Measurement of Lepton-Jet correlation in DIS with H1 at HERA, using machine learning for unfolding

Miguel Arratia
DIS Born-level configuration

\[ \gamma^* q \rightarrow q \]
A new channel to probe for quark transverse-momentum distributions (TMDs) and evolution

Liu et al. PRL. 122, 192003, Gutierrez et al. PRL. 121, 162001

“The advantage of the lepton-jet correlation as compared to the standard SIDIS processes is that it does not involve TMD fragmentation functions.”
Existing TMD data

Bacchetta et al. 1703.10157
Constraining TMD evolution with HERA data

Bridging DIS from fixed-target exp. and high Q2 Drell-Yan at colliders.
Fixing open issues of TMD factorization & universality
The H1 experiment at HERA

- Tracking system
  (silicon tracker, jet chambers, proportional chambers)
- LAr calorimeter (em/had)
- Scintillating fiber calorimeter

Both combined using an energy flow algorithm

1% Jet energy scale
0.5-1% lepton energy scale
Unfolding with Omnifold (via machine-learning).
Andreassen et al. PRL 124, 182001 (2020)
Closure tests (Pseudo Data: Django, Response: Rapgap)

![Graphs showing closure tests for H1 with pseudo data representing Django and Rapgap responses.](image)

- For $Q^2 > 150$ GeV$^2$, $0.2 < y < 0.7$, $p_T^{jet} > 10$ GeV, and $k_T, R = 1.0$, the graphs demonstrate the comparison between Django truth, Django reconstruction, and the unfolded with Rapgap responses, highlighting the relative differences in $1/\sigma_{jet}$ and $d\sigma/dp_T^{jet}$.
Systematic uncertainties

**H1**

$Q^2 > 150 \text{ GeV}^2$, $0.2 < y < 0.7$, $p_T^{\text{jet}} > 10 \text{ GeV}$

- QED rad. corr.
- HFS scale (in jet)
- HFS scale (remainder)
- HFS $\phi$ angle
- Lepton energy scale
- Lepton $\phi$ angle
- Model
- Stat.
- Total Syst.
Systematic uncertainties

**H1** $Q^2 > 150$ GeV$^2$, $0.2 < y < 0.7$, $p_T^{\text{jet}} > 10$ GeV

<table>
<thead>
<tr>
<th>Description</th>
<th>Systematic Uncertainty [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>QED rad. corr.</td>
<td>Blue</td>
</tr>
<tr>
<td>HFS scale (in jet)</td>
<td>Green</td>
</tr>
<tr>
<td>HFS scale (remainder)</td>
<td>Red</td>
</tr>
<tr>
<td>HFS $\phi$ angle</td>
<td>Purple</td>
</tr>
<tr>
<td>Lepton energy scale</td>
<td>Red</td>
</tr>
<tr>
<td>Lepton $\phi$ angle</td>
<td>Pink</td>
</tr>
<tr>
<td>Model</td>
<td>Gray</td>
</tr>
<tr>
<td>Stat.</td>
<td>Yellow</td>
</tr>
<tr>
<td>Total Syst.</td>
<td>Black (dotted)</td>
</tr>
</tbody>
</table>

**Graph 1: Systematic Uncertainty vs. $q_T^{\text{jet}}/Q$**

**Graph 2: Systematic Uncertainty vs. $\Delta\phi^{\text{jet}}$ [rad]**
Measurement of Lepton-Jet Correlation in Deep-Inelastic Scattering with the H1 Detector Using Machine Learning for Unfolding

V. Andreev et al. (H1 Collaboration)
Phys. Rev. Lett. 128, 132002 – Published 31 March 2022
Jet transverse momentum

Well described by NNLO calculation, and some MCs like Herwig and Djangoh
Jet pseudorapidity

Not well described at large pseudorapidity by NNLO, missing higher-order terms.

Well described by Rapgap
TMD calculation does a great job at low $q_T$; collinear calculation does a great job at large $q_T$.

Large overlap between collinear and TMD frameworks
Textbook example of “matching” between collinear and TMD frameworks

First time seen in DIS!

(not seen in fixed-target DIS)
TMD calculation, without free parameters, describes data over wide kinematic range

\[
\frac{d^5 \sigma(p \rightarrow p')}{dy \, d^2 k_\perp \, d^2 q_\perp} = \sigma_0 \int d^2 k_\perp d^2 \lambda_\perp \, x f_q(x, k_\perp, \zeta_c, \mu_F) \times H_{\text{TMD}}(Q, \mu_F) S_J(\lambda_\perp, \mu_F) \times \delta^{(2)}(q_\perp - k_\perp - \lambda_\perp).
\]

- TMD calculations by F. Yuan and Z. Kang, TMD PDFs and soft factors extracted from low Q2 DIS and DY data. Sun et al. arXiv:1406.3073
TMD calculation does a great job at low qT; collinear calculation does a great job at large qT. Large overlap between collinear and TMD frameworks

Lepton-jet azimuthal correlations

H1
$Q^2 > 150 \text{ GeV}^2$
$0.2 < y < 0.7$
$p_T^{\text{jet}} > 10 \text{ GeV}$
$k_T, R = 1.0$

$1/\sigma_\text{jet} \frac{d\sigma}{d\Delta\phi^{\text{jet}}}$

Data
PYTHIA 8.3
HERWIG 7.2
DJANGOH
RAPGAP

CASCADE set 1
CASCADE set 2
NNLO $\Theta$ NP
TMD (LO + NLL')

Model / Data

artificial horizontal marker offsets added for clarity

$\Delta\phi^{\text{jet}} [\text{rad}]$
Azimuthal correlation

Textbook example of “matching” between collinear and TMD frameworks

First time seen in DIS!

(not seen in fixed-target DIS)
Omnifold allowed us to do a simultaneous, unbinned “unfolding”

First-ever measurement that uses machine-learning to correct for detector effects.
**Correlation matrix**

- Simultaneous Unfolding of these observables
- Unbinned (binned here for reference)
Summary

- New measurement of lepton jet momentum and azimuthal imbalance in DIS, which provide a new way to constrain TMD PDFs and their evolution

- Pure TMD calculation does a great job at low $q_T$; Pure collinear calculation does a great job at large $q_T$. Large overlap. Data can constrain matching between TMD and collinear frameworks

- First-ever measurement that uses machine-learning to correct for detector effects. (using Omnifold method)

- This is the first measurement in a series of studies that aim at creating a pathfinder program for the future EIC
backup
Reweighting the reco-level distributions

We use simple fully connected networks with a few hidden layers.

The distribution is binned for illustration, but the reweighting is unbinned.
All these distributions are simultaneously reweighted.
Jet performance (energy flow reconstruction)
Closure tests (Pseudo Data: Django, Response: Rapgap)
Hadronization corrections (applied to NNLO calculation)

Small, and consistent with Pythia8 and Herwig despite different models of hadronization
Response matrices
(not actually used as our results are unbinned, but just for reference)
Response matrices
(not actually used as our results are unbinned, but just for reference)
Lepton-jet imbalance $q_T = |\vec{k}_{l\perp} + \vec{p}_j\perp|$

In Born-level configuration
Probes quark TMD PDFs

Liu et al. PRL. 122, 192003 (2019)

\[
\frac{d^5 \sigma(\ell p \to \ell' J)}{dy_\ell d^2 k_{\ell\perp} d^2 q_\perp} = \sigma_0 \int d^2 k_\perp d^2 \lambda_\perp x f_q(x, k_\perp, \zeta_c, \mu_F) \\
\times H_{\text{TMD}}(Q, \mu_F) S_J(\lambda_\perp, \mu_F) \\
\times \delta^{(2)}(q_\perp - k_\perp - \lambda_\perp).
\]
Evolution: Endgame