Supplementary information

1. Sum rules in x-ray magnetic circular dichroism (XMCD)
XMCD is a powerful tool for obtaining spin and orbital contributions to the magnetization in ferromagnets. Sum rules relate the difference in x-ray absorption between left and right circularly polarized x-rays at, for instance, the L\textsubscript{2,3} absorption edges of 3d transition metals to ground state expectation values of orbital, L, and spin, S, angular momentum operators (in units of η) along the sample magnetization direction as\textsuperscript{1,2}

\[
L = -\frac{4}{3} \frac{\Delta A_1 + \Delta A_2}{A_3 + A_2} n_h \tag{A1}
\]

\[
S + \frac{7}{2} T = -\frac{\Delta A_1 - 2\Delta A_2}{A_3 + A_2} n_h \tag{A2}
\]

Here \(A_{2,3}\) and \(\Delta A_{2,3}\) are the integrated intensities over the corresponding L\textsubscript{2,3} absorption edges for the sum and difference spectra taken with opposite x-ray helicity, respectively. \(n_h\) is the number of empty states (holes) in the 3d shell. The so-called magnetic dipole operator, \(T\), describes the anisotropy of spin and charge distribution of 3d electrons in the unit cell. Stöhr and König\textsuperscript{3} and independently Dürr and van der Laan\textsuperscript{4} showed that this term averages to zero in angle integrated XMCD measurements or for polycrystalline samples (as used in the main paper).

The sum rules in eqs. (A1) and (A2) show that XMCD measurements at two spin-orbit split absorption edges, L\textsubscript{2} and L\textsubscript{3}, can be used to separate \(L\) and \(S\). If only one absorption edge is used, such as the L\textsubscript{3} edge in the main paper, a linear combination of the two quantities is obtained as

\[
\Delta A_i = -\frac{A_3 + A_2}{n_h} \left(\frac{1}{2} S + \frac{1}{2} L\right) \tag{A3}
\]

or as \(\Delta A_i \propto (S + \frac{3}{2} L)\) as used in the main paper.

The time resolved experiments in the main paper use eq. (A3) to follow the evolution of \((S + \frac{3}{2} L)\) as a function of pump-probe time delay. This is exact if the proportionality factor
\( \frac{1}{2}(A_2 + A_3)/n_h \) remains unaffected by the laser excitation. In an earlier publication (ref [10] of the main paper) we showed that less than 0.1 Ni 3d electrons per atom are repopulated by the fs laser pulse. The resulting relative change in the number of 3d holes, \( n_h \), is <10% (typically there are about 9 3d electrons in the Ni ground state). The evolution of \( (S + \frac{1}{2}L) \) in the main paper assumes that \( \frac{1}{2}(A_2 + A_3)/n_h \) remains unchanged by laser excitation. This translates into a ~10% confidence limit via possible changes in \( n_h \) which is below the statistical uncertainty of the data shown in Fig. 3B of the main paper.

We conclude this section with a comment on the validity of sum rule analysis of static and dynamic measurements. For static measurements, i.e. those taken in thermal equilibrium, the spin and orbital magnetisation derived from sum rule analysis was found to agree very well with theory and other experimental methods \(^5\). For 3d transition metals the orbital magnetization is typically only 10% of the spin moment \(^1,^5\) (this translates into a 20% ratio for the corresponding angular momenta as used in the main paper). Sum rules were originally derived by elegant methods of angular momentum theory \(^1,^2\). However, subsequent simpler derivations, e.g. of the orbital sum rule \(^6\) are more transparent for grasping their physical content as well as the underlying assumptions. The characteristic feature of sum rules is that only magnetic properties of an initial state, \(|i>\), are determined by spectroscopically probing x-ray transitions into excited final states, \(|f>\). This is achieved by integration of the measured x-ray spectra over all possible final states, \(|f>\), (i.e. integration over the whole absorption edge) \(^6\). This feature and the underlying assumptions remain unaffected by laser excitation since the sum rule integration is performed also over the laser excited final states. Laser induced changes of electronic level population, state hybridization, exchange interaction etc. are, therefore, all included in the sum rule analysis.
2. Modelling of x-ray absorption spectroscopy (XAS) line shape changes

In this section we describe the modelling of the observed XAS line shape changes (Fig. 2 of the main paper) following ultrafast laser excitation of the Ni 3d electrons. We discuss the influence of single- and many-particle excitations on the L3-edge absorption spectra of Ni metal using different theoretical approaches.

2.1 Many-particle transitions in x-ray absorption spectra

Electron-electron interactions in the Ni 3d shell for the XAS initial and final states lead to a multiplet structure discernible as satellite peaks at the L2,3 absorption edges. The multiplet structure can be reproduced by a cluster model (see Fig. A1), based on the Anderson impurity description of electron correlation effects\(^7,8\). The Ni initial state without laser excitation is given by a coherent superposition of n-electron wave functions, 3d\(^n\), containing 35% 3d\(^{10}\), 51% 3d\(^9\) and 14% 3d\(^8\) valence electronic configurations\(^8\). The smaller 3d\(^{10}\) and 3d\(^8\) admixtures to the dominant 3d\(^9\) configuration describe hopping induced valence electron fluctuations between the Ni 3d shell of atoms probed by the x-ray absorption events and 4sp orbitals of neighbour atoms\(^8\). Compared to that, the 3d - 3d orbital overlap between neighbour Ni atoms is much smaller and was not taken into account in the model. Only the initial state 3d\(^9\) and 3d\(^8\) configurations can contribute to the L2,3 XAS process, 3d\(^n\) → 2p\(^5\)3d\(^{n+1}\). Following ref. [8] this results in the main XAS peak as being mainly of 3d\(^9\) initial state character and in a satellite feature of mainly 3d\(^8\) weight (see Fig. A1). At 0K
only the lowest energy multiplet initial state is occupied. (blue solid line at the bottom of Fig. A1) With increasing temperature and also with laser heating excited states become populated. However, we found in our calculation that this does not affect the Ni XAS line shape, even for electronic temperatures of 1000 K a value observed in Ni at 200 fs (see ref [14] of the main paper). We can reproduce the experimentally observed ~130 meV XAS peak shift (Fig. 2A of the main paper) when we reduce the wave function overlap between Ni 3$d$ and ligand 4$s$p orbitals by 9%. This corresponds to a reduced electronic hopping between these orbitals, i.e. to increased electron localization. Changing the Ni 3$d$ - ligand 4$s$p charge transfer energy by a similar relative amount than the electronic hopping matrix element does not result in a shifted XAS main peak.

2.2 One electron processes in x-ray absorption spectra

In the above cluster model emphasis is put on a treatment of electron correlations, however, the valence band structure is not explicitly taken into account. The effect of dispersing valence bands and their temperature dependent electron population on XAS can be estimated by considering single electron transitions. Here we follow the approach of Erskine and Stern due to its transparency. Using the 3$d$ density of states of paramagnetic Ni the XAS near threshold was calculated for unpolarized x-rays as shown in Fig. A2 for two temperatures of 300K (blue line) and 1000K (red line). This method is strictly valid only for non-interacting valence electrons. However, this is a reasonable approximation for Ni XAS near threshold which is dominated by single particle like transition from a 3$d^9$ initial state configuration, characterized by a single hole in the 3$d$ shell, into a completely filled 3$d^{10}$ shell final state configuration (see also Fig. A1).
The XAS temperature dependence in this model is dominated by a broadening of the leading edge clearly visible in Fig. A2. There is also a slight reduction of the XAS peak intensity. However, no energy shift of the XAS features is discernible.

Together with the results shown in Fig. A1, we can conclude that the observed XAS peak shift is a consequence only of changing electron correlation effects.

3. Time-resolved data analysis

In order to describe the measured time-dependent absorption, $A$, in Fig. 3 of the main paper, we assume a response function consisting of two exponential curves, convolved with the probe pulse duration, approximating the rate equations of the three-temperature model\textsuperscript{10, 11} as

$$A = H(t) \cdot \left\{ K_1 \left[1 - \exp(-t/\tau_1)\right] \cdot \exp(-t/\tau_2) + K_2 \left[1 - \exp(-t/\tau_2)\right]\right\}$$

(A4)

Here $\tau_1$ is the thermalization time describing the rising edge of the curve, and $\tau_2 > \tau_1$ is the decay time that describes the transfer of thermal energy to other sub-systems (spin and lattice). $H(t)$ is the Heaviside function ($H(t)=1$ if $t>0$, 0 otherwise) describing the energy transfer from the fs laser pulse, the duration of which is $<$50 fs. The laser pulse length is much shorter than the x-ray pulse duration of $100\pm20$ fs (see ref. [19] of the main paper).

We, therefore, convoluted eq. (A4) with a Gaussian of $l=100\pm20$ fs length (FWHM) which
describes the temporal resolution of our setup. This value consists mainly of the x-ray pulse
length, but also includes any jitter of the time delay during the measurements. The point of
zero time delay \((t=0)\) in Fig. 3 of the main paper is determined from this fit. The parameter
\(K_2 \ll K_1\) accounts for a small residual temperature rise of the whole system and was
neglected in our fit shown in the main paper as it did not have any influence determining the
time constants.

With these parameters the fit of the data in Fig. 3A of the main paper results in a rise time \(\tau_l\)
\(= 120\pm 50\) fs. This value is very similar to the electronic thermalization timescales of 54 fs
for CoPt\(^3\)\(^{10}\) and 150 fs for Ni films\(^{11}\). It is also consistent with measurements of the
evolution of the laser induced electron-hole excitations obtained by photoemission
measurements (see ref. [10] of the main paper).

The measured initial response of the XMCD signal, \(D\), is characterized by a monotonous
drop, which is approximated by an exponential decrease convolved with the probe pulse:

\[
D = 1 - H(t) \cdot \left[1 - \exp\left(-t/\tau_D\right)\right]
\] (A5)

Within our time delay range we do not observe a recovery of the dichroic signal, \(D\). Hence,
we use only one time constant \(\tau_D\), describing the ultrafast demagnetization. Any energy
transfer to the lattice on the ps time range as included in the three temperature model was
found not to influencing \(\tau_D\) and consequently was neglected for simplicity. The values for
time-zero and temporal resolution obtained by fitting \(A\) were used also when fitting \(D\). The
resulting time constant \(\tau_D = 120\pm 70\) fs is identical with \(\tau_l\) within the experimental
uncertainty.

References:

1. Thole, B.T., Carra, P., Sette, F. & van der Laan, G., X-ray circular dichroism as a


