The Discovery of Direct CP Violation

L. Iconomidou-Fayard and D. Fournier

LAL, Univ Paris-Sud, CNRS/IN2P3, Orsay 91400, France

lyfayard@in2p3.fr
daniel.fournier@cern.ch

Soon after the discovery in 1964 of the non-conservation of CP symmetry in the neutral kaon system, the hunt was launched for a component arising from direct violation on top of the dominant effect due to mixing. It took almost 20 years until the first evidence of a signal was reported by NA31 and another 10 years to establish the effect with a significance of more than 5 standard deviations. This article describes the beams, detectors and analysis methods used by the two CERN experiments, NA31 and NA48, which made key contributions to these results and established new standards for precision measurements.

1. Introduction

1.1. The early days of CP violation

In a world where CP is conserved, a $K^0$ produced by the interaction of a proton beam on a target evolves as the coherent superposition, in equal proportions at the time of production, of $K^0_1$ and $K^0_2$, where $K^0_1 = (K^0 + \bar{K}^0)/\sqrt{2}$ is the positive CP eigenstate, and $K^0_2 = (K^0 - \bar{K}^0)/\sqrt{2}$ is the negative one. Being positive under CP, the $K^0_1$ decays rapidly to the $\pi\pi$ final state. On the contrary $K^0_2$ has a much longer lifetime being forced to decay mainly to $3\pi$ or semi-leptonic $\pi\ell\nu$ final states, disfavoured by phase space. The two lifetimes are very different, $c\tau_S = 2.68$ cm for the former and $c\tau_L = 15.34$ m for the latter, which allows a very clean separation of the two eigenstates: far enough from the target, for example after 20 or more $\tau_S$, one does not expect to find any two-pion decay.

In 1964, at BNL, Christenson, Cronin, Fitch and Turlay observed for the first time a significant number of two pion decays after enough $c\tau_S$ to demonstrate that CP is not conserved. The simplest explanation, confirmed in 1967 by the observation of an asymmetry between $\pi^+\ell^-\bar{\nu}$ and $\pi^-\ell^+\nu$ rates in neutral kaon decays, is to consider that the long and short lifetimes mass eigenstates, $K^0_L$ and $K^0_S$, are not pure CP eigenstates, but contain a small admixture of the opposite CP species, described by a complex parameter $\epsilon$:

$$K^0_L = \frac{(K^0_2 + \epsilon K^0_1)}{\sqrt{(1 + \epsilon^2)}}$$

$$K^0_S = \frac{(K^0_1 + \epsilon K^0_2)}{\sqrt{(1 + \epsilon^2)}}$$

(1)

(2)
This is called “CP violation in the mixing” or “indirect CP violation”. The CP violating parameter $\epsilon$ was soon after measured with good precision. Today $|\epsilon| = (2.228 \pm 0.011) \times 10^{-3}$ and $\phi_\epsilon = (43.52 \pm 0.02)^\circ$.

1.2. Basic phenomenology

L. Wolfenstein\(^3\) postulated in 1964 that there exists a “superweak” interaction which manifests itself only in the $\Delta S = 2$ mixing between $K^0$ and $\bar{K}^0$, and does not appear in the $\Delta S = 1$ transitions of the neutral kaon decays.

A direct consequence of this assumption is that $\eta^{00} = \eta^{+ -}$ with:

\[
\eta^{00} = \frac{A(K_L \to \pi^0\pi^0)}{A(K_S \to \pi^0\pi^0)},
\]

\[
\eta^{+ -} = \frac{A(K_L \to \pi^+\pi^-)}{A(K_S \to \pi^+\pi^-)}.
\]

Contrary to the superweak hypothesis, one can expect on general grounds that CP violation may be present in any weak decay, leading to a difference between the amplitude of a particle $X$ decaying to a final state $f$, $A(X \to f)$ and the amplitude of the CP-conjugate particle $\bar{X}$ to $\bar{f}$, $A(\bar{X} \to \bar{f})$. In the case of the neutral kaon system this means that the decay amplitude of the CP-odd combination $K^0_2$ to a $\pi\pi$ final state might differ from zero, corresponding to a “direct CP violation”.

Given the conservation of CPT, direct CP violation may not lead to visible effects. In a general manner, the direct CP-violating transition to a given final state can be non-zero provided there are at least two amplitudes leading to this same final state, each with different phases. This is the case for a kaon decaying into two-pions through two decay amplitudes, one in the $I = 0$ and the second in the $I = 2$ isospin states. As the Clebsch–Gordan coefficients projecting the $I = 0$ and $I = 2$ states onto the $\pi^+\pi^-$ and $2\pi^0$ final states are different, in the presence of direct CP violation $\eta^{00}$ and $\eta^{+ -}$ are no longer equal.

More precisely:

\[
\eta^{+ -} = \epsilon + \epsilon', \quad \eta^{00} = \epsilon - 2\epsilon'
\]

with

\[
\epsilon' = \frac{i}{\sqrt{2}} \text{Im} \left( \frac{A_2}{A_0} \right) e^{i(\delta_2 - \delta_0)}
\]

where $A_{0,2}$ and $\delta_{0,2}$ are the amplitudes and the strong phases of the two-pion final isospin states $I = 0$ and $I = 2$. Conversely:

\[
\text{Re} \left( \frac{\epsilon'}{\epsilon} \right) = \frac{1 - \left| \frac{A_2}{A_0} \right|^2}{6} = \frac{(1 - \text{RR})}{6}
\]
The Discovery of Direct CP Violation

where RR is the so-called “double ratio”, defined as:

\[
RR = \frac{\Gamma(K^0_L \rightarrow \pi^0\pi^0)}{\Gamma(K^0_S \rightarrow \pi^0\pi^0)} / \frac{\Gamma(K^0_L \rightarrow \pi^+\pi^-)}{\Gamma(K^0_S \rightarrow \pi^+\pi^-)}
\] (8)

Measurements of \(\pi\pi\) phase shifts\(^4\) show that the phase of \(\epsilon'\) is approximately equal to that of \(\epsilon\).

As demonstrated by Kobayashi and Maskawa, CP violation is possible in the Standard Model when three generations of weakly interacting quark doublets are present.\(^5\) The quark mass eigenstates and the quark flavour eigenstates are related by the \(3 \times 3\) complex and unitary CKM matrix \(V_{ij}\), parametrised by three rotation angles and one phase \(\delta\). The amplitude of CP violation is determined by the Jarlskog invariant,\(^6\)

\[
J_{CP} = \sin \phi_{12} \times \sin \phi_{13} \times \cos \phi_{23} \times \cos \phi_{13}^2 \times \cos \phi_{23} \times \sin \delta
\] (9)

representing the common area of triangles resulting from the unitarity conditions \(V_{ij}^* V_{ik} = 0\) (with \(j \neq k\)). The index \(i\) \((j, k)\) runs over the three up (down) quarks.

The key parameter is the phase \(\delta\), origin of the CP violation.

In the late 70s to early 80s, several calculations of \(\epsilon'\) and \(\epsilon\) were available, which however were not very precise for the latter, mainly due to “long range” effects. The most representative calculations of \(\epsilon'/\epsilon\) at this time\(^7–9\) were ranging from 0.002 to 0.02.

Later on in the early 90s, new theoretical developments brought the various teams working in the field\(^10,\ 11\) to converge towards a quite small central value, typically below \(5 \times 10^{-4}\). This was due to large cancellations between the amplitudes of two \(\Delta S = 1\) diagrams, namely the so-called “penguins” electromagnetic and QCD diagrams,\(^12\) and to the increasing lower bound on the top quark mass.\(^13\) The uncertainties on the predictions were, however, quite important.

1.3. Experimental situation on \(\epsilon'/\epsilon\) in the 80s–90s

The experimental precision is, in general, statistically limited by the measurement of \(\eta^{00}\) which requires the detection of the four photons from the two \(\pi^0\) decays. The most precise results available in the early 1980s were: \(|\eta^{00}/\eta^{+-}| = 1.00 \pm 0.06\) obtained at the CERN PS\(^14\) with about 45 reconstructed \(K^0_L \rightarrow \pi^0\pi^0\) decays, \(|\eta^{00}/\eta^{+-}| = 1.03 \pm 0.07\)\(^15\) obtained at the BNL AGS with about 120 reconstructed \(K^0_L \rightarrow \pi^0\pi^0\) decays, and \(|\eta^{00}/\eta^{+-}| = 1.00 \pm 0.09\) obtained at the BNL AGS\(^16\) in the region of \(K^0_S-K^0_L\) interference which allows one to measure both the magnitude and phase of \(\eta^{00}\) (see Section 3.3).

Based on the rather large theoretical predictions that were valid in early 80s, several experiments were proposed in the US, at BNL\(^17\) and at Fermilab,\(^18,\ 19\) At CERN, an initiative was undertaken in the early 80s which gave birth to the NA31 experiment, which was formally approved in 1982. NA31 was aiming at a total
accuracy on $\epsilon'/\epsilon$ of about one permil, using more intense beams than in the past and methods that partially cancel the systematic uncertainties of the measurement.

After this round, in the early 90s, the need for a new $\epsilon'/\epsilon$ measurement with precision as high as $10^{-4}$ became clear from both the theoretical and the experimental status and this resulted in a new generation of experiments, KTeV at Fermilab and NA48 at CERN.

The NA31 and NA48 experiments were both installed on beams from the 450 GeV Super Proton Synchrotron (SPS), in the EHN1 and EHN2 areas respectively.

1.4. The main challenges in the measurement of $\epsilon'/\epsilon$

A precise measurement requires both high accumulated statistics for all the four channels, especially for the most suppressed $K^0_L \to 2\pi^0$ mode, and small and well-controlled systematic uncertainties.

The statistical accuracy can be improved by means of intense proton beams hitting fixed targets and producing $K^0_S$ and $K^0_L$ beams. At CERN, this was possible due to the availability of the 450 GeV SPS. Neutral beams produced by proton interactions on a target contain practically equal amount of short-lived and long-lived kaons. Because of the very different lifetimes, the $K^0_S$ component will rapidly decay, unlike the $K^0_L$ one. This requires the production of the $K^0_S$ and $K^0_L$ beams at two different distances, close and far from the detectors respectively. To create the $K^0_S$ beam, both CERN experiments prefered the solution of protons hitting a second target, instead of regenerating the $K^0_S$ component from the $K^0_L$ beam as done traditionally by the pioneering groups and also in the Fermilab experiments. The CERN choice was made in order to avoid the drawbacks of the regeneration technique: the inelastic $K^0_L$ interactions and the interference occuring between the original CP violating $K^0_L$ component and the regenerated $K^0_S$ one, that requires several auxiliary measurements of the regeneration parameters.

A second challenge concerns the collection of the events of interest. The measurement of $\epsilon'/\epsilon$ consists in counting events from the four modes, $K^0_S$ and $K^0_L$ decaying into $2\pi^0$ and $\pi^+\pi^-$, in order to build the double ratio $RR$ defined in Eq. (8). While in $K^0_S$ the two-pion channels saturate the decays, in $K^0_L$ these final states count for less than 0.3% of the total decay rate. The CP-violating channels in $K^0_L$ decay must therefore be efficiently identified in the presence of the large amount of three-body final states which have to be suppressed by means of a powerful trigger and accurate measurements.

\[60 \text{ Years of CERN Experiments and Discoveries} \]
Third, despite the severe collimation, intense beams are often accompanied by particles that can affect a good event resulting in its loss. While small, these effects may impact the measurement of the double ratio if they alter differently the four decay modes. The measurement of the accidental activity was one of the challenges for precision experiments which developed a series of subtle analyses, in particular the so-called overlay method, to investigate the effect. This method consists in superimposing by software specific so-called “random” triggers onto each event in the data, where the “random” triggers are recorded in proportion to the beam intensities to give an accurate picture of the ambient activity.

Finally, particular care is needed to reduce the systematic effects, for example by concurrently recording — whenever possible — the four modes of interest thus leading to cancellations, and by designing the experiments so as to minimise the need for corrections.

2. First Generation: The NA31 Beams and Detectors

The project proposed by four institutes aggregated from 1982 to 1993 up to 60 physicists from seven European institutes. The first important data taking period took place in 1986. Additional data were recorded in 1988 and 1989, after beam and detector upgrades, in order to improve both the statistical and the systematic uncertainties.

2.1. The $K_0^L$ and $K_S^0$ beams

The choice of the NA31 collaboration was to alternate the data taking between coaxial $K_S^0$ and $K_L^0$ beams, with a typical cycle of 30 hours. As the apparatus registers concurrently charged and neutral decays, a part of systematic uncertainties cancels out.

The $K_L^0$ beam is produced from an extracted proton beam from the SPS hitting a far target placed $\sim 120 \text{ m}$ upstream the decay volume. The construction of a high intensity $K_L^0$ beam requires the sweeping of the charged secondaries and careful multistage collimation of the neutral component to precisely define the beam aperture and to remove scattered halo particles. The $K_S^0$ beam is created by transporting an attenuated proton beam to a second target, located very close to the decay region. The beam parameters are listed in Table 1 and a sketch of the experimental layout is shown in Fig 1. In the $K_S^0$ beam, it is essential to define precisely the beginning of the decay volume against the earlier decays inside the collimators. This is done by positioning a scintillator, used as veto, right after the $K_S^0$ defining collimator. In order for this counter to veto also neutral decays with high efficiency, a 7 mm lead sheet is positionned in front of it. This “antiKS counter” (AKS) also plays an essential role in the energy scale determination and control.
Table 1 Parameters of $K^0_L$ and $K^0_S$ beams. In parentheses the modified values used after the 1986 run are shown.

<table>
<thead>
<tr>
<th>Beam type</th>
<th>$K^0_L$</th>
<th>$K^0_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam energy (GeV)</td>
<td>450</td>
<td>450 (360)</td>
</tr>
<tr>
<td>Length of beam from target to exit of final collimator (m)</td>
<td>48.0</td>
<td>120.0</td>
</tr>
<tr>
<td>Dist final coll-LAr Calo(m)</td>
<td>123.8</td>
<td>76.7 to 124.7</td>
</tr>
<tr>
<td>Be target diam/length (mm)</td>
<td>2/400</td>
<td>2/400</td>
</tr>
<tr>
<td>Production angle (mrad)</td>
<td>3.6 (2.5)</td>
<td>3.6 (4.5)</td>
</tr>
<tr>
<td>Beam acceptance (mrad)</td>
<td>±0.2</td>
<td>±0.5</td>
</tr>
<tr>
<td>Protons on target per pulse</td>
<td>$1 \times 10^{11}$</td>
<td>$3 \times 10^{7}$</td>
</tr>
<tr>
<td>$K^0$ per pulse, at prod., in beam accept.</td>
<td>$1.8 \times 10^9$</td>
<td>$3.3 \times 10^3$</td>
</tr>
<tr>
<td>Neutron per pulse in beam accept.</td>
<td>$1.5 \times 10^7$</td>
<td>$3 \times 10^4$</td>
</tr>
</tbody>
</table>

Fig. 1. Layout of the NA31 experiment with, from left to right: the $K^0_L$ beam, the $K^0_S$ beam, the evacuated decay volume with anticounters, the thin window, the two wire chambers, the liquid argon calorimeter, the hadronic calorimeter and the muon veto.

In the $K^0_L$ mode the distribution of the decay vertices along the beam axis is an exponential with a slope of about 5 m at 100 GeV. This is by far different from the flat $K^0_L$ distribution. To mitigate this effect, the $K^0_S$ beam elements were installed on supports movable on rails, the “XTGV”, positioned for the necessary running time at fixed stations every 1.2 m along the 50 m decay region. Thus, after summing the data from all stations the overall vertex distribution of the $K^0_S$ events became to a large extent similar to the $K^0_L$ one.

Despite this trick, the energy spectra still differ between the two beams, because the $K^0_S$ collimation length selects preferentially the more energetic kaons. This
difference amplifies the sensitivity of the result by about a factor of 3 to a possible
difference of the energy scale between charged and neutral events. To reduce this
effect, some beam parameters were modified in 1988: the energy of the proton beam
striking the $K_S^0$ target and the incident angle of the protons on both $K_S^0$ and $K_L^0$
targets were tuned so that the two energy spectra measured in the decay volume
became more similar, reducing in a significant way the sensitivity of the result to
their difference.

2.2. The NA31 experimental layout

The $K_S^0$ XTGV is enclosed in a 130 m long cylinder of 2.4 m diameter that contains
the decay volume. The first 100 m are evacuated down to about $3 \times 10^{-3}$ Torr
in order to prevent $K_S^0$ regeneration from $K_L^0$ and multiple scattering of charged
decay pions. A thin kevlar window separates the evacuated part from an enclosure
filled with Helium where the two wire chambers are installed. The Helium enclosure
ends with an Aluminium cover, shaped to withstand the forces generated by a
breakdown of the kevlar window in case of an accident. Downstream of it are the
scintillator hodoscopes, the electromagnetic liquid argon calorimeter (LAC) and
hadronic (HAC) calorimeters, and the muon veto. All detectors are traversed by a
central evacuated pipe for the beam, ending at the beam dump.

Four rings of anticounters are used to veto three-body decays. A hodoscope of
scintillation counters is placed in front of the LAC to start the trigger for charged
decays. A valid charged trigger requires two hits in opposite quadrants. A second
hodoscope, inserted at mid depth in the LAC is used to trigger the acquisition of
neutral candidates. A valid neutral trigger requires a left–right coincidence.

Downstream of the first (second) 80 cm iron slab of the muon filter which
follows the calorimeters is a hodoscope of horizontal (vertical) scintillator slabs
each covering the full width of 2.7 m and read out at both ends. This detector is
used as muon veto.

The measurement of both decay modes is based on calorimetry (electromagnetic
and hadronic) and uses the wire chambers for reconstructing the decay vertex and
trajectories of charged particles.

2.3. Measuring the neutral decays: Liquid argon calorimeter

To achieve the needed discrimination between the $2\pi^0$ signal and the many times
more abundant $3\pi^0$ background requires a calorimeter with a high level of energy
and space resolution. A liquid argon ionisation calorimeter (LAC) has been used
for this purpose. Its body consists of a stack of lead conversion plates, aluminium
clad, alternating with printed-circuit boards for signal readout. Together with the
liquid argon, its total thickness is about 24 radiation lengths ($X_0$). The calorimeter
is physically divided into a left and a right part, each of them split in a front and a
back part, with the neutral hodoscope fitted in between. The important parameters
Table 2  Main parameters of the liquid argon calorimeter.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of the Pb/Al sandwich plates (mm)</td>
<td>1204 × 2408 × 2.3</td>
</tr>
<tr>
<td>Size of the readout boards (mm)</td>
<td>1200 × 2400 × 0.8</td>
</tr>
<tr>
<td>Length of one cell (mm/rad.length)</td>
<td>7.3/0.3</td>
</tr>
<tr>
<td>Number of cells in depth (front+back)</td>
<td>40 + 40</td>
</tr>
<tr>
<td>Strip pitch (mm)</td>
<td>12.5</td>
</tr>
<tr>
<td>Number of cells per quadrant (X–Y)</td>
<td>96–96</td>
</tr>
<tr>
<td>Number of electronics channels</td>
<td>1536</td>
</tr>
</tbody>
</table>

Fig. 2. Display of a $3\pi^0$ candidate event. On the borders of the squared area representing the LAC surface are drawn peaks corresponding to the energy deposition in the horizontal and vertical calorimeter strips. The reconstructed photon positions are denoted by the letters $G_i$. The positions of the three $\pi^0$ are indicated by $\pi_i$. Different colours are used for each $\pi^0$ and its decay products.

are given in Table 2. The readout boards are etched to form 96 parallel strips per quadrant, alternately vertical and horizontal. Corresponding strips are ganged by groups of 20 in depth, to form readout cells. With this scheme showers are measured by their horizontal and vertical projections, separately for each quadrant.

In order to optimise the signal to noise ratio, each readout cell is coupled to its preamplifier through a low impedance cable and a transformer. The electronic circuit features a 1.6 $\mu$s differentiation and two 50 ns integration steps. As a result, the charge signal fed to a 12-bit ADC has a parabolic rise, lasting about 500 ns, followed by a slow decay. A second output is fed to the “peak-finder” system used to count the number of photons in the calorimeter, at the trigger level. Figure 2 shows the event display of a $3\pi^0$ candidate.

A zero-suppression system, with a threshold of 150 MeV (about 9 times the noise of typically 16 MeV per channel) combined to a neighbourhood logic, is used to limit the readout bandwidth.
The whole calorimeter front face was exposed several times to an electron beam to measure \textit{in situ} the response on a grid of points. From these measurements, the sampling term of the energy resolution was found to be equal to $7.5\% / \sqrt{E(\text{GeV})}$, the uniformity within $\pm 0.5\%$, and the linearity within $\pm 0.3\%$ between 12 and 120 GeV. The electronic noise contributes about 100 MeV in the resolution.

Electronic calibrations were taken regularly and revealed small drifts of the pedestal and channel gains, the latter being associated to temperature variations in the hall (typically 0.15\% per degree). An hourly correction based on the measured temperature was applied to compensate this effect.

For the 1989 run, it was decided to complement the charge readout by a precise time measurement based on “zero-crossing” TDCs (ZTDC). Any calorimeter activity in the time interval from 0.9 $\mu$s before to 3.0 $\mu$s after the passage of particles is recorded, with an average time resolution for photons of 10 ns. This implementation allowed to cross check the effect of accidental activity in the measurement of neutral decays, otherwise estimated by the overlay method.

2.4. \textit{Measuring the charged mode}

The direction of charged particles is measured by the two wire chambers. Each track energy is computed combining the energies measured in the LAC and HAC. Each chamber consists of four planes with wires (at a 6 mm pitch) along the vertical, horizontal, U (53 degrees from horizontal) and V (perpendicular to U) directions, positioned so as to minimise ambiguities in space-point reconstruction. Three hits among four planes are enough to reconstruct a space point. The average efficiency per plane is 99.3\% with a precision on the space points of 750 $\mu$m. The decay vertex is reconstructed with a longitudinal precision of 80 cm and an rms of 5 mm on the closest distance of approach.

The HAC is a sampling device, with 25 mm iron sheets alternating with planes made out of $1.3\, \text{m} \times 0.12\, \text{m}$ scintillator slabs of 4.5 mm thickness forming a quadrant structure. There are 24 (25) planes, alternately with horizontal and vertical scintillator slabs in the front (back) module with a total of 176 channels.

After the analysis of the first data taking period, a transition radiation detector was built and installed in 1988 between the second wire chamber and the Aluminium cover of the large tube, to independently cross check the estimation of the $K^0_L \rightarrow \pi^+\pi^-\nu\bar{\nu}$ background. The detector consists of four identical ensembles, each made of a radiator of polypropylene foils followed by a chamber operated with a Xe–He–CH$_4$ mixture. The detector offered an additional electron–pion separation resulting in a rejection power of 10 against electrons for a 98\% pion efficiency.$^{22}$

2.5. \textit{Trigger, online background rejection and data acquisition}

The trigger and data acquisition were designed to cope with the memory time of the detector (2.1 $\mu$s, determined by the LAC), the single rate in the detectors, about
100 kHz, to allow a throughput of about 1 Mbyte/s corresponding to about 2000 accepted events per burst in the $K_0^L$ mode, mostly of three-body decay background despite a high rejection by the three levels of the trigger.

The first level performs a fast selection of events consistent with either two charged particles, or at least two electromagnetic showers using signals from the hodoscopes, LAC and HAC. Furthermore, for the neutral trigger, less than 5 peaks in both the horizontal and the vertical strips of the front part of the LAC are required, using the “peakfinder” system.

At the second level, a custom made hard-wired processor (AFBI) with 150 ns cycle time treats the neutral and the charged conditions in two parallel streams. The neutral stream uses all readout strips to calculate the total energy, the barycenter position in the transverse plane and the longitudinal decay vertex position $Z_V$. This latter quantity is calculated from the second moments, assuming that the clusters correspond to the decay of an object with the kaon mass. The $3\pi^0$ decays with missing clusters are reconstructed with a vertex closer to the calorimeter than the real decay position, and are thus preferentially rejected by the condition $Z_V < 50$ m counted from the $K_0^S$ collimator. The charged stream requires conditions on the energy deposited in the LAC and HAC compartments. About 50% (30%) of neutral (charged) triggers are rejected by the AFBI, while the loss of good $2\pi$ decays is less than 0.1%. Events which have both a neutral and a charged trigger are all accepted.

Finally, at the third level, data from events accepted by the AFBI are loaded into the input memory of a dual 168E processor where some of the offline tracking calculations are made for charged triggers, which allows to cut on $Z_V$, on the acoplanarity, and on the energy assuming a kaon mass. In the $K_0^L$ mode, the 168E accepts about 50% of the charged triggers, including 15% which are too complicated to be treated in the allowed time of 1ms, and 10% of downscaled events used for efficiency calculations. Accepted events are written to tape for offline analysis.

3. The NA31 Analysis and Result

In the following, the descriptions and numerical results are given mainly for the first data taking period (1986). Tables allow the comparison between the running periods.

3.1. Analysis

The $K^0 \rightarrow 2\pi^0 \rightarrow 4\gamma$ decays are reconstructed from the measured energies and positions of photons in the calorimeter. The energy of a photon results from the sum of the cluster energies in the two projections within a fixed size window, and is corrected for leakage outside. Positions result from a barycenter using a maximum of 15 strips in each projection. Events with a fifth photon of more than 2.5 GeV are rejected. Valid photons have to be above 5 GeV, and to be more than 5 cm
The energy is measured as the sum of energies of the four photons with about 1% precision. Assuming a kaon mass, the distance from the decay vertex to the calorimeter is computed with a similar relative precision. To reject the remaining background mostly due to $K^0 \rightarrow 3\pi^0 \rightarrow 6\gamma$ in the $K_S^0$ beam, a $\pi^0$ mass constraint on the three two-photon combinations is applied. The best two-photon pairing is given by the smallest $\chi^2$, shown in Fig. 3. The signal region is taken as $\chi^2 < 9$. The residual background is estimated by linear extrapolation from the large $\chi^2$ region to the signal region, to be $(4.0 \pm 0.2)\%$, where the error includes statistics and systematics. The numbers of neutral events available in the $K_S^0$ and $K_L^0$ beams are given in Table 3.

The $K^0 \rightarrow \pi^+ \pi^-$ decays are reconstructed from space points in the chambers. Events with more than two space points in the first chamber are rejected. The $\pi^+ \pi^-$ invariant mass, calculated from the angle between the two tracks and the calorimetric energy associated with each of them, is required to be within $2.1\sigma$

---

The resolution is about 2 MeV.

For consistency, neutral events with one or more space points in the first chamber are also rejected.
### Table 3

Event statistics accumulated in the NA31 data taking periods for the four modes, together with the estimated background fractions.

<table>
<thead>
<tr>
<th>Decay mode</th>
<th>1986 data Events($\times 1000$)</th>
<th>Background(%)</th>
<th>1988 + 1989 data Events($\times 1000$)</th>
<th>Background(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^0_L \rightarrow \pi^0\pi^0$</td>
<td>109</td>
<td>4.0</td>
<td>319</td>
<td>2.67</td>
</tr>
<tr>
<td>$K^0_L \rightarrow \pi^+\pi^-$</td>
<td>295</td>
<td>0.6</td>
<td>847</td>
<td>0.63</td>
</tr>
<tr>
<td>$K^0_S \rightarrow \pi^0\pi^0$</td>
<td>932</td>
<td>&lt;0.1</td>
<td>1322</td>
<td>0.07</td>
</tr>
<tr>
<td>$K^0_S \rightarrow \pi^+\pi^-$</td>
<td>2300</td>
<td>&lt;0.1</td>
<td>3241</td>
<td>0.03</td>
</tr>
</tbody>
</table>

($\sigma \sim 20$ MeV) from the neutral kaon mass. The $K^0$ energy is calculated from the opening angle of the two tracks, and the ratio $R$ of their calorimetric energies, assuming a kaon mass

$$E = \frac{1}{\theta} \sqrt{\left(2 + \frac{1}{R}\right) \left(m_K^2 - m_\pi^2 \left(2 + \frac{1}{R}\right)\right)}. \quad (10)$$

The ratio $R$ is limited to 2.5 in order to achieve a 1% resolution and to reduce $\Lambda$ decays to a negligible level. Events with additional photons are rejected. $K_\Xi$ decays are suppressed by requiring that none of the two tracks pass “electron-like” shower-shape criteria. About half of $\pi^+\pi^-$ decays are rejected by the cuts. The response of the HAC, monitored by sending laser pulses to all photomultipliers at a regular pace, was estimated to stay constant within ±0.5%, which ensures a stability of the acceptance ratio $K_L \rightarrow \pi^+\pi^- : K_S \rightarrow \pi^+\pi^-$ better than 0.1%. The remaining background from three-body $K^0_L$ decays is estimated using the distribution of the transverse distance $d_T$ from the target to the decay plane (Fig. 3). Extrapolating from the control region $7 < d_T < 12$ cm under the signal region defined as $d_T < 5$ cm, a background fraction of $(0.65 \pm 0.2)%$ is estimated and subtracted. The numbers of charged events available in the $K^0_S$ and $K^0_L$ samples are given in Table 3.

Fitting the decay vertex spectrum of $K^0_S$ decays in charged or neutral mode to the simulated distribution shaped by the AKS located at different longitudinal positions allows to fix the relative energy scale between charged and neutrals with an uncertainty of ±0.1%. An illustration of this fit, for energies between 70 and 170 GeV, is given in Fig. 4(left).

#### 3.2. The NA31 results

Once the energy scale is fixed, events are counted in 10 bins of energy between 70 and 170 GeV, and 32 bins in vertex position between 10.5 and 48.9 m. The chosen decay range is defined to minimise the pollution from kaon scattering and regeneration on the final collimator and to limit the $3\pi^0$ background which increases with $Z_V$. The weighted average of the double ratio, background subtracted and corrected for acceptance and resolution, is $RR = 0.977$. The effect of accidental
coincidences between a kaon decay and some activity in the beam is estimated by overlaying data with random events. The net effect (losses-gains) on the double ratio is $-0.34 \pm 0.10\%$. Other small corrections are applied for acceptance and scattering differences between the four modes, trigger and AKS anticounter inefficiencies, leading to an overall correction of 0.003. The final result is $RR = 0.980 \pm 0.004 \text{ (stat)} \pm 0.005 \text{ (syst)}$, giving:

$$\text{Re}(\epsilon'/\epsilon) = (3.3 \pm 1.1) \times 10^{-3}. \quad (11)$$

For the first time a 3-standard deviation effect had been observed. This result was published in 1988\textsuperscript{23} with the conservative title “First evidence for direct CP violation”.

The main sources of systematic uncertainties are related to the accidental losses, to the energy scale difference between charged and neutral modes and to the background estimation.

After this first successful run, the collaboration decided to improve the beam and the detector in order to reduce both the main systematic and the statistical uncertainties, in two new data-taking periods which took place in 1988 and 1989.

In practice, the use of a lower proton energy to produce the $K^0_S$ beam and the tuning of the production angle of both beams (Table 1), resulted in a much flatter ratio of $K^0_S/K^0_L$ energy spectra, as shown in Fig. 4(right). This change reduced the systematic uncertainty from the energy scale from 0.3\% to 0.12\%.

On the charged decay mode, the information provided by the new transition radiation detector allowed an independent estimation of the main component of the background, that is $K_0^0$ decays. With this additional handle, the systematic uncertainty on the subtracted background in the signal region was reduced to 0.1\% (see Table 4).
The installation in 1989 of the ZTDC system provided the time measurement of the calorimetric clusters. Using this information the effect of accidentals was found to be in nice agreement with the overlay method. Overall, the systematic uncertainty associated with accidental activity was reduced to 0.14%.

The evolution of the systematic uncertainties between the run periods is shown in Table 4.

The final double ratio obtained with the upgraded detectors and beams using the 1988 + 1989 data taking is $RR = 0.9878 \pm 0.0026$ (stat) $\pm 0.0030$ (syst). This value of $RR$ translates into $\text{Re}(\epsilon' / \epsilon) = (2.0 \pm 0.7) \times 10^{-3}$. Taking into account that some systematic uncertainties are common to the two data taking periods, the average is

$$\text{Re}(\epsilon' / \epsilon) = (2.30 \pm 0.65) \times 10^{-3}. \quad (12)$$

This final result, now 3.5-standard deviation away from zero, was published in November 1993.

A few months earlier, the E731 collaboration at Fermilab came out with a result of a similar precision, but much lower central value, namely $\epsilon' / \epsilon = (0.74 \pm 0.61) \times 10^{-3}$. This E731 result was both compatible with zero and with the NA31 result and thus rather inconclusive. Given the smallness of the observed effect, another round of experiments was justified.

### 3.3. Phase measurement

In 1987, the NA31 experiment devoted a special run to the measurement of the phases, $\phi^{00}$ and $\phi^{+-}$.

As early as 1965, it was shown by Bell and Steinberger using unitarity and CPT conservation, that the phase of $\epsilon$, $\phi_\epsilon$, should have the “natural value” given by

$$\phi_\epsilon = \tan^{-1}(2\Delta M / \Gamma_S). \quad (13)$$

When CPT is conserved, and neglecting the small contribution of $\text{Im}(\epsilon' / \epsilon)$, the phases of $\eta^{00}$ and of $\eta^{+-}$ should be both equal to $\phi_\epsilon$.\(^4\) In 1987, $\phi^{+-}$ was

\(^4\)With $\Delta M = M_L - M_S = 3.48 \times 10^{-12}$ MeV = $0.53 \times 10^{10}$ s$^{-1}$ and $\Gamma_S = 1.11 \times 10^{10}$ s$^{-1}$ one has $\Delta M / \Gamma_S = 0.477$ resulting to $\phi_\epsilon = 43.5^\circ$. 

---

**Table 4** Systematic uncertainties on the double ratio (in %) for the two NA31 data taking periods.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Background in $K^0_L \rightarrow \pi^0\pi^0$</td>
<td>0.2</td>
<td>0.13</td>
</tr>
<tr>
<td>Background in $K^0_L \rightarrow \pi^+\pi^-$</td>
<td>0.2</td>
<td>0.10</td>
</tr>
<tr>
<td>Energy scale</td>
<td>0.3</td>
<td>0.13</td>
</tr>
<tr>
<td>Accidental losses</td>
<td>0.2</td>
<td>0.14</td>
</tr>
<tr>
<td>Monte Carlo acceptance</td>
<td>0.1</td>
<td>0.10</td>
</tr>
<tr>
<td>Trigger and AKS efficiency</td>
<td>0.2</td>
<td>0.09</td>
</tr>
<tr>
<td>Total systematic uncertainty</td>
<td>0.5</td>
<td>0.3</td>
</tr>
</tbody>
</table>
already rather well measured, unlike $\phi^{00}$ which was not precisely known, leading to a difference $\phi^{00} - \phi^{+-}$ of $12 \pm 6\degree$.

In a $K^0$ beam, after about $12\tau_S$, the $\pi\pi$ rate from the $K_S^0$ component and the CP violating component of $K_L^0$ decays become comparable and can interfere. In the interference region, the two-pion decay rate as a function of time in the kaon rest frame can be written as

$$I(t) = S(p)[e^{-t/\tau_S} + |\eta|^2e^{-t/\tau_L} + 2D(p)|\eta|e^{-t/2(1/\tau_S+1/\tau_L)}\cos(\Delta M t - \phi)]$$ (14)

where $S(p)$ is the momentum spectrum of $(K^0 + \bar{K}^0)$ and $D(p) = (K^0 - \bar{K}^0)/(K^0 + \bar{K}^0)$ is the dilution factor.

In order to have an optimised acceptance in the region between $10\tau_S$ and $15\tau_S$, the most sensitive to the interference term, NA31 used an improved beam scheme with two target stations, “far” and “near”, which were installed at 48 m and 33.6 m respectively, upstream the $K_L^0$ cleaning collimator. This beam set-up, with two-stage collimation and improved shielding, allowed intensities up to $2 \times 10^{10}$ protons per pulse. The same target dimensions, proton energy, and incidence angle as for the 1986 period shown in Table 1 were used.

Data were taken with the same apparatus and trigger as for $\epsilon'/\epsilon$, but downscaling all events with early decay vertices ($<7\tau_S$), in order to maximise the number of events in the interference region. The data sample consists of about 140 million triggers written to tape. Some additional data with the XTGV placed in its most upstream position were also taken at regular intervals to determine the energy scale of the neutral sample. An error of 0.1% on the energy scale induces a phase difference of about $1\degree$. The measured kaon decay rates in the charged and neutral modes are shown in Fig. 5.

The value of the phase difference is obtained by a combined fit (charged and neutral modes) to the ratio of event rates between the two targets, as a function of lifetime counted from the mid-point between the two targets.

The result is published in 1990: $\phi^{00} - \phi^{+-} = (0.20 \pm 2.6 \text{ (stat)} \pm 1.3 \text{ (syst)})\degree = (0.20 \pm 2.9\%)$, where the dominant systematic uncertainty comes from the energy scale. Combined with the exceedingly small value of $\Delta M$ compared to $M_K$, the phase difference measurements provide the strongest test of CPT conservation, at the level of $\pm 4 \times 10^{-19}$.

CP violation associated to CPT conservation implies that $T$ is violated. Such a violation has been observed with a significance of 6 standard deviations in the CPLear experiment at CERN, by comparing the rates of tagged $K^0 \rightarrow \pi^- e^+\nu$ and $\bar{K}^0 \rightarrow \pi^+ e^-\bar{\nu}$ in $p\bar{p}$ annihilations. This observation implies that the transition amplitudes $K^0 \rightarrow \bar{K}^0$ and $K^0 \rightarrow K^0$ are not identical, demonstrating the violation of $T$ symmetry.

$^6$This result is now superseded by KTeV with $\phi^{00} - \phi^{+-} = (0.29 \pm 0.31)\degree$. 
4. The Second Generation: The NA48 Beams and Detectors

The NA48 experiment aimed at achieving a precision of a few $10^{-4}$ on $\epsilon' / \epsilon$. Ten new European institutes on top of the six from the NA31 community contributed to this new collaboration. Data for the $\epsilon' / \epsilon$ measurement have been recorded in three periods in 1997, 1998–1999 and in 2001. A special run took place in 2000, allowing auxiliary studies to improve the control of specific systematic effects.

The collaboration focussed on a method allowing an even better cancellation of systematic uncertainties than in NA31, by collecting at the same time and in the same decay volume all the four decay modes. To allow this, the decay volume had to be shortened to $3.5 \tau_S$ counted from the end of the $K^0_S$ defining collimator.

Acceptance differences between $K^0_L$ and $K^0_S$ decays are minimised by weighting each $K^0_L$ event by a function of its proper time, so that the $K^0_S$ and $K^0_L$ decay spectra become almost identical. Unlike NA31, a magnetic spectrometer is used for the measurement of charged decays with a better precision. A quasi-homogeneous liquid krypton calorimeter with high granularity provides excellent position, energy and time information for the neutral decays. A schematic drawing of the NA48 beam line and detectors is given in Fig. 6.
4.1. The NA48 beams

The $K^0_L$ beam is produced by a ten times more intense 450 GeV proton beam than for NA31, impinging on a beryllium target under an angle of 2.4 mrad. The protons...
Fig. 7. Left: Schematic view of the bent crystal. Right: Schematic view of the tagger detector.

for the $K^0_S$ beam are derived from the non-interacting protons in the $K^0_L$ target hitting a silicon mono-crystal which is cut to dimensions of $(60, 18, 1.5)$ mm$^3$ parallel to the (110) crystalline plane. The crystal is bent and mounted on a motorised goniometer that allows a precise positioning to select, deflect and guide along the crystalline planes the desired amount of $3 \times 10^7$ ppp (Fig. 7(left)). The channelling properties of a bent crystal produce a clean beam with well defined emittance, deflected into the desired direction by applying the equivalent of a $14.4$ T⋅m bending power in a short length of $6$ cm, and containing only a $2 \times 10^{-5}$ fraction of protons without the presence of a heavy collimator system. The transmitted proton beam traverses a tagging station which registers precisely the passage time of each proton, and is then transported through a series of deflecting magnets to the $K^0_S$ target, located $6$ m upstream from the beginning of the fiducial decay region and $72$ mm above the $K^0_L$ beam axis. At the aperture of the final collimator, the $K^0_S$ beam goes through an anti-counter used to precisely define the beginning of the $K^0_S$ fiducial region, like in NA31. The $K^0_S$ beam direction is tuned so as to overlap the $K^0_L$ one at the entrance of the main detectors, $120$ m downstream. Both $K^0_S$ and $K^0_L$ beams are transported in vacuum inside a large evacuated tank, closed off at its end by a thin Kevlar window to separate the evacuated decay region from the detectors.

4.2. The tagger

The tagging detector placed in the path of the proton beam selected by the bent crystal consists of 24 scintillators mounted in a carbon-fibre structure, alternating in horizontal and in vertical directions and staggered so as to share the beam profile among different counters. For the incoming protons, the scintillators offer in each direction a geometrical overlap of $50$ µm between successive detectors to reduce possible small misalignment issues (Fig. 7(right)). Despite the radiation hardness of the material the effect of $10^{13}$ protons striking the detector in 100 days caused
a 50% decrease of the emitted light. This was monitored and corrected for by the appropriate change of the photomultiplier high voltage. After digitisation by a 8-bit 1 GHz flash-ADC the data are transferred to a ring buffer and are extracted in case of a trigger. The tagger showed excellent performance, offering a reconstructed time resolution of 140 ps and a separation of two close pulses down to 5 ns.

4.3. The liquid Krypton calorimeter

The liquid Krypton calorimeter consists of \(\sim 13000\) cells of \(2\, \text{cm} \times 2\, \text{cm}\) cross-section, defined by Cu–Be–Co ribbons of \(40\, \mu\text{m} \times 18\, \text{mm} \times 125\, \text{cm}\) forming longitudinal projective towers pointing to the centre of the decay region and immersed in a bath of \(\sim 10\, \text{m}^3\) of liquid Krypton. Each cell is made out of two cathodes and one central anode. Liquid Krypton was adopted as active medium because its radiation length (4.7 cm) is short enough to allow building a compact homogeneous calorimeter while being still affordable. The intrinsic stability of the ionisation signal and the small Moliere radius, ensuring compact lateral shower size, are further assets. A high voltage of 3 kV was applied to generate the drift field. Preamplifiers and calibration system are directly installed at the downstream end of the anodes in the cold volume, minimising the noise and allowing fast charge transfer. The ionisation signal is shaped outside the cryostat and digitised asynchronously by a 10 bit 40 MHz flash ADC. A four gain-switching scheme driven by the signal pulse-height, covers the total dynamic range from 3 MeV to 50 GeV. The readout is restricted to active cells grouped into clusters, by applying zero-suppression using dedicated algorithms.

Energy and time are reconstructed using a digital filter method\(^{34}\) applied to the three samples around the maximum. Clusters of \(3 \times 3\) cells are formed around the most energetic cell. The barycentre of the energy deposition is used as estimator of the cluster position and the time of the most energetic cell is considered as time of the cluster. After cell inter-calibration using Ke3 decays, \(E/p\) studies result in an energy resolution of \(3.2\%/\sqrt{E}\) (GeV), with additional contributions of 90 MeV from the electronic noise and 0.42% from the constant term. The energy response is measured to be linear within 0.1% between 5 and 100 GeV. The position resolution is better than 1 mm in both directions and the time resolution is 500 ps.

4.4. The spectrometer

Charged decays are triggered by a scintillator hodoscope. The tracks are reconstructed from the hits measured in the four octagon-shaped wire chambers (DCH), two located before and two after the dipole magnet. Each chamber contains eight planes of grounded sense wires tilted by 45° with respect to each other that guarantees a high detection efficiency with sufficient redundancy.\(^{35}\) All anode wires are instrumented with amplifiers and TDC circuits. The magnet integral field of 0.883 Tm between the second and the third chamber induces a transverse momentum kick of 265 MeV/c in the horizontal direction. The track
momentum is reconstructed with a resolution measured in electron beams to be
\[ \sigma(p)/p = 0.48\% + 0.009 \times p\% \ (p \ \text{in GeV}/c), \]
where the first term is due to multiple scattering in the DCH and the surrounding Helium gas, and the second one comes from the position accuracy of the chamber hits. The resolution of the reconstructed kaon mass is 2.5 MeV/$c^2$. For two-track events, the decay vertex is computed from hits in the two upstream chambers, with a longitudinal (transverse) resolution of 50 cm (2 mm) and a spread of 7 mm on the closest distance of approach.

After the 1999 run, an implosion of the beam tube caused severe damages to the drift chambers. New chambers were installed for the 2001 data taking.

4.5. The NA48 trigger and data acquisition systems

The goal of the NA48 trigger is to reduce the 500 kHz rate of particles hitting the detector to a few kHz of accepted events, with minimal inefficiencies and dead time. The first level of the pipelined charged trigger is based on signals in the hodoscope, hit multiplicity in the wire chambers and an energy threshold in the electromagnetic calorimeter. The second level computes the track coordinates, the proper decay time and the invariant mass using positions of the hits in the first, second and fourth chamber.

A pipelined design was also used for the neutral trigger. Using the analogue sums made on $2 \times 8$ calorimetric cells in both, horizontal and vertical projections, the kaon energy, the number of photons, their arrival time and the proper decay time are estimated every 25 ns. The trigger requires less than six peaks in each projection to reject $K^0_L \rightarrow 3\pi^0$ decays, a total energy greater than 50 GeV and a decay vertex reconstructed at less than 5\tau_S from the final collimator.

The trigger decisions from the sub-detectors are assembled by the trigger supervisor system that defines the event timestamp relative to the 40 MHz clock used to synchronise all the detectors. Event fragments are built up to complete events in on-site PC farms and are transmitted to the CERN computing centre via a Gigabit optical link with a speed of about 10 MB/s. There, raw data are stored on disk and in parallel monitored, reconstructed and further selected in an offline PC farm.36

4.6. The NA48 analysis

This section provides a description of the analysis of the statistically most precise data period, 1998–1999. Final results obtained combining all the periods are given at the end.

4.6.1. The neutral decays

Clusters are chosen in the energy range from 3 to 100 GeV and must be well separated from each other. The 4-cluster time is computed as an energy-weighted average using the two most energetic cells in each cluster. Each cluster time must
be compatible within 5 ns with the average time. The longitudinal vertex position is computed from the energies and the positions of the four clusters assuming a kaon decay. The \( m_{\gamma\gamma} \) values obtained from the 2-by-2 photon associations are combined in a \( \chi^2 \) discriminator. Out of the three possible pairings, the one with the lowest \( \chi^2 \) is kept. The main background in the \( 2\pi^0 \) decay channel results from unidentified \( K^0_L \rightarrow 3\pi^0 \) with undetected or merged photons, resulting in only four measured clusters. The background is studied in a control region defined as \( 36 < \chi^2 < 135 \) and subtracted for each energy bin from the signal region, defined as \( \chi^2 < 13.5 \) (see Fig. 8(left)). The effect of the subtracted background on the double ratio is \((-5.9 \pm 2.0) \times 10^{-4}\). The large reduction with respect to NA31 is due to both the better \( m_{\gamma\gamma} \) resolution and the lower overall background acceptance within the significantly shorter decay volume.

The energy scale is adjusted comparing the upstream edge of the reconstructed vertex distribution of \( K^0_S \rightarrow 2\pi^0 \) candidates with the AKS counter position. Additional checks are done using \( \pi^0 \) and \( \eta \) decays into two photons, produced in special runs where a \( \pi^- \) beam hits a target placed in different known positions. The agreement of the reconstructed and the nominal target positions was found to be within 1 cm. Comparing the energy scales obtained with two or more distant targets gives constraints on the energy non-linearity.

4.6.2. *The charged decays*

Tracks are reconstructed using the hits and the drift time information. These tracks must come from a common vertex, have momenta greater than 10 GeV and be at least 12 cm away from the centre of each DCH. The tracks must be within the
acceptance of the LKr and of the muon vetoes to allow proper particle identification. The reconstructed kaon energy is computed from the momentum ratio of the two tracks, and their relative opening angle $\theta$, assuming a $K \to \pi^+\pi^-$ decay (see Eq. (10)). The vertex reconstruction is controlled comparing the nominal position of the AKS counter to the upstream edge of the reconstructed spectrum of $K^0_S \to \pi^+\pi^-$ decays.

Three-body decays are the main background in the $K^0_L$ sample. $K\pi\pi$ decays are rejected by an $E/p$ test applied to both tracks. Removing events with in-time hit in the muon vetoes suppresses the $K\mu\pi$ background. Further suppression of both modes is obtained by demanding the two-tracks invariant mass to be compatible with the kaon mass. The final discrimination is done using the transverse missing momentum $P_T^2$ constructed in a way to be independent of the vertex resolution. Figure 8 right shows the $P_T^2$ distribution. The signal region is defined as $P_T^2 < 0.0002$ (GeV/c)$^2$. The tails are populated with remaining $K\pi\pi$ events, as verified with $K\pi\pi$ enriched samples, respectively. The background in the signal region corresponds to a residual fraction of $10 \times 10^{-5}$ $K\pi\pi$ and $6 \times 10^{-4}$ $K\mu\pi$. It is evaluated in each of the twenty energy bins. The $K^0_S$ sample is also contaminated by $\Lambda \to \pi^-p$ decays which are suppressed by exploiting the characteristic asymmetric momenta of the two tracks.

All requirements against the backgrounds are applied to both $K^0_S$ and $K^0_L$ samples in order to symmetrise the losses affecting genuine $2\pi$ decays.

4.6.3. Corrections: Tagging inefficiency and dilution

The separation of the $K^0_S$ and $K^0_L$ data samples is based on the time coincidence of the tagger with the LKr (for neutral decays) and the charged hodoscope (for charged decays).

As the tagger signal is used for both charged and neutral decays the uncertainties related to its time measurement are symmetric. Two mistagging effects remain:

- The first type of mistagging occurs when a $K^0_S$ event is labelled as $K^0_L$ because the reconstructed detector time falls outside the coincidence window. This inefficiency can be measured independently in the charged mode, by selecting $K^0_S$ decays from the reconstructed vertical position of the decay vertex. The inefficiency, $\alpha_{SL}$, is the fraction of events lying outside the coincidence window of $\pm 2$ ns (Fig. 9(left)). To measure the inefficiency in the neutral mode, the neutral decays ($K^0_S$ and $K^0_L$ into $2\pi^0$ and $3\pi^0$) with a photon converted into an electron–positron pair are considered. On an event-by-event basis, the charged time computed from the electron and positron hits in the hodoscope, is compared with the neutral time, reconstructed as the average time of the photons.

Overall, the mistagging probability $\alpha_{SL}$ has been found to be the same for charged and neutral decays within an uncertainty of $\pm 0.5 \times 10^{-4}$.

- The second type, denoted $\alpha_{LS}$, concerns $K^0_L$ events assigned to the $K^0_S$ sample because of an accidental coincidence with the tagger. This effect, called “dilution”,...
The Discovery of Direct CP Violation

is assessed by measuring the probability for a $K^0_L$ event to be accompanied by a tagged proton within a ±2 ns window located in the sidebands before or after the event time. This probability is as large as 10%. An additional correction is applied for the small intensity difference between in-time and sideband windows, estimated using independently identified $K^0_L$ samples: charged decays chosen by their vertex to be $K^0_L$ and 3π$^0$ neutral events. The measured value is $\Delta \alpha_{LS} = \alpha^{00}_{LS} - \alpha^{+ -}_{LS} = (4.3 \pm 1.4\text{ (stat)} \pm 1.0\text{ (syst)}) \times 10^{-4}$. The origin of the difference between charged and neutral dilution has been identified to be due to the higher sensitivity of the charged mode to accidentals, at both the trigger and reconstruction levels. This value has been confirmed by the overlay method and varies within the SPS spill as the instantaneous intensity (Fig. 9(right)).

4.6.4. Corrections: Beam activity, scattering and acceptance

Most of the accidental activity originates from the high-intensity $K^0_L$ beam. The concurrent collection of the four modes minimises the sensitivity of the result to this effect. Overlaying the data with specific triggers recorded by the beam monitors in proportion to the intensity, allows to estimate the residual effect. The net (losses-gains) measured effect is larger in the charged mode with respect to the neutral one by (1.4 ± 0.7)% (Fig. 10(left)). Because the intensity variations turn out to be similar in both beams within ±1%, this effect cancels in the double ratio. The overall uncertainty from accidental effects on the double ratio is $\Delta RR = \pm 4.2 \times 10^{-4}$.

Scattered kaons appear as tails in the distribution of the radius of the centre-of-gravity for neutral and charged events. For $K^0_L$ these tails are dominated by events
which have suffered double scattering, resulting in a correction of $-(9.6 \pm 2.0) \times 10^{-4}$ to account for the difference between the losses in the two decay modes.

Despite the convergence of the two beams towards the LKr centre and the use of a short common decay region for the measurement, the different lifetimes of $K^0_S$ and $K^0_L$ would imply a large acceptance correction on the double ratio, as shown in Fig. 10(right). To cancel the bulk of this correction, a weighting factor is applied to each $K^0_L$ event as a function of its proper time resulting into very similar decay spectra for the two beams. This reduces the Monte Carlo correction to $(26.7 \pm 4.1({\rm stat}) \pm 4.0({\rm syst})) \times 10^{-4}$ at the cost of an increased statistical uncertainty.

### 4.7. NA48 results

The result is computed in 20 bins of kaon energy in the range 70–170 GeV. In each bin the double ratio is estimated using the numbers of $K^0_S$ and proper-time weighted $K^0_L$ charged and neutral decays and applying the various corrections estimated per bin. Table 5 shows the effect of the various corrections on the raw double ratio, together with their uncertainties for the 1998 + 1999 result.\(^{37}\)

The final double ratio is extracted by averaging the 20 numbers with a logarithmic estimator.

The overall NA48 statistics and results for the three run periods are given in Table 6. The combined final result of the experiment is $(14.7 \pm 2.2) \times 10^{-4},^{38}$ confirming the observation of direct CP violation with a significance of 6.6 standard deviations.
Table 5  Corrections and systematic uncertainties on the double ratio for the 1998 + 1999 data sample.

<table>
<thead>
<tr>
<th>Source</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi^+\pi^-$ trigger inefficiency</td>
<td>$-3.6 \pm 5.2$</td>
</tr>
<tr>
<td>AKS inefficiency</td>
<td>$+1.1 \pm 0.4$</td>
</tr>
<tr>
<td>Reconstruction $\pi^0\pi^0$</td>
<td>$0 \pm 5.8$</td>
</tr>
<tr>
<td>Reconstruction $\pi^+\pi^-$</td>
<td>$+2.0 \pm 2.8$</td>
</tr>
<tr>
<td>Background $\pi^0\pi^0$</td>
<td>$-5.9 \pm 2.0$</td>
</tr>
<tr>
<td>Background $\pi^+\pi^-$</td>
<td>$+16.9 \pm 3.0$</td>
</tr>
<tr>
<td>Beam scattering</td>
<td>$-9.6 \pm 2.0$</td>
</tr>
<tr>
<td>Accidental tagging</td>
<td>$+8.3 \pm 3.4$</td>
</tr>
<tr>
<td>Tagging inefficiency</td>
<td>$0 \pm 3.0$</td>
</tr>
<tr>
<td>Acceptance</td>
<td>$+26.7 \pm 4.1 \pm 4.0$</td>
</tr>
<tr>
<td>Accidental activity</td>
<td>$0 \pm 4.4$</td>
</tr>
<tr>
<td>Long term variations of $K^0_S/K^0_L$</td>
<td>$0 \pm 0.6$</td>
</tr>
<tr>
<td>Total</td>
<td>$35.9 \pm 12.6$</td>
</tr>
</tbody>
</table>

Table 6  Number of selected events after background subtraction and corrected for mistagging for the three data taking periods. The $K^0_L$ statistics are given without lifetime weighting. The corresponding results on $\epsilon'/\epsilon$ are also given together with their uncertainties.

<table>
<thead>
<tr>
<th>Source</th>
<th>1997</th>
<th>1998 + 1999</th>
<th>2001</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb of $K_L \rightarrow \pi^0\pi^0$ (x1000)</td>
<td>489</td>
<td>3290</td>
<td>1546</td>
<td>5325</td>
</tr>
<tr>
<td>Nb of $K_S \rightarrow \pi^0\pi^0$ (x1000)</td>
<td>975</td>
<td>5209</td>
<td>2159</td>
<td>8343</td>
</tr>
<tr>
<td>Nb of $K_L \rightarrow \pi^+\pi^-$ (x1000)</td>
<td>1071</td>
<td>14453</td>
<td>7136</td>
<td>22660</td>
</tr>
<tr>
<td>Nb of $K_S \rightarrow \pi^+\pi^-$ (x1000)</td>
<td>2087</td>
<td>22221</td>
<td>9605</td>
<td>33913</td>
</tr>
<tr>
<td>$\epsilon'/\epsilon$ (x10^{-4})</td>
<td>18.5</td>
<td>15.0</td>
<td>13.7</td>
<td>14.7</td>
</tr>
<tr>
<td>Stat error (x10^{-4})</td>
<td>4.5</td>
<td>1.7</td>
<td>2.5</td>
<td>1.4</td>
</tr>
<tr>
<td>Syst error (x10^{-4})</td>
<td>5.8</td>
<td>2.1</td>
<td>1.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Total error (x10^{-4})</td>
<td>7.3</td>
<td>2.7</td>
<td>3.1</td>
<td>2.2</td>
</tr>
</tbody>
</table>

The KTeV experiment at Fermilab published its final result in 2011 giving a combined value for $\epsilon'/\epsilon$ of $(19.2 \pm 2.1) \times 10^{-4}$,\[40\] in fair agreement with both NA31 and NA48 results.

5. Concluding Remarks

5.1. The world average of $\epsilon'/\epsilon$

The comparison of the experimental results from CERN and Fermilab is shown in Fig. 11. The world average computed by the Particle Data Group is $(16.8 \pm 2.0) \times 10^{-4}$, where the quoted error is inflated by 1.4 to take into account the dispersion of results. The existence of a direct component in CP Violating amplitudes in the kaon system is experimentally established at a level of 8.4 standard deviations.\[40\]
Using formulas (3)–(8), this fundamental result can be expressed as:

\[
\frac{\Gamma(K^0 \to \pi^+\pi^-) - \Gamma(K^0 \to \pi^+\pi^-)}{\Gamma(K^0 \to \pi^+\pi^-) + \Gamma(K^0 \to \pi^+\pi^-)} = 2\text{Re } \epsilon' = (5.3 \pm 0.6) \times 10^{-6},
\]

\[
\frac{\Gamma(K^0 \to \pi^0\pi^0) - \Gamma(K^0 \to \pi^0\pi^0)}{\Gamma(K^0 \to \pi^0\pi^0) + \Gamma(K^0 \to \pi^0\pi^0)} = -4\text{Re } \epsilon' = (-10.6 \pm 1.2) \times 10^{-6}
\]

demonstrating the occurrence of direct CP violation.

5.2. **CP violation in kaons: A portal to heavy meson systems**

Since its unexpected discovery in 1964, CP violation opened new horizons in experiments and it has been searched for in heavier meson systems. $D^0\bar{D}^0$ and $B^0\bar{B}^0$ offer a rich phenomenology, which has been exploited in the last 20 years by dedicated experiments at $e^+e^-$ colliders (CESR, LEP, PEP-2, KEK-B) and at $p\bar{p}$ colliders (Tevatron, LHC). Given the high mass of these mesons, many final states are allowed so that the decay widths of the “heavy” and “light” mass eigenstates are much closer than in the kaon system.

In the D system, oscillations due to mixing have been observed and measured recently,\(^{41}\) with $x = (0.41 \pm 0.14)\%$ and $y = (0.63 \pm 0.07)\%$, where $x = \frac{\Delta M}{2\Gamma}$ and
The Discovery of Direct CP Violation

\[ y = \frac{\Delta \Gamma}{\Gamma} \]

with \( \Delta M \) and \( \Delta \Gamma \) being the mass and decay width difference of the two masse eigenstates and \( \bar{\Gamma} \) the average decay width. No CP violation was found, in agreement with SM expectations.

The \( B^0 \bar{B}^0 \) phenomenology is much richer. For the \( B^0_d \), \( x = 0.774 \pm 0.006 \) while \( y \), expected very small in the SM, is experimentally compatible with zero. The first observation of CP violation was made by Babar and Belle\(^{42, 43} \) in the \( \psi K_S \) final state where it is due to interference between mixing and decay amplitudes. Direct CP violation has been observed in several final states. For example, the measured asymmetry between \( B^0 \rightarrow K^+ \pi^- \) and \( \bar{B}^0 \rightarrow K^- \pi^+ \) was first observed by the LHCb experiment\(^{44} \) with a value of \(-0.082 \pm 0.013\), widely different from the direct CP violation in the kaon system.

For \( B_s \) it took quite a long time before oscillations were observed, because \( \Delta M_s \) is large compared to \( \Gamma_s \) resulting into oscillations with a very short wave length, beyond the capabilities of early vertex detectors. The first observation was made by CDF.\(^{45} \) Today this is well established with \( x_s = 26.85 \pm 0.13 \) and \( y_s = 0.137 \pm 0.012 \).

CP violation in the \( B_s^0 \bar{B}_s^0 \) system has been established as well.\(^{46} \)

On the rare decay side, a step forward was the observation of the \( B_s \rightarrow \mu \mu \) mode which has recently been measured with a rate of \((2.9 \pm 0.7) \times 10^{-9}\),\(^{47} \) compatible with the accurate SM prediction.\(^{48} \) With additional data this decay will become sensitive to new physics.

5.3. CP violation in kaons: A portal for theory

The observation of CP violation fueled spectacular ideas in theory. The CKM formalism and the Standard Model were largely inspired by this discovery. The non-conservation of this symmetry is now a well established phenomenon in the weak decays of hadrons, parametrised by the CKM matrix elements. CP violation was demonstrated by Sakharov\(^{49} \) in 1967 to be one of the three necessary conditions for baryogenesis. While it is tempting to associate the observed CP violation to the corresponding Sakharov’s condition, it is commonly admitted that its magnitude is far too small to play this role. But sources of CP violation other than the CKM matrix are possible. In Higgs-boson mediated transitions with multiple Higgs bosons, CP violation can appear; the value of \( \epsilon \) in the kaon system imposes however strong constraints on this possibility.\(^{50} \) Finally, CP violation is also possible and currently being searched for in the neutrino sector with particular emphasis now that \( \sin^2 2\theta_{13} \) is known with good precision\(^{51} \) and has a sizeable value (\( \approx 0.09 \)).

5.4. The legacy of CERN kaon experiments

CERN experiments have had a leading role in the discovery of the direct CP violation component, although the competition with Fermilab was essential. Both NA31 and NA48 groups have improved continuously the beams, the detectors and the analysis methods to better control the systematic uncertainties. In their hunt
for precise measurements, the groups developed original tools and methods that are nowadays widely used: the overlay procedure to account for accidental effects, the multiple sampling of calorimetric signals for fast and precise readout, the use of online and offline computer farms, the remote data control, the use of crystal channelling for beam selection and transport, etc.

In parallel with the $\epsilon'/\epsilon$ data takings, NA31 and NA48 had both undertaken a wide program of studies of rare kaon decays, with original measurements and first observations of several channels.$^{52-54}$

The important contribution of these experiments to the worldwide physics landscape has been awarded by prestigious prizes, given — in particular — to the NA31 collaboration for the first observation of the direct CP component, published in the 1988 paper.$^{23}$

Today, the theoretical precision of $\epsilon'/\epsilon$ is still unsatisfactory and cannot be usefully confronted with the measurements. Hopes for more accurate predictions are based on ongoing lattice QCD computations.$^{55}$

In the meanwhile, other kaon decay channels are used to push further the limits of the Standard Model. The NA48/2 Collaboration has accumulated in 2003–2004 about $\simeq 4 \times 10^9 K^{\pm} \rightarrow \pi^{\pm} \pi^+ \pi^-$ and $\simeq 10^8 K^{\pm} \rightarrow \pi^{\pm} \pi^0 \pi^0$ decays in order to study the Dalitz plot of the final states, where direct CP violation could induce an asymmetry. The null result obtained by this experiment is consistent with Standard Model expectations.$^{56}$

Even more interesting would be the study of the very rare decay $K^0 \rightarrow \pi^0 \nu \bar{\nu}$, not yet observed, for which the calculated branching ratio $(2.4 \pm 0.1) \times 10^{-11}$, together with the elusive final state raise extreme experimental challenges. The precision of the SM prediction is however a strong motivation for a dedicated experiment. Several proposals were made in the past; the only one still alive is the E14 proposal at JPARC. Also very clean theoretically, but not dominated by a CP-violating transition, is the $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ mode with an expected rate of about $10^{-10}$. Also experimentally challenging, this channel allows a precise measurement of the $|V_{td}|$ CKM parameter. A dedicated experiment, NA62,$^{57}$ installed in the EHN2 area at the CERN SPS, had its first run in 2014 and expects to collect in two years about 80 $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ decays. Results from these two very rare decay channels would allow an independent and accurate determination of the Unitarity triangle,$^{58}$ probing at the same time physics scenario beyond the Standard Model.

Acknowledgments

The authors are grateful to Ivan Mikulec and Arthur Schaffer for the careful reading and their suggestions. The technical expertise of Catherine Bourge and Bruno Mazoyer contributed efficiently to the preparation of this document.
The Discovery of Direct CP Violation

References

14. M. Holder et al., On the decay $K_L \rightarrow \pi^0\pi^0$, Phys. Rev. B 40, 141 (1972).
29. G. D. Barr et al., Proposal for a precision Measurement of $\epsilon'/\epsilon$ in CP Violating $K^0 \to 2\pi$ decays, CERN/SPSC/90-22 (1990).

47. V. Khachatryan et al., Observation of the rare $B^0_\tau \rightarrow \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data, submitted to *Nature*, arXiv:1411.4413 (2014).


55. T. Blum et al., The $K \rightarrow (\pi\pi)_{\tau=2}$ decay amplitude from Lattice QCD, arXiv: 1111.1699 (2011).

56. R. Batley et al., Search for direct CP violating charge asymmetries in the $K^\pm \rightarrow \pi^+\pi^0\pi^-$ and $K^\pm \rightarrow \pi^0\pi^+\pi^0$ decays, *Eur. Phys. J. C* **52**, 875–891 (2007).
