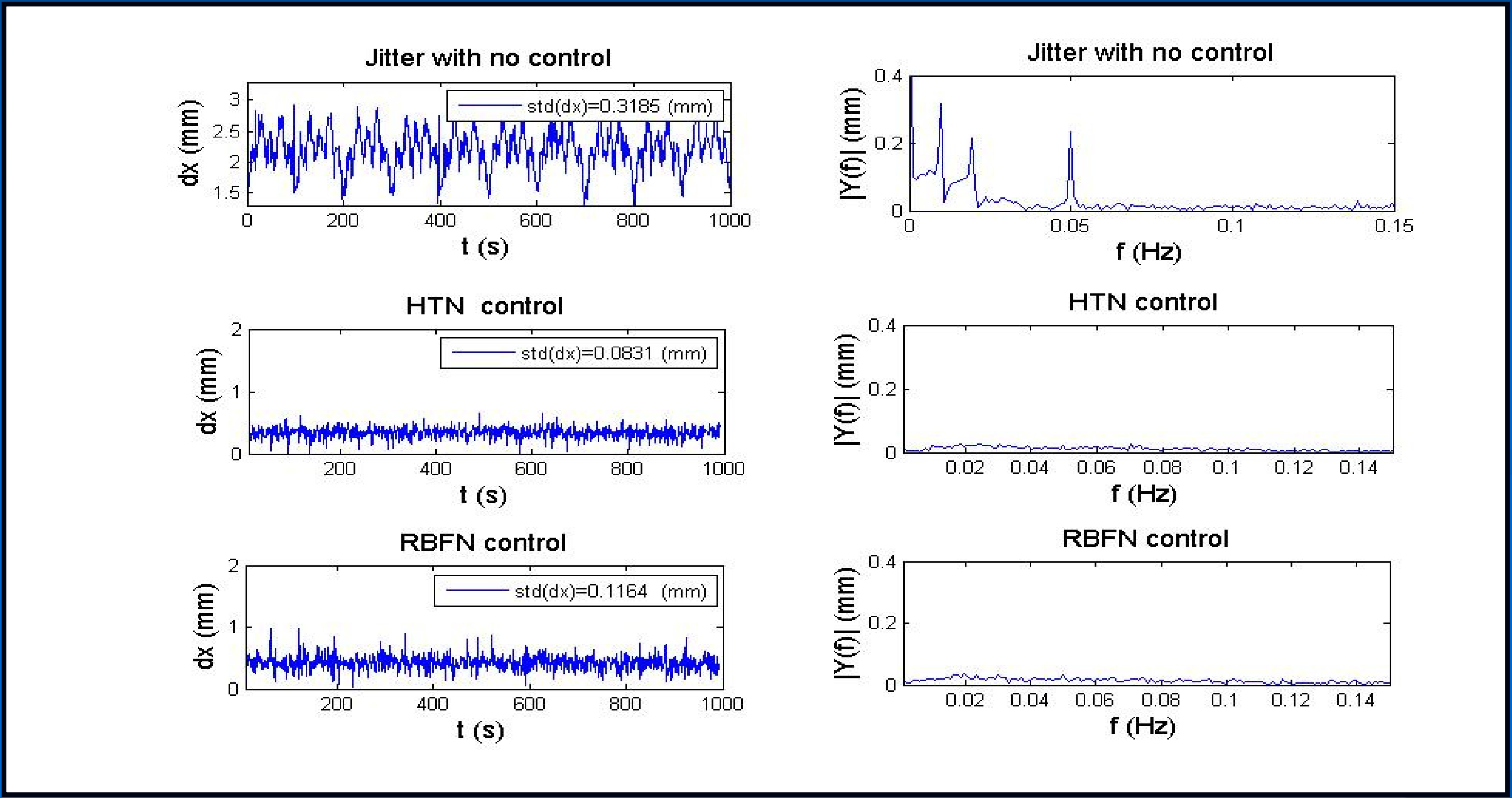
Linac Assembly



- ACC1 and ACC2 provide 100 MeV beam @ 1-10Hz.
- BPM resolution ~50 µm, white noise level ~ 0.11 mm rms.
- Jitter are induced in klystron 1 phase and voltage and corrected using klystron 2 voltage.
- Slow actuator response limits the experiment to max. ~0.075 Hz.

4. Australian Synchrotron Linac Studies

(1) Real time energy control

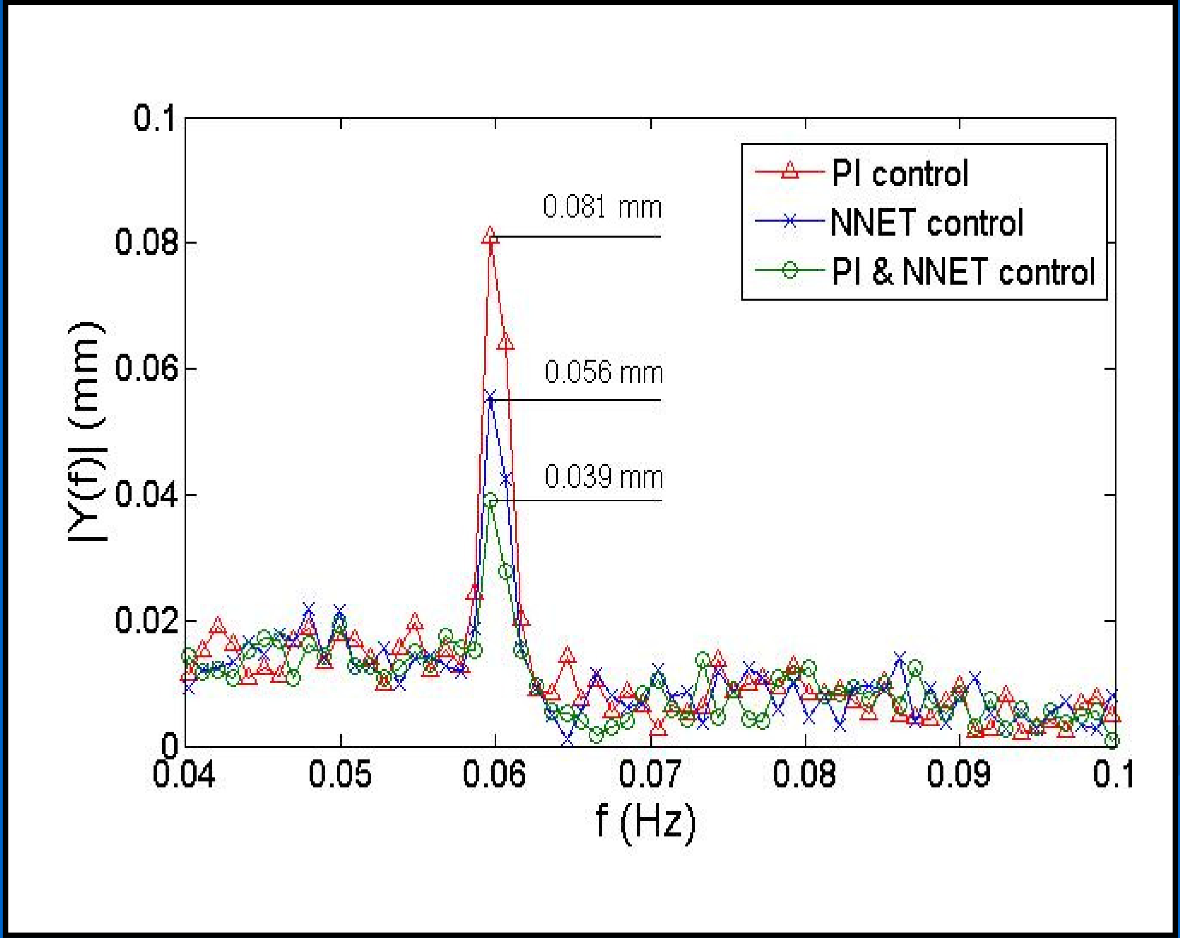


- Both the hyperbolic tangent network (HTN) and radial basis function network (RBFN) received 7 lagged values of V_1 and 5 lagged values of φ_1. The HTN and RBFN had 7 and 76 hidden neurons, respectively.

4. Australian Synchrotron Linac Studies

(2) Real time energy feed forward-feedback combined control:

- The network is trained for a 0.1 kV and 0.04 Hz jitter.
- It is then brought online to correct a 0.1 kV and 0.06 Hz jitter.
- FFT for frequency change: The combined system brings further attenuation.
- Even when mis-tuned the NNET performs better than the PI controller.



5. LCLS studies

- *Aims:*
  - Cancel higher frequencies (0.1-2.5 Hz) – limit due to the correcting actuators’ response.
  - Act at higher repetition rate (5 Hz with Matlab).
  - Couple energy and bunch length control.
  - Develop a more adaptive system.

- *Experiments:*

  (1) Couple energy and bunch length control.

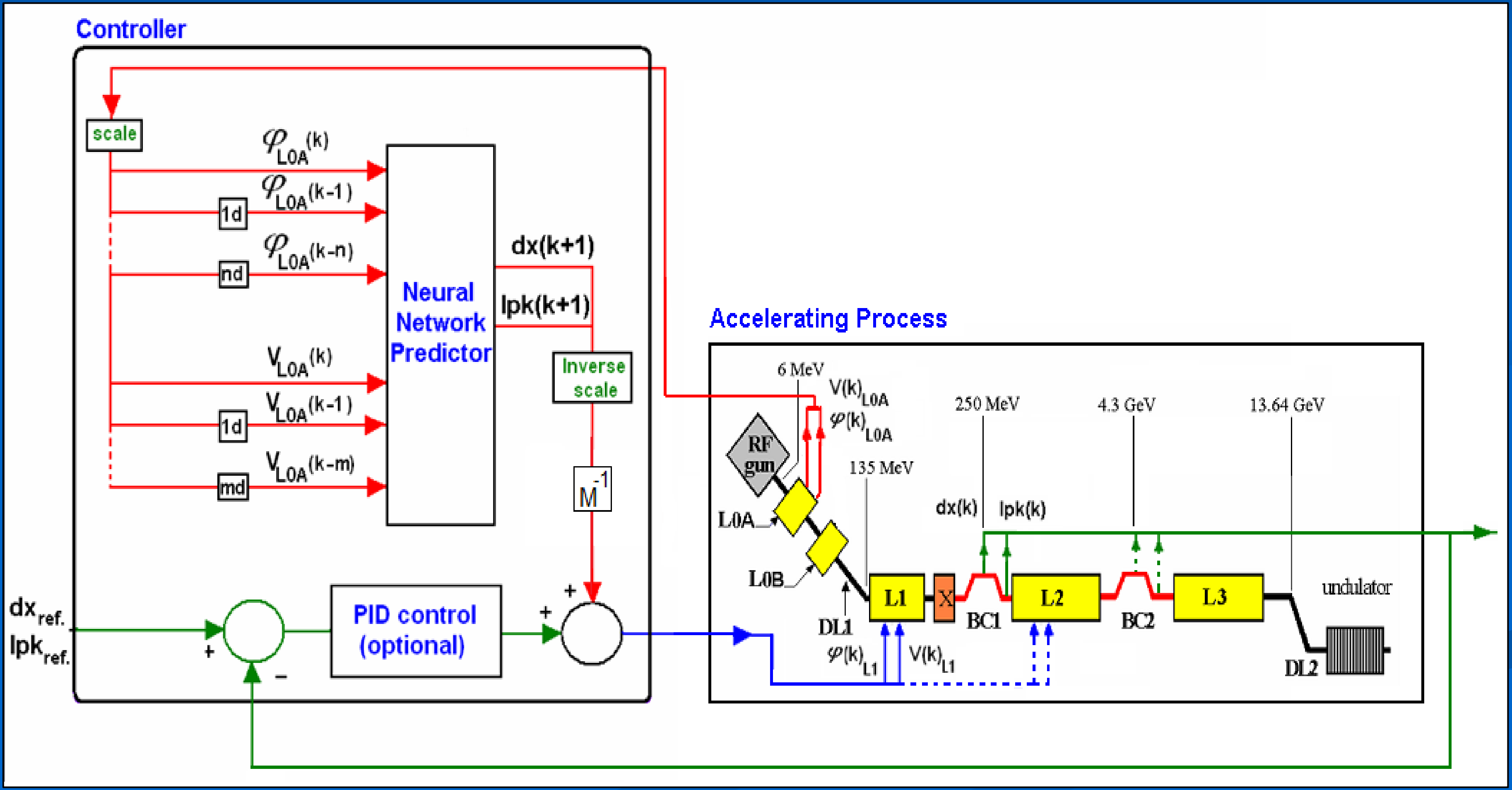
  (2) Comparison of NNET and PI algorithm performances.

  (3) Online adjustments of the response based on the data FFT to optimize the control.

  - Selected Network topology: 8-10 hidden neurons 8 lagged values of V_{L0A} and 8 lagged values of φ_{L0A}.

5. LCLS studies

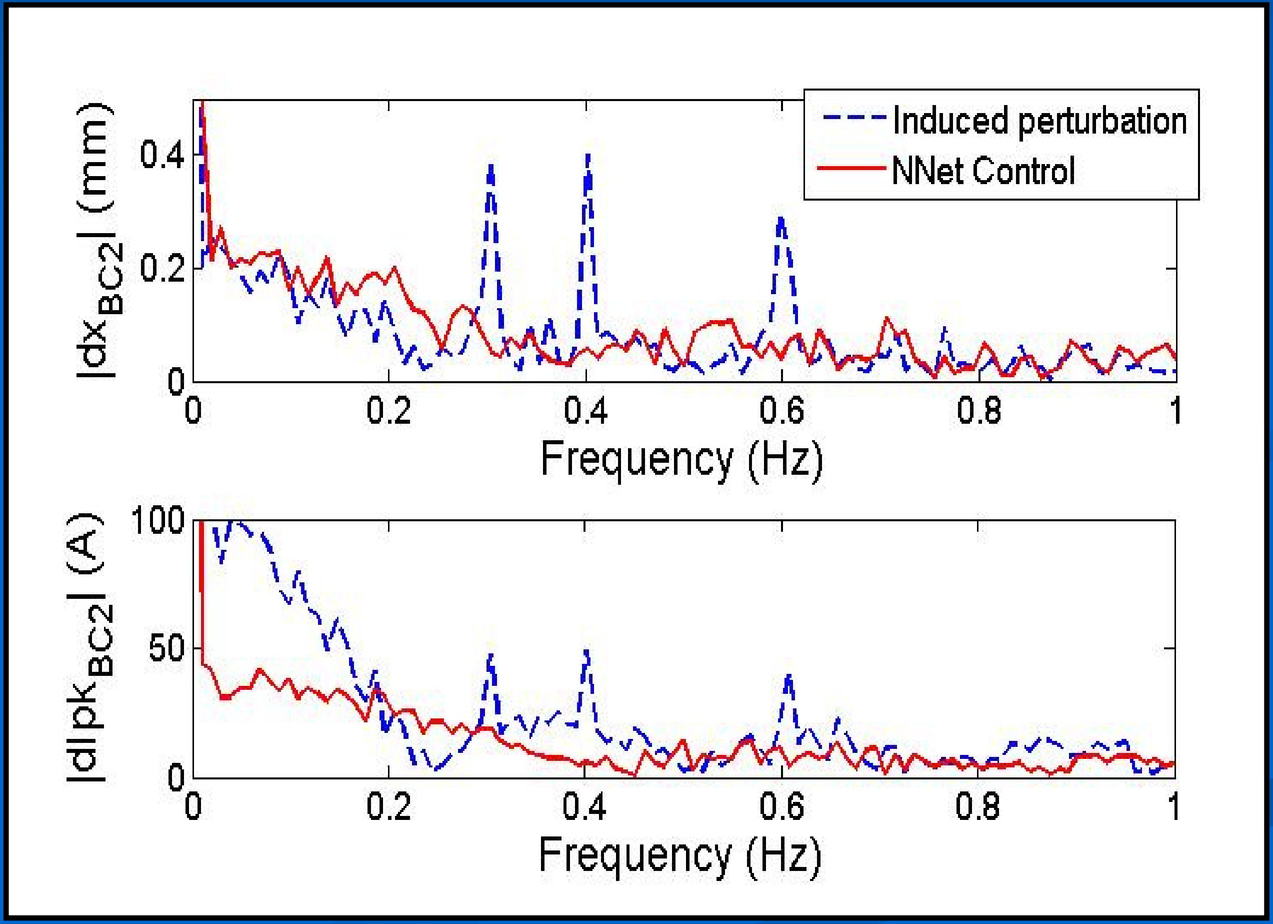
- *Machine assembly and control scheme:*



5. LCLS studies: Results

(1) Coupled control results at BC2

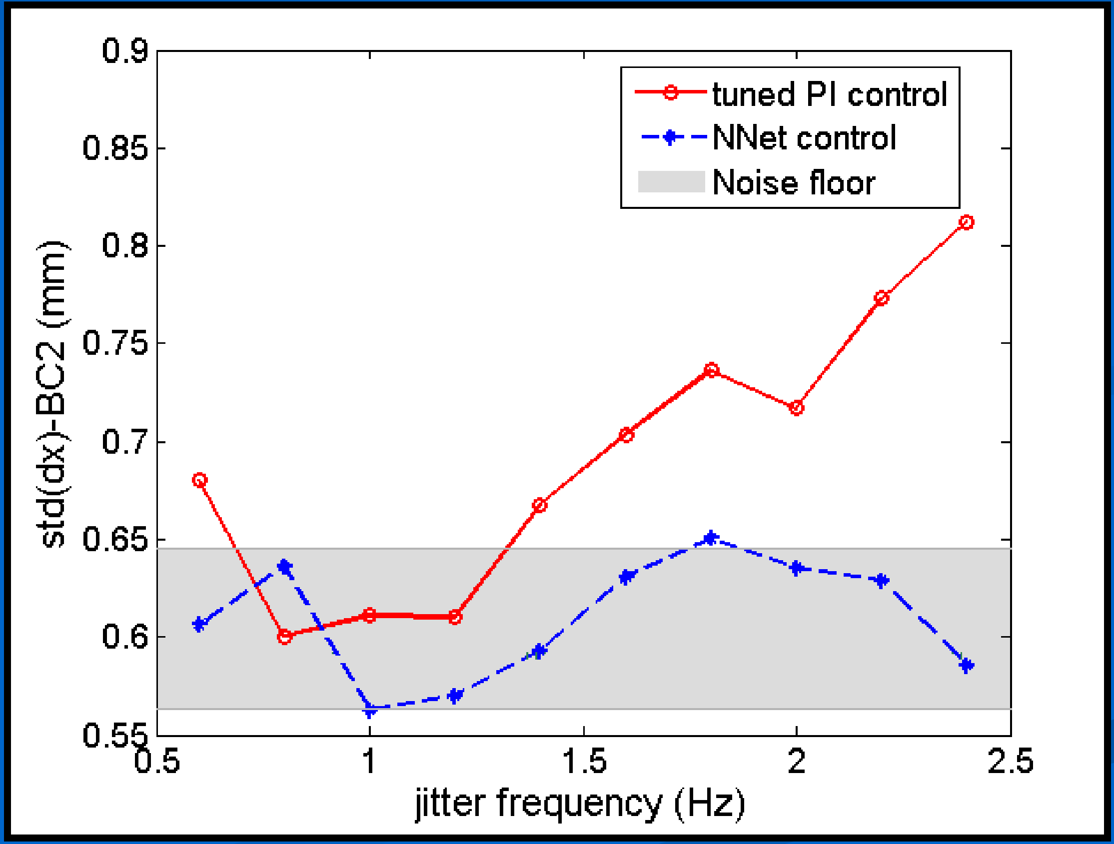
- Deviation recorded at BC2, with perturbations and corrections applied to two of the L2 klystrons.
- Successful cancellation of all frequencies was observed.
- Modulation in the frequency domain induced by the slow communication with the actuators through the Matlab channel access.



5. LCLS studies: Results

(2) PI and NNET performances

- Deviation recorded at BC2, with deviation and correction applied to klystrons of the L2 section.
- Original imposed deviation amplitude ~1 mm rms.
- NNET and PI re-tuned to optimize the correction for each frequency.
- Better performance of the NNET above 1.5 Hz.



5. LCLS studies

(3) Developing a more adaptive system

- The correction is computed according to:

$$dC(k+1) = -M^{-1}dO(k+1) - P_g M^{-1}dO(k) - I_g M^{-1} \sum_{l=k-R+1}^{k} dO(l)$$

- Accurate feed forward control of the machine depends on the accuracy of:
  - The network response: dO(k+1)
  - The response matrix: M

- The NNET prediction dO(k+1) can be imprecise if jitter conditions change. In this case re-training will resolve the problem.

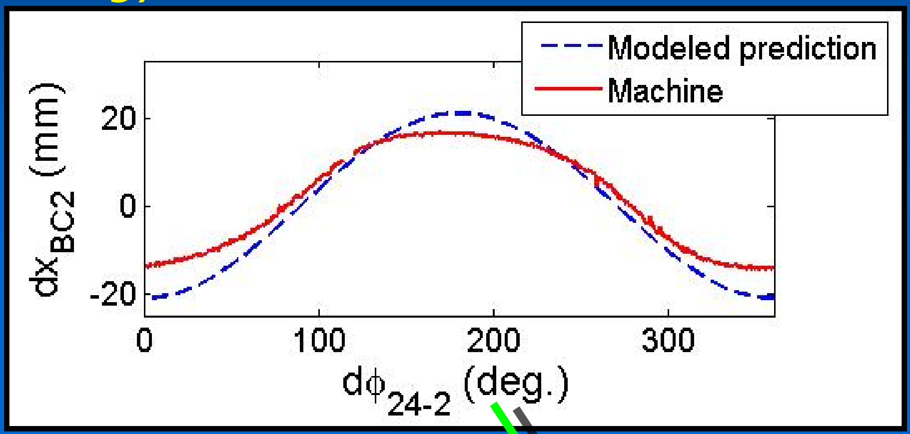
- Because of non-linearities of the actuator response and change in the klystron settings M is not static!

→ *M must be adjusted dynamically !!*

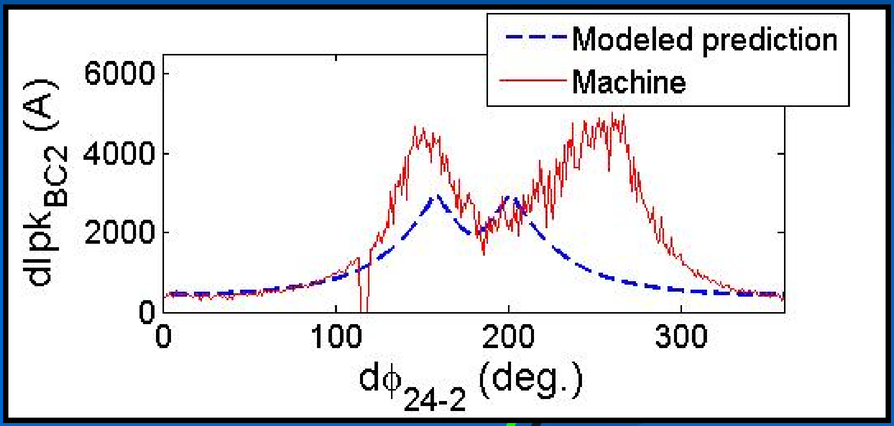
5. LCLS studies: Results

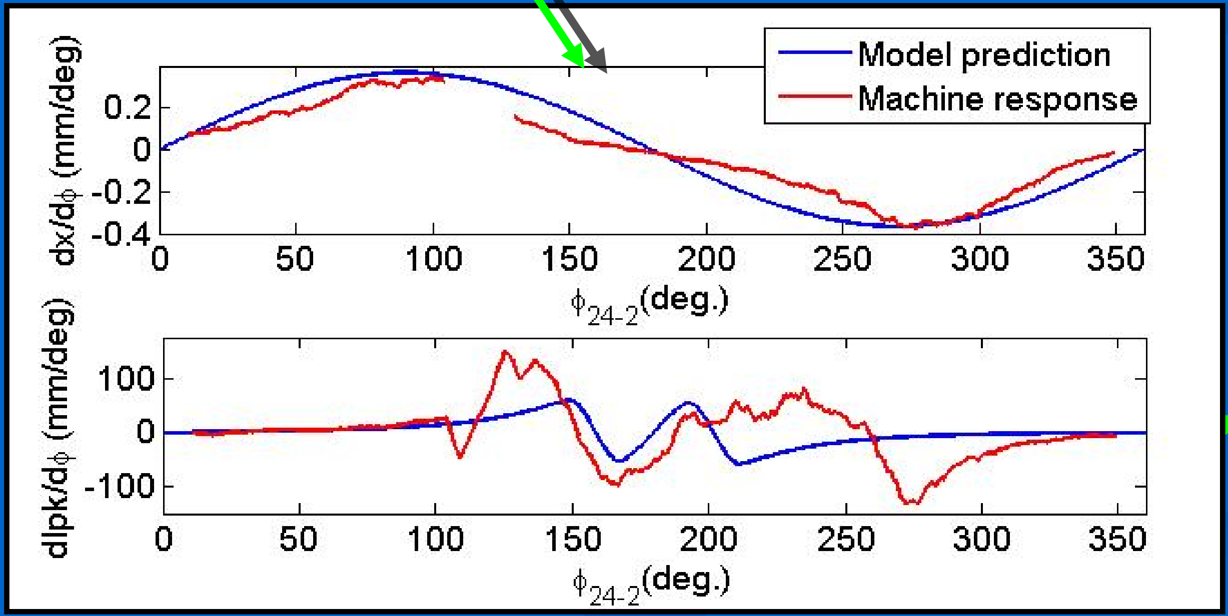
(3) Machine Model

Energy model



Peak current model



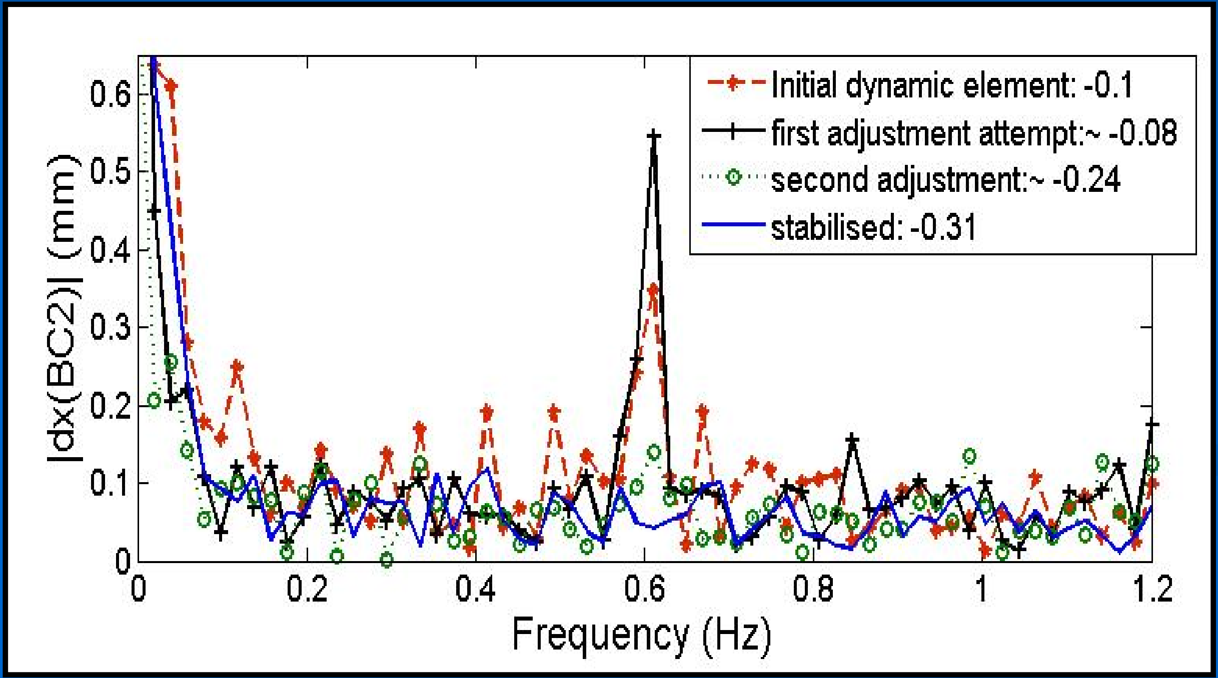


- Energy model sufficient but peak current modeling requires further development.

5. LCLS studies: Results

(3) Online ajustement of M

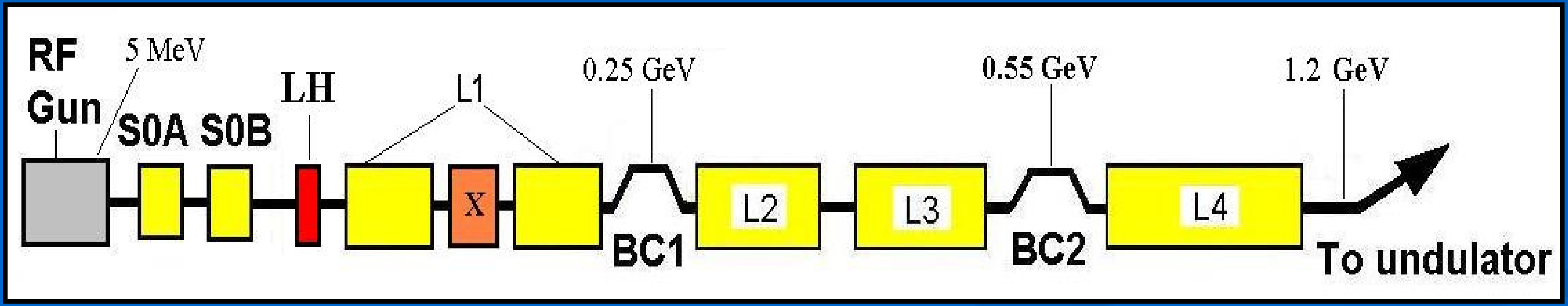
- During control one of the klystron setting is changed from 60° to 300°.
- The corresponding value of the actuator response M=dx/dφ changes from 0.3 (mm/°) to -0.3 (mm/°).
- To test the algorithm, M is set to -0.1 when the phase change occurs.



> - The algorithm then performs a small change to evaluate the proper sign for the adjustment.
> - Based on the position FFT further adjustments are performed.

5. Application to the FERMI case

- *The FERMI machine assembly*



- 1 BC configuration: 3 Observables: E_{BC1}, E_{end}, I_{BC1}
  and 3 Controllables: VL1, VL4, φL1

- 2 BC configuration: 5 Observables: E_{BC1}, E_{BC2}, E_{end}, I_{BC1}, I_{BC2}
  and 5 Controllables: V_{L1}, V_{L4}, φ_{L1}, ?, ?

5. Application to the FERMI case

- *Knowledge gained:*

  - The system can be applied to a multi Observable – multi Controllable case for control over a wider range of frequencies than the PI .
  - Limitations are presently due to the slow communication with the actuators and saturation of the klystrons.
  - Further development of the system is required to ensure adaptability.

- *Next steps:*

  - Definition of the control system requirements and capabilities: actuators availability speed and location.
  - Implementation of a Matlab GUI for the machine up to BC1.

Acknowledgments

➢ FERMI@Elettra*

-Laura Badano, Sandra Biedron, Paolo Craievich, Gerardo D'Auria, Simon Di Mitri, Mario Ferianis, Marco Lonza, Stephen Milton, Giuseppe Penco, Marco Veronese and others

This work was supported in part by the Italian Ministry of University and Research under grants FIRB-RBAP045JF2 and FIRB-RBAP06AWK3

➢SLAC - LCLS

-Ronald Akre, Paul Emma, Franz-Josef Decker, Josef Frish, Juhao Wu and others

➢Australian Synchrotron

-Greg Le Blanc, Rohan Dowd, Eugene Tan

➢Monash University

-Michael Morgan, Charles Grief

(A) Network structure optimization

Criterion 1: The error index

- Measures the quality of the prediction, with *y(t)* the model output and $\hat{y}$*(t)* the targeted value.

$$(1)\quad E.I. = \left[\frac{\sum (\hat{y}(t) - y(t))^2}{\sum y^2(t)} \right]^{1/2}$$

Criteria 2: The correlation tests

- Helps evaluate the number of lagged inputs. If enough lags are provided the conditions (2), (3) and (4) should hold. In practice we use the 95% interval confidence.

- E is the expectation value, ε is the error between the NNET prediction and the desired output, *u* is the input variable (klystron 1 phase or voltage).

$$(2)\quad \phi_{u\epsilon}(\tau) = E(\epsilon(t)u(t-\tau)) = 0$$

$$(3)\quad \phi_{u^2\epsilon}(\tau) = E(\epsilon(t)(u^2(t-\tau) - \bar{u}^2(t))) = 0$$

$$(4)\quad \phi_{u^2\epsilon^2}(\tau) = E(\epsilon^2(t)(u^2(t-\tau) - \bar{u}^2(t))) = 0$$

$$\hat{\phi}_{u\epsilon}(\tau) = \frac{\sum_{t=1}^{N-\tau} u(t)\epsilon(t+\tau)}{\left[\sum_{t=1}^{N} u^2(t) \sum_{t=1}^{N} \epsilon^2(t+\tau) \right]^{\frac{1}{2}}}$$