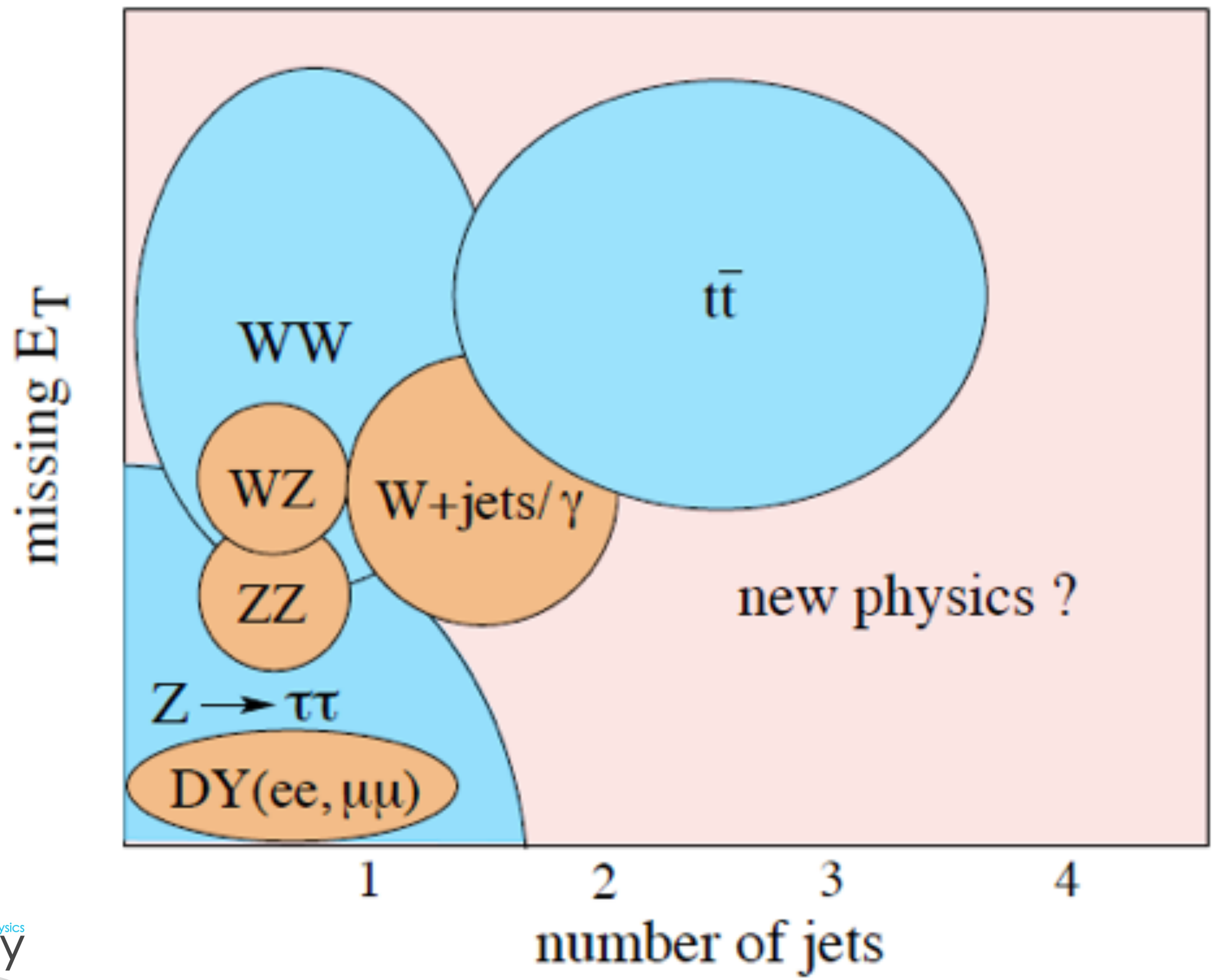
Resonance searches in multilepton final states in Atlas

Ask E. Løvchall-Jensen

Motivation for performing general searches

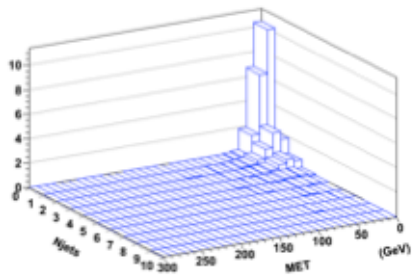
- The advantages of specific searches are numerous.
- Various approaches for more general searches however exist and can be interesting for various reasons.
- Resonance searches can be performed on large parameter sets using “bump hunters”
- Deviances can be seen in low statistics regions by comparing shapes of various processes simultaneously.

An example of an shape comparison. (CERN-THESIS-2011-004)

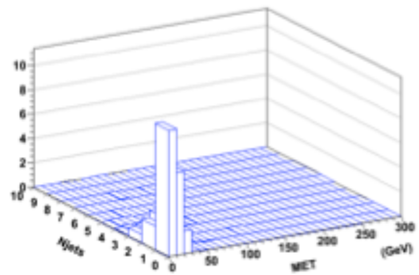


Monte Carlo generated pseudo data

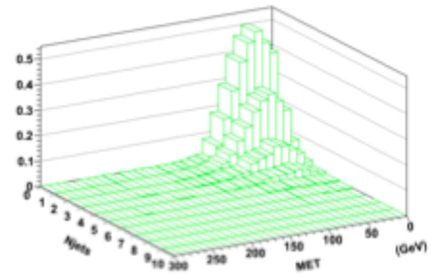
“Medium isolated” leptons



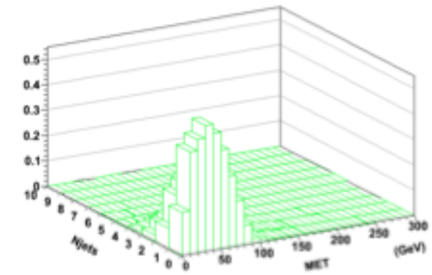
(a) Alpgen $Z\tau\tau \rightarrow e\mu \leq 5$partons



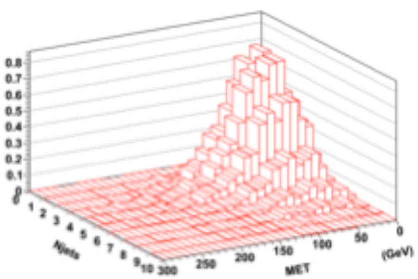
(b) Alpgen $Z\tau\tau \rightarrow e\mu \leq 5$partons



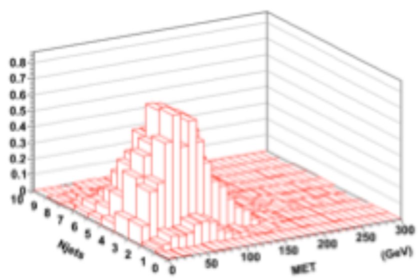
(e) Herwig $WW \rightarrow e\mu$



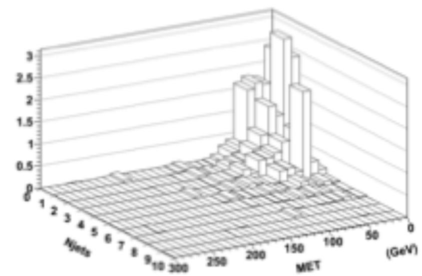
(f) Herwig $WW \rightarrow e\mu$



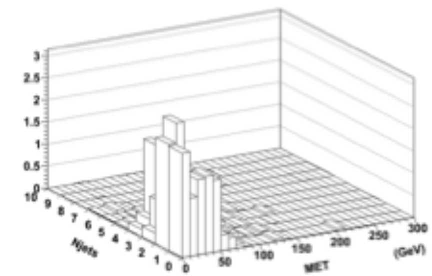
(c) MC@NLO $t\bar{t} \rightarrow e\mu$



(d) MC@NLO $t\bar{t} \rightarrow e\mu$

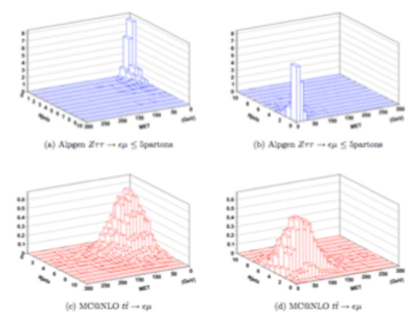


(g) $e\mu$ from combined $Wl\nu + jets, WZ, ZZ, Zee, Z\mu\mu$, single top $Wt, W\mu\gamma$ and $QCD_{b\bar{b}}$



(h) $e\mu$ from combined $Wl\nu + jets, WZ, ZZ, Zee, Z\mu\mu$, single top $Wt, W\mu\gamma$ and $QCD_{b\bar{b}}$

Log likelihood maximization



- The Likelihood function

$$L = \prod_i f(x_i^{pred}(\theta), \sigma_i)$$

- Probability in each bin

$$L = \prod_i \rho_i = \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}$$

$$\mu_i = \alpha N_{t\bar{t}_i} + \beta N_{WW_i} + \gamma N_{Z\rightarrow\tau\tau_i} + n_{other_i}$$

- Systematic errors as Gaussian smearing

$$L(\mu_{t\bar{t}}, \mu_{WW}, \mu_{\tau\tau}) = \prod_i^{N_{bin}} \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \times \prod_j G(x_j, \sigma_j)$$

centre for particle physics Discovery

Plots from my thesis – fit against 10pb^{-1} pseudodata.

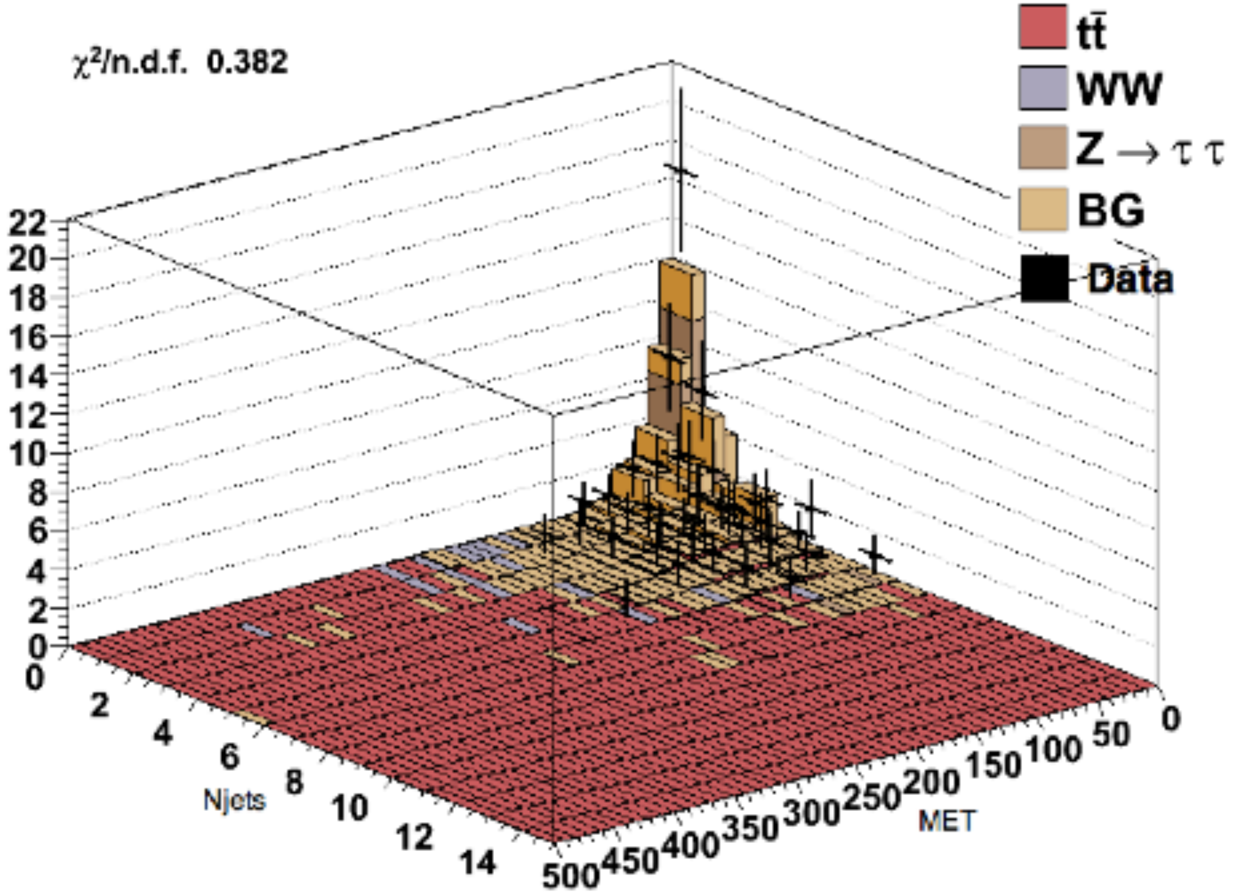


Figure 6.1: Data points from control sample with statistical errors and fit values for all signals floating.

Fit of 20 pb^{-1} real data

Process	$\sigma_i = \frac{\alpha_i N_i}{\Gamma_i A_i \mathcal{L}_i}$ (pb)	stat. errors	syst. error	expected events	events in fit of data
$t\bar{t}$	195	$+58-46$	$+16-10$	46	43
WW	33	$+36-24$	$+6-3$	12	6
$Z\rightarrow\tau\tau$	843	$+254-199$	$+73-48$	63	62

Table 6.6: Fitted cross sections for all processes floating after medium, isolated selection in 20 pb^{-1} real data

From CERN-PH-EP-2011-103: (35pb^{-1})

Channel	$\sigma_{t\bar{t}}$ (pb)	b-tag $\sigma_{t\bar{t}}$ (pb)
ee	$195\,^{+77}_{-65}\,^{+34}_{-30}$	$185\,^{+73}_{-62}\,^{+39}_{-30}$
$\mu\mu$	$186\,^{+54}_{-48}\,^{+19}_{-16}$	$193\,^{+55}_{-49}\,^{+30}_{-23}$
$e\mu$	$166\,^{+28}_{-25}\,^{+13}_{-17}$	$187\,^{+34}_{-31}\,^{+19}_{-14}$
eTL	$170\,^{+92}_{-81}\,^{+65}_{-58}$	–
μTL	$106\,^{+46}_{-40}\,^{+35}_{-31}$	–
Combined	$171 \pm 20 \pm 14$	$188 \pm 26\,^{+20}_{-16}$

Table 3: Measured cross sections in each dilepton channel, and the combination of the untagged and tagged channels with their statistical and systematic uncertainties. The luminosity uncertainty is not included here.

Usage of method

- The method has been used as an alternative test in measurement of the ttbar cross section measurement in the dilepton channel.
- MET uncertainties problematic for WW shape – otherwise could have been an interesting method for measuring WW cross section with low statistics.
- The method is being extended to eg. trilepton channels and to a variety of parameters.
- If discrepancies can be found using a general approach many parameters spaces can be probed.
- Caution should be taken with eg. look elsewhere effect.

Bumphunters

- Instead of overall shape deviances, statistical test can be performed to find actual peaks or excess of events in distribution tails – given the above framework this is easily added and could obviously yield interesting results.

- A range of “bumphunting” tools have already been developed and tested.

- Discovery potential has been shown to match those of the specific searches in CDF and D0.
  - This is of course an important point - it might not be the case for all LHC searches

The different codes

- VISTA

  Probes the bulk of kinematic distributions of high-Pt data. It searches for significant deviations from the null-hypothesis (SM) in regions confined by sidebands.
- Sleuth

  Examine the tails of SumPt distributions of high-Pt final states.
- BumpHunter

  A package that includes a series of tests one of which is entitled BumpHunter. BumpHunter also searches for discrepancies in regions confined by sidebands.
- TailHunter

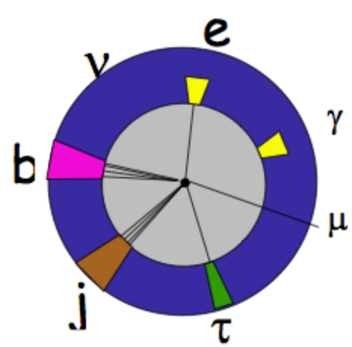
  A tool under the BumpHunter package similar to Sleuth in that it looks for excess events in tails of high-Pt final states. It has no sideband regions
- Bard

  More theoretically minded program, that looks for possible theories (Lagrangians) that fit with a found peak.

Final states in VISTA

Select High-Pt Objects

- Data was collected by online triggers:
  e, central, Et>18
  μ, central, Pt> 18
  γ, cen or plug, Et>25
  jet, Et>20 (prescaled), Et>100
  central e, Et>4, central μ, Pt>4
  central e or μ, Et,Pt>4, plug e, Pt>8
  γ γ, cen or plug, Et>18
  τ τ, central, Pt>10



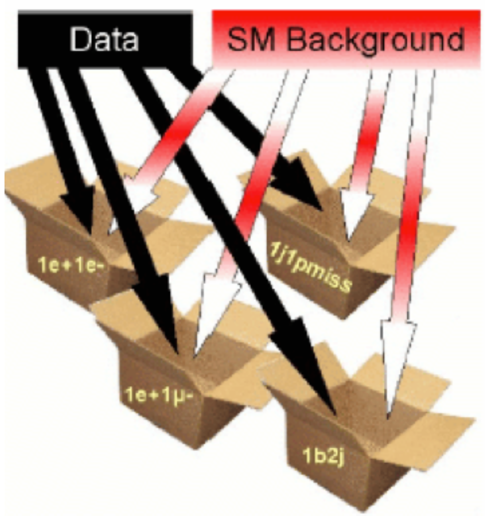
- Offline, reduce the sample size, require :
  electron>25
  OR photon>60 GeV
  OR two leptons Et>17 GeV
  OR jet > 200 GeV, etc
  OR one of 10 di-object selections, some prescaled

Data sorted by exclusive final state

- each identified object is exclusive
- each njet is exclusive
- require 10 events to create a new box

For all final states, histogram:

- Pt, η , ϕ of objects
- ΔR, Δϕ of pairs
- mass or m_t of subsets of particles
- other specialized variables



General idea of approach to finding peaks in bulk of distribution - here from *BumpHunter*.

- The test statistic is obtained by defining a central window, W_c (vary between 68% and 95% signal) and sidebands of non-discrepant data.
- Number of events in data and expected events from the tested hypothesis (SM) are then counted in window and in sidebands.
- test statistic is defined as:

$$t = \begin{cases} 0 & if\, d_c \leq b_c or \mathcal{P}(d_L, b_L) \leq 10^{-3} or \mathcal{P}(d_R, b_R) \leq 10^{-3}, \\ f(d_c - b_c) & otherwise. \end{cases}$$

- basically it says test is 0 if data < bg in central window or sidebands have excess in data.
- otherwise the test is $f(d_c-b_c)$ where f can be any positive, monotonically increasing function, such as $(d_c-b_c)^2$ or $(d_c-b_c)^{100}$
- The p-value can be calculated from the test statistic, t.

Bumphunters

- Shift central window and repeat calc. of p-value
- after all possible regions (given defined step-size) are tested the BumpHunter test statistic is calculated:

$$t = -log(p - value^{min})$$

- where $p - value^{min}$ is the smallest of all found p-values.

- Tailhunters
  - (Sleuth, TailHunter ...)
  - without sidebands - last bin to right in window is last bin with data in distribution.
  - i.e. only requirement is excess in data.
  - Very simple yet force lies in ability to test many parameter combinations.

Backup

Sleuth

Lesser generation equivalence

$e^+ \equiv \mu^+$

Quark jets come in pairs

$2j \equiv 3j$

$4j \equiv 5j$

$bj \equiv bb \equiv bjj \equiv bbj$

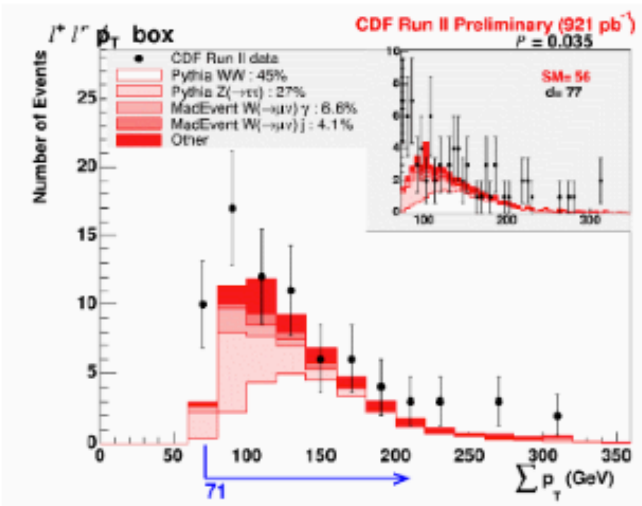
- Search for excess of events in defined region
- At least 3 data events
- from cut to infinity

$$p_d = \sum_{i=d}^{\infty} \frac{b^i}{i!} e^{-b}$$

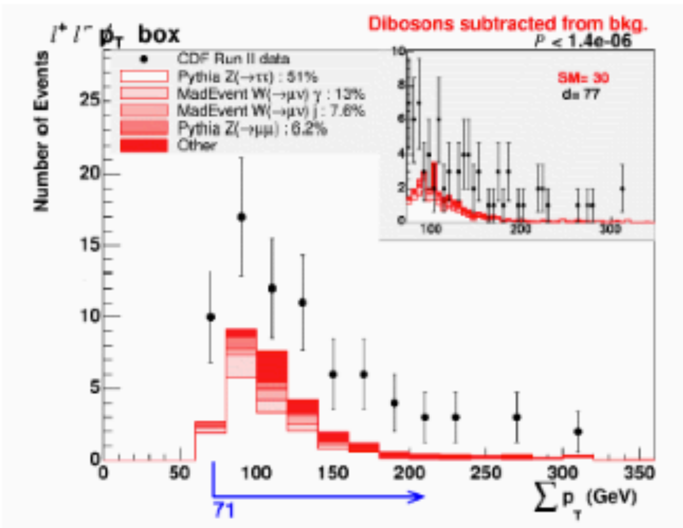
- for SM exp. b to fluctuate up to or above #events in region d.
- Pseudo-experiments are generated for most interesting regions in each final state - $\mathcal{P}$ is then the fraction of PE in this final state more interesting than most interesting region in data.
- $\mathcal{P}$ small = new physics!
- Further calculated for all final states finds smallest $\mathcal{P}$ and checks for most probability of this summing over all sleuth final states... !

Does this work?

ll MEt normally



ll MEt after removing WW



- it would discover WW, even after refitting
- reasonably sensitive to high mass Higgs
- not as sensitive to light Higgs as dedicated analysis
- less sensitive to single top than dedicated analysis

Discovery centre for particle physics

BumpHunter

- Remember p-value is defined $P(t \geq t_0|H_0)$. $t \geq t_0$ only happens when $d_c \geq d_{c0}$ for some observed data in the central band d_{c0}, while no excess is found in the sidebands. The p-value can be calculated from these three independent criteria:

$$p - value = \begin{cases} 1, \\ \mathcal{P}(d_{c_0}, b_c)(1 - 10^{-3})^2 \end{cases}$$

for same ranges but with obs. compared to bg

- $\mathcal{P}$ can be calculated as it can be shown, that

$$\mathcal{P}(d,b) \begin{cases} \Gamma(d,b) & \text{if } d \geq b, \\ 1 - \Gamma(d+1,b) & \text{if } d \leq b. \end{cases}$$

where

$$\Gamma(d,b) = \frac{1}{\Gamma(d)} \int_0^b t^{d-1} e^{-1} dt$$ and $$\Gamma(d) = \int_0^\infty t^{d-1} e^{-1} dt$$

Discovery centre for particle physics

Where to get it and find out more:

- VISTA/SLEUTH: dunno... but try
  - http://www-cdf.fnal.gov/physics/exotic/r2a/20070426.vista_sleuth/publicPage.html
- BumpHunter package:
  - https://svnweb.cern.ch/trac/atlasgrp/browser/Institutes/UChicago/DijetMassAndChi/trunk/macros/package
- Bard
  - http://arxiv.org/abs/hep-ph/0602101
- Marmoset:
  - http://marmoset.jthaler.net/wiki/doku.php?id=installing