EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

CERN - PS DIVISION

CERN/PS 95-29 (OP)

DETERMINATION OF BEAM PARAMETERS FOR LEAR WITH NEURAL NETS

M. Ludwig and G. Stimpfl-Abele

Abstract

The transversal local tunes for an unbunched randomly-coasting beam can be determined analytically from Schottky scans. This method is rather slow since an integral equation has to be solved. A very fast approach is described here using neural feed-forward nets. Results obtained with simulated Schottky scans from LEAR are presented. Very good performance is achieved compared to the analytical method.

Fourth International Workshop on Software Engineering and Artificial Intelligence for High Energy Physics, Pisa, Italy, 3-8 April 1995

Geneva, Switzerland
28 juin 1995

CERN PS 95-29 (OP)

DETERMINATION OF BEAM PARAMETERS FOR LEAR WITH NEURAL NETS

*Georg Stimpfl-Abele**

Departament Estructura i Constituents Matèria, Universitat de Barcelona, E-08028 Barcelona, Spain

and

Michael Ludwig

PS Division, CERN, CH-1211, Geneva 23, Switzerland

ABSTRACT

The transversal local tunes for an unbunched randomly-coasting beam can be determined analytically from Schottky scans. This method is rather slow since an integral equation has to be solved. A very fast approach is described here using neural feed-forward nets. Results obtained with simulated Schottky scans from LEAR are presented. Very good performance is achieved compared to the analytical method.

1. Introduction

The root mean-square current of a randomly-coasting particle beam shows tiny fluctuations with time, according to the well known theory of Schottky noise. The frequency spectrum of this current is usually observed with a spectrum analyzer. The measurement and analysis of the so-called Schottky scans serve as a standard tool to evaluate the tune and emittance distribution of such a beam.

Carefully analyzing the Schottky signal with a method described by S. v. d. Meer[1,2] and S. Baird[3] yields the transversal betatron tune qf across the beam for all beam revolution frequencies f.

This method was initially used at CERN's Antiproton Accumulation Complex (AAC) as a nondestructive measurement tool for tune and emittance across the beam. A modified version is currently implemented at CERN's Low Energy Antiproton Ring (LEAR)[4]. This method implies the stepwise solution of an integral equation and is unfortunately quite indirect and rather slow, but it nevertheless has the advantage of being nondestructive.

In the AAC, the evaluation of the transversal local tune qf from Schottky scans is nowadays routinely done with a relative error $\Delta qf/qf$ of the order of 10^{-3}. It is desired to reach the same precision for LEAR.

* current address : PPE Division, CERN, CH-1211 Geneva 23, Switzerland
internet address : stimpfl@cernvm.cern.ch

In this study a fast approach is tried out based on neural nets (NN's). The aim is to feed NN's with Schottky scans and to train them to answer some of the beam parameters. The crucial point is to test whether fast NN's can achieve a similarly high precision as the slow standard algorithm.

2. Input Data

2.1. *Simulation*

An algorithm has been developed to simulate Schottky scans at LEAR based on the machine design and a simplified version of the Schottky-scan theory.[4] The tune distribution q_f can be approximated by

$$q_f = q\left[1 + \xi\frac{\Delta p}{p} + \xi_2\left(\frac{\Delta p}{p}\right)^2\right]. \tag{1}$$

$\Delta p/p$ is the local momentum spread for a subset of particles in the beam, q the mean betatron tune, ξ the mean linear chromaticity and ξ_2 the second-order chromaticity.

q, ξ, ξ_2 and the noise level, specified by the noise-to-signal ratio R, belong to the *variable* input parameters of the simulation algorithm. Other input parameters, like machine design properties, remain constant. For the sake of simplicity the study is restricted to horizontal tunes (constant ξ_2).

2.2. *Data samples*

The input parameters q, ξ and R are varied over a relatively wide range. Three data samples (A, B and E) are generated using a grid of parameters, i.e. each of the three parameters q, ξ and R is changed with constant stepsize. 100 q, 10 ξ and 10 R points are generated giving a total number of 10^4 scan sets in each sample. It is assumed that the spectrum analyzer has 501 channels. Hence each scan sets contains 501 values serving as input for the NN's.

To test the interpolation capability of the nets another sample is generated using flatly-distributed random values of the parameters (sample C).

The samples A, B and C cover an unrealistically large range of tunes. Hence an additional sample (D) is produced by merging the samples A and B and cutting off small and large tunes. Table 1 lists the parameters used in the simulation of the data samples. The minimal and maximal values are given as well as the step sizes (except from the continuous sample C).

Table 1. Variation of the parameters in the data samples.

Sample	Size	q_{min}	q_{max}	Δq	ξ_{min}	ξ_{max}	$\Delta\xi$	R_{min}	R_{max}	ΔR
A	10000	2.	2.4950	0.005	0.	0.9	0.1	0.	0.18	0.02
B	10000	2.0025	2.4975	0.005	0.05	0.95	0.1	0.1	0.19	0.02
C	10000	2.	2.5		0.	1.		0.	0.2	
D	11500	2.1600	2.4650	0.0025	0.	0.95	0.05	0.	0.19	0.01
E	10000	2.2950	2.3049	0.0001	0.	0.18	0.02	0.	0.18	0.02

3. Tests and Results

3.1. *Nets*

Very simple NN's are used with one hidden layer. Since the size of the data samples is small the number of hidden neurons is kept low. Reasonable results are obtained with 25 or 50 hidden neurons. The NN's are used to estimate either one or all three beam parameters. Each output neuron corresponds to one parameter.

The hidden neurons have a sigmoid activation function and the output neurons a linear one. The target values of each output neuron are chosen such, that 0 and 1 correspond to the minimal and maximal value of the corresponding parameter, respectively.

The error of an output neuron is the difference between the values of its output and its target. The nets are trained by minimizing the the sum of the quadratic errors using standard error back-propagation.

3.2. *Basic tests*

Since the number of scan sets in the data samples is rather small compared to the number of energy channels, reduced input sets are used at the beginning. Only half of the energy channels are taken omitting even channel-numbers. Furthermore only 25 hidden neurons are used.

A NN is trained on sample A to estimate the parameters q, ξ and R and is tested with sample B. The error of the parameters q and ξ is shown in fig. 1a and 1b as a function of q. The large fluctuations are not surprising since their are many resonances for small and high tunes.

In a second test the regions with the most pronounced resonances are cut out. Using the same data samples as before only scans with a tune between 2.16 and 2.465 are selected for training and testing. The fluctuations are much smaller now (fig. 1c and 1d).

3.3. *Differences between the learning and the test samples*

The tests described above are repeated using the learning sample (A) for the test instead of the test sample (B). The results are very similar to those obtained previously although the parameter grids of sample A and B are shifted with respect to each other. This shows the ability of the net to interpolate and the absence of overlearning.

Then the samples A and B are merged and a net is trained and tested with this data. An additional test with the random sample C does not reveal significant differences.

Sample E shows the same features. The results obtained by training and testing with the full sample are very similar to those obtained by splitting the sample randomly into a learning and a test sample.

The similarity of the performances using the learning and the test data can be

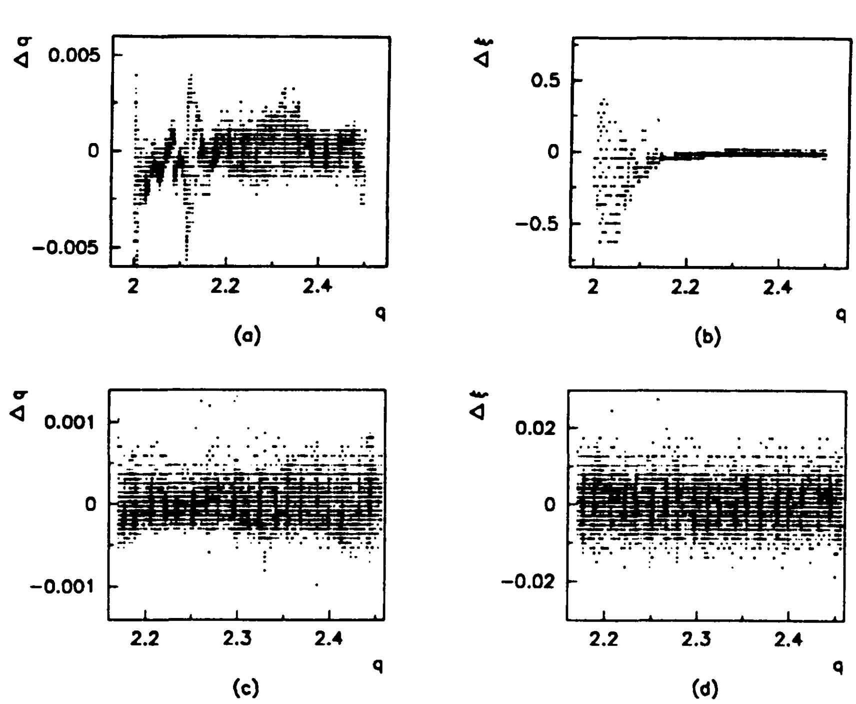

Fig. 1. The errors Δq and Δξ versus q for sample B before (a,b) and after q cut (c,d).

explained by the fact that the data samples contain a complete set of equidistant parameter-values. Hence the nets are trained with all relevant combinations of the parameters. This completeness is normally not achievable with very complicated simulation programs which depend on many parameters with very different distributions. In such cases very many input sets are needed to obtain similar results with learning and test data.

In order to increase the size of the data samples for learning and testing, the same data is used for both purposes in the following tests. The performance of the nets for both a wide and a small range of tunes is studied with samples D and E, respectively.

The noise-to-signal ratio R can easily be determined by the NN's. A resolution (root mean-square error) of about 4×10^{-4} is obtained in all tests, which is very good compared to other methods. This parameter is therefore omitted in the following discussion of the results.

3.4. *Output correction and error normalization*

To probe the results more deeply two separate NN's are trained with sample D to estimate independently q and ξ. The distribution of the error Δq versus q now shows a wave pattern (fig. 2a). The fact that Δq is small compared to q allows to iron out the waves in the test by correcting for these systematic deviations (fig. 2b).

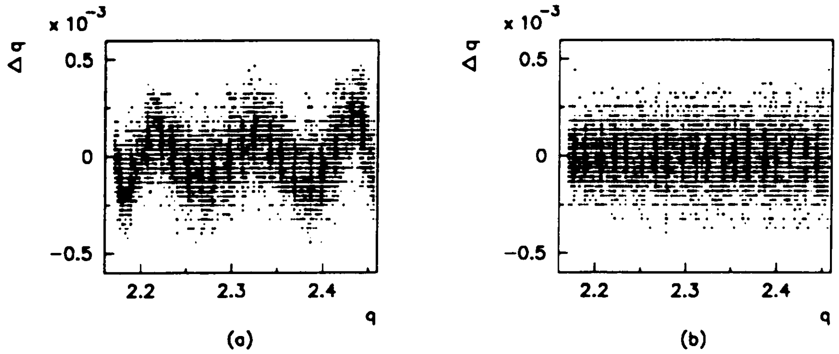

Fig. 2. The error Δq versus q for sample D before (a) and after (b) correction.

The poorer performance of the net with three outputs compared to the net exclusively trained for q may be explained by the differences in the resolution of the parameters. Since the nets are trained to minimize the sum of the quadratic errors of all outputs, a parameter with higher resolution (and hence smaller errors) has less influence on the weight modification than one with lower resolution. Taking this into account in the training by minimizing normalized output-errors (like in a χ^2 fit) almost the same results are obtained with a single NN as with special nets.

Similar effects are observed for sample E. Higher resolution is achieved because the parameter space is much smaller than for sample D. The resolution (root mean-square error) of q and ξ obtained in these tests is given in Table 2.

Table 2. Resolutions of q and ξ obtained with sample D and E using nets with N_{out} outputs.

Sample	N_{out}	Δq	Δq_{corr}	$\Delta\xi$
D	1	1.5×10^{-4}	1.2×10^{-4}	4.9×10^{-3}
D	3	1.6×10^{-4}	1.3×10^{-4}	6.0×10^{-3}
E	1	8.4×10^{-5}	5.0×10^{-5}	2.6×10^{-3}
E	3	8.2×10^{-5}	5.0×10^{-5}	3.6×10^{-3}

The dependencies of Δq and $\Delta\xi$ on q and ξ for sample E are shown in fig. 3. The variations with q are still important.

3.5. *Increasing the size of the nets*

Increasing the number of hidden neurons from 25 to 50 does not improve the results. This may be due to the limited size of the data samples. But taking all 501 channels instead of only 251 increases the performance significantly since more information is presented to the nets. Again no difference is found between using 25 and 50 hidden neurons.

The resolutions obtained with sample D are 1.1×10^{-4} for q, 9.3×10^{-5} for q_{corr} and 4.1×10^{-3} for ξ. With sample E the results for q, q_{corr} and ξ are 4.0×10^{-5}, 3.7×10^{-5} and 4.0×10^{-3}.

The beam parameters are determined by scaling (and shifting) the NN outputs.

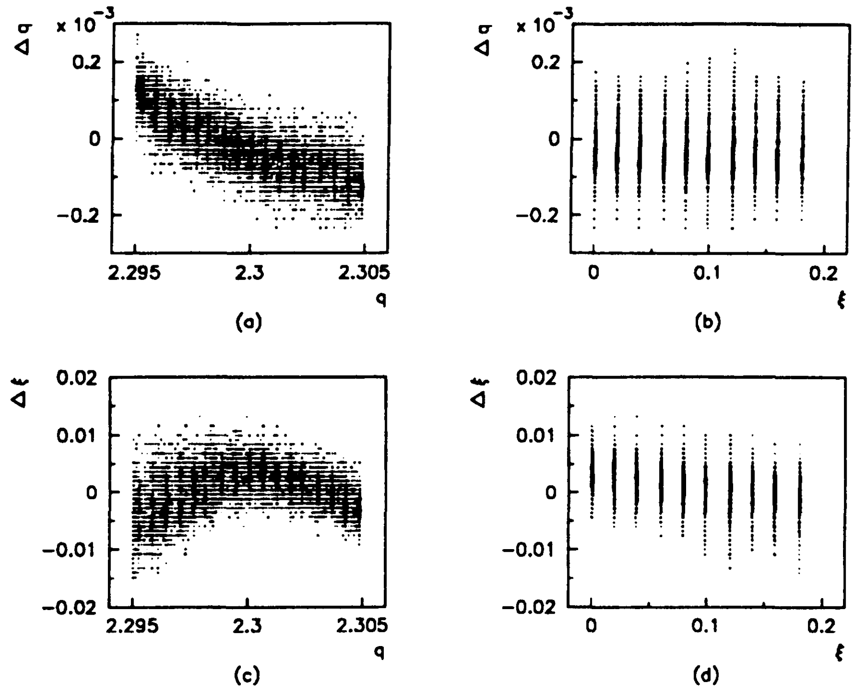

Fig. 3. The errors Δq and $\Delta\xi$ versus q and ξ for sample E.

The highest output-resolution achieved in the tests is 3.1×10^{-3} obtained for q_{corr} using sample D with 501 inputs.

4. Conclusions

Very high precision can be obtained with simple NN's. Since the relative errors on the tune are very small it is possible to use NN's with a very limited number of hidden neurons and to correct the output for systematic deviations. The best resolution achieved in this study is smaller than 4×10^{-5}. The normalization of the output errors allows one single net trained for the determination of several parameters to produce a similar performance as nets trained to estimate only one parameter. Work is continuing to estimate the whole tune distribution q_f with only one net.

References

1. S.v.d. Meer, Diagnostics with Schottky noise, CERN/PS Report 88-60 (AR), 1988.
2. S.v.d. Meer, Measurements of transverse stack emittances from Schottky scans, CERN/PS Report 84-11 (AR), 1984.
3. S. Baird, Schottky noise: what can we observe ?, CERN/PS Report 89-06 (OP), 1989.
4. M. Ludwig, Measurement of local tune and emittance from Schottky scans in LEAR, CERN/PS Note 94-63 (OP), 1994.