**Online Supplement for “Logistic Regression for Crystal Growth Process Modeling through
Hierarchical Nonnegative Garrote based Variable Selection”**

Hongyue Sun¹, Xinwei Deng², Kaibo Wang³, and Ran Jin¹

¹Grado Department of Industrial and Systems Engineering, Virginia Tech., Blacksburg,
VA 24061, USA

²Department of Statistics, Virginia Tech., Blacksburg, VA 24061, USA

³Department of Industrial Engineering, Tsinghua University, Beijing 100084, China

Supplemental Material A

The detailed simulation results are summarized in Tables A1-A3. The values in the tables are calculated as the averages and the standard errors (in the parenthesis) based on 50 simulation replicates under each simulation scenarios. In the tables, n_{tr} shows training sample size for the models. The validation sample size is the same as training sample size and testing sample size twice of training sample size. The training, validation and testing samples are generated from the same model. Testing error is the error for the testing data. Overall variable selection error is calculated as the percentage of total incorrectly selected variables in the final estimated model among all predictors. Type I variable selection error is the percentage of significant variables not being selected in the final estimated model, and Type II variable selection error is the percentage of insignificant variables being selected in the final estimated model. The result is highlighted in bold if its performance is the best among all methods under that simulation setting. If multiple results are comparable (fall into 1 standard deviation of the minimum error), then all these comparable results are regarded as the best.

Table A1. Testing Errors for 50 Replications

Sample Size	Methods	$\rho=0, \tau=0$		$\rho=0, \tau=0.3$		$\rho=0.6, \tau=0$		$\rho=0.6, \tau=0.3$	
		S=0.1	S=0.4	S=0.1	S=0.4	S=0.1	S=0.4	S=0.1	S=0.4
$n_{tr}=20$	LR	0.475 (0.016)	0.446 (0.014)	0.470 (0.013)	0.449 (0.015)	0.484 (0.014)	0.452 (0.018)	0.472 (0.017)	0.424 (0.018)
	Lasso	0.418 (0.018)	0.418 (0.018)	0.388 (0.020)	0.342 (0.017)	0.429 (0.018)	0.259 (0.014)	0.402 (0.018)	0.240 (0.015)

	Ridge	0.438 (0.014)	0.370 (0.017)	0.404 (0.019)	0.292 (0.014)	0.439 (0.017)	0.211 (0.014)	0.396 (0.015)	0.202 (0.013)
	NNG	0.423 (0.018)	0.399 (0.017)	0.353 (0.016)	0.324 (0.015)	0.399 (0.015)	0.264 (0.018)	0.384 (0.016)	0.215 (0.013)
	GrpLasso	0.417 (0.015)	0.324 (0.017)	0.350 (0.016)	0.266 (0.014)	0.377 (0.016)	0.192 (0.012)	0.345 (0.015)	0.172 (0.012)
	HLasso	0.408 (0.017)	0.375 (0.020)	0.403 (0.020)	0.321 (0.016)	0.403 (0.015)	0.231 (0.012)	0.448 (0.017)	0.246 (0.016)
	HNNG	0.400 (0.019)	0.369 (0.018)	0.350 (0.015)	0.289 (0.012)	0.378 (0.016)	0.212 (0.013)	0.366 (0.016)	0.196 (0.012)
$n_{tr}=100$	LR	0.305 (0.003)	0.219 (0.003)	0.300 (0.003)	0.199 (0.002)	0.309 (0.002)	0.176 (0.003)	0.309 (0.003)	0.177 (0.003)
	Lasso	0.278 (0.003)	0.202 (0.002)	0.278 (0.003)	0.173 (0.002)	0.274 (0.003)	0.153 (0.002)	0.275 (0.002)	0.135 (0.002)
	Ridge	0.311 (0.003)	0.210 (0.002)	0.302 (0.002)	0.173 (0.002)	0.297 (0.003)	0.154 (0.003)	0.297 (0.002)	0.135 (0.002)
	NNG	0.275 (0.003)	0.219 (0.003)	0.276 (0.003)	0.181 (0.002)	0.277 (0.003)	0.155 (0.002)	0.280 (0.003)	0.140 (0.002)
	GrpLasso	0.293 (0.003)	0.192 (0.002)	0.287 (0.003)	0.160 (0.002)	0.285 (0.002)	0.141 (0.002)	0.282 (0.002)	0.121 (0.002)
	HLasso	0.353 (0.005)	0.192 (0.003)	0.283 (0.003)	0.163 (0.003)	0.286 (0.003)	0.145 (0.002)	0.286 (0.003)	0.126 (0.002)
	HNNG	0.266 (0.003)	0.200 (0.002)	0.265 (0.003)	0.165 (0.002)	0.263 (0.002)	0.137 (0.002)	0.274 (0.003)	0.125 (0.002)
$n_{tr}=200$	LR	0.286 (0.001)	0.181 (0.001)	0.285 (0.001)	0.163 (0.001)	0.285 (0.002)	0.138 (0.001)	0.283 (0.001)	0.133 (0.001)
	Lasso	0.265 (0.001)	0.176 (0.001)	0.270 (0.001)	0.154 (0.001)	0.270 (0.002)	0.130 (0.001)	0.263 (0.001)	0.122 (0.001)
	Ridge	0.286 (0.001)	0.181 (0.001)	0.285 (0.001)	0.158 (0.001)	0.285 (0.002)	0.130 (0.001)	0.278 (0.001)	0.123 (0.001)
	NNG	0.262 (0.001)	0.181 (0.001)	0.264 (0.001)	0.158 (0.001)	0.266 (0.001)	0.131 (0.001)	0.262 (0.001)	0.123 (0.001)
	GrpLasso	0.275 (0.001)	0.176 (0.001)	0.276 (0.001)	0.151 (0.001)	0.275 (0.002)	0.125 (0.001)	0.271 (0.001)	0.117 (0.001)
	HLasso	0.264 (0.001)	0.172 (0.001)	0.267 (0.002)	0.150 (0.001)	0.275 (0.002)	0.133 (0.001)	0.285 (0.003)	0.121 (0.001)
	HNNG	0.257 (0.001)	0.169 (0.001)	0.260 (0.001)	0.149 (0.001)	0.264 (0.001)	0.122 (0.001)	0.263 (0.001)	0.117 (0.001)

Table A2. Type I and Type II Variable Selection Errors for 50 Replications

Sample Size	Methods	$\rho=0, \tau=0$				$\rho=0, \tau=0.3$			
		S=0.1		S=0.4		S=0.1		S=0.4	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
$n_{tr}=20$	LR	0.690 (0.333)	0.509 (0.125)	0.605 (0.160)	0.490 (0.151)	0.620 (0.372)	0.500 (0.091)	0.612 (0.184)	0.505 (0.118)
	Lasso	0.460 (0.376)	0.180 (0.142)	0.627 (0.218)	0.180 (0.151)	0.360 (0.365)	0.161 (0.120)	0.587 (0.230)	0.222 (0.146)
	Ridge	0.090 (0.261)	0.821 (0.305)	0.050 (0.134)	0.930 (0.183)	0.090 (0.241)	0.801 (0.326)	0.015 (0.041)	0.970 (0.047)
	NNG	0.570 (0.391)	0.136 (0.186)	0.650 (0.297)	0.198 (0.297)	0.440 (0.345)	0.127 (0.197)	0.563 (0.336)	0.298 (0.351)
	GrpLasso	0.370 (0.438)	0.558 (0.393)	0.140 (0.248)	0.610 (0.337)	0.230 (0.307)	0.603 (0.298)	0.120 (0.238)	0.730 (0.287)
	HLasso	0.520 (0.349)	0.363 (0.141)	0.420 (0.234)	0.163 (0.144)	0.640 (0.268)	0.177 (0.115)	0.540 (0.222)	0.243 (0.205)
	HNNG	0.510 (0.357)	0.146 (0.199)	0.557 (0.324)	0.273 (0.289)	0.420 (0.341)	0.127 (0.181)	0.415 (0.276)	0.365 (0.272)
		$\rho=0.6, \tau=0$				$\rho=0.6, \tau=0.3$			
		S=0.1		S=0.4		S=0.1		S=0.4	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
	LR	0.610 (0.339)	0.520 (0.088)	0.545 (0.131)	0.527 (0.107)	0.640 (0.336)	0.500 (0.095)	0.550 (0.158)	0.495 (0.111)
	Lasso	0.450 (0.381)	0.179 (0.143)	0.487 (0.191)	0.225 (0.113)	0.470 (0.397)	0.182 (0.126)	0.455 (0.169)	0.222 (0.132)
	Ridge	0.090 (0.241)	0.841 (0.267)	0.005 (0.025)	0.980 (0.043)	0.060 (0.218)	0.896 (0.197)	0.005 (0.025)	0.983 (0.037)
	NNG	0.610 (0.368)	0.113 (0.157)	0.465 (0.313)	0.343 (0.353)	0.560 (0.424)	0.124 (0.147)	0.320 (0.259)	0.425 (0.366)
	GrpLasso	0.300 (0.378)	0.572 (0.336)	0.020 (0.099)	0.755 (0.283)	0.220 (0.306)	0.591 (0.307)	0.020 (0.099)	0.672 (0.301)
	HLasso	0.380 (0.312)	0.448 (0.151)	0.130 (0.222)	0.203 (0.135)	0.800 (0.247)	0.106 (0.116)	0.270 (0.252)	0.138 (0.085)
	HNNG	0.510 (0.385)	0.138 (0.202)	0.325 (0.276)	0.333 (0.247)	0.540 (0.389)	0.147 (0.159)	0.277 (0.231)	0.360 (0.253)
$n_{tr}=100$		$\rho=0, \tau=0$				$\rho=0, \tau=0.3$			
		S=0.1		S=0.4		S=0.1		S=0.4	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
	LR	1.000 (0.000)	0.499 (0.110)	0.995 (0.025)	0.488 (0.115)	1.000 (0.000)	0.487 (0.072)	0.980 (0.053)	0.498 (0.123)
	Lasso	0.010 (0.071)	0.344 (0.164)	0.012 (0.046)	0.552 (0.194)	0.020 (0.099)	0.358 (0.178)	0.022 (0.055)	0.530 (0.194)
	Ridge	0.000 (0.000)	0.973 (0.041)	0.000 (0.000)	0.988 (0.029)	0.000 (0.000)	0.988 (0.028)	0.000 (0.000)	0.997 (0.017)
	NNG	0.040 (0.137)	0.090 (0.085)	0.085 (0.142)	0.348 (0.316)	0.060 (0.164)	0.124 (0.148)	0.080 (0.131)	0.350 (0.302)

	GrpLasso	0.000 (0.000)	0.872 (0.170)	0.000 (0.000)	0.933 (0.176)	0.000 (0.000)	0.933 (0.120)	0.000 (0.000)	0.892 (0.203)
	HLasso	0.540 (0.222)	0.204 (0.099)	0.020 (0.099)	0.172 (0.062)	0.050 (0.152)	0.550 (0.154)	0.010 (0.071)	0.207 (0.122)
	HNNG	0.020 (0.099)	0.089 (0.104)	0.052 (0.076)	0.303 (0.217)	0.030 (0.120)	0.136 (0.155)	0.052 (0.095)	0.278 (0.211)
		$\rho=0.6, \tau=0$				$\rho=0.6, \tau=0.3$			
		S=0.1		S=0.4		S=0.1		S=0.4	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
	LR	1.000 (0.000)	0.503 (0.084)	0.942 (0.077)	0.502 (0.114)	0.990 (0.071)	0.493 (0.068)	0.907 (0.094)	0.483 (0.097)
	Lasso	0.000 (0.000)	0.269 (0.163)	0.035 (0.057)	0.465 (0.146)	0.000 (0.000)	0.344 (0.220)	0.070 (0.076)	0.433 (0.149)
	Ridge	0.000 (0.000)	0.970 (0.056)	0.000 (0.000)	0.990 (0.027)	0.000 (0.000)	0.970 (0.045)	0.000 (0.000)	0.990 (0.027)
	NNG	0.050 (0.152)	0.129 (0.178)	0.063 (0.095)	0.495 (0.313)	0.020 (0.099)	0.192 (0.204)	0.058 (0.092)	0.535 (0.316)
	GrpLasso	0.000 (0.000)	0.856 (0.188)	0.000 (0.000)	0.908 (0.174)	0.000 (0.000)	0.911 (0.153)	0.000 (0.000)	0.875 (0.227)
	HLasso	0.000 (0.000)	0.656 (0.214)	0.000 (0.000)	0.308 (0.216)	0.200 (0.247)	0.356 (0.110)	0.000 (0.000)	0.175 (0.059)
	HNNG	0.020 (0.099)	0.086 (0.105)	0.037 (0.073)	0.242 (0.155)	0.010 (0.071)	0.147 (0.183)	0.027 (0.058)	0.305 (0.178)
$n_{tr}=200$		$\rho=0, \tau=0$				$\rho=0, \tau=0.3$			
		S=0.1		S=0.4		S=0.1		S=0.4	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
	LR	1.000 (0.000)	0.507 (0.104)	1.000 (0.000)	0.493 (0.144)	1.000 (0.000)	0.493 (0.087)	1.000 (0.000)	0.513 (0.137)
	Lasso	0.000 (0.000)	0.371 (0.186)	0.000 (0.000)	0.627 (0.203)	0.000 (0.000)	0.328 (0.215)	0.000 (0.000)	0.595 (0.187)
	Ridge	0.000 (0.000)	0.986 (0.025)	0.000 (0.000)	0.997 (0.017)	0.000 (0.000)	0.979 (0.032)	0.000 (0.000)	0.993 (0.023)
	NNG	0.000 (0.000)	0.140 (0.224)	0.003 (0.018)	0.505 (0.384)	0.000 (0.000)	0.103 (0.109)	0.005 (0.025)	0.533 (0.331)
	GrpLasso	0.000 (0.000)	0.944 (0.138)	0.000 (0.000)	0.933 (0.154)	0.000 (0.000)	0.917 (0.140)	0.000 (0.000)	0.950 (0.161)
	HLasso	0.000 (0.000)	0.478 (0.091)	0.000 (0.000)	0.425 (0.237)	0.000 (0.000)	0.456 (0.055)	0.000 (0.000)	0.233 (0.176)
	HNNG	0.000 (0.000)	0.082 (0.076)	0.012 (0.038)	0.145 (0.132)	0.000 (0.000)	0.090 (0.137)	0.002 (0.018)	0.285 (0.208)
		$\rho=0.6, \tau=0$				$\rho=0.6, \tau=0.3$			
		S=0.1		S=0.4		S=0.1		S=0.4	
		Type I	Type II	Type I	Type II	Type I	Type II	Type I	Type II
	LR	1.000 (0.000)	0.509 (0.092)	0.997 (0.018)	0.498 (0.090)	1.000 (0.000)	0.513 (0.086)	0.975 (0.057)	0.520 (0.106)
	Lasso	0.000 (0.000)	0.377 (0.174)	0.002 (0.018)	0.530 (0.203)	0.000 (0.000)	0.349 (0.203)	0.015 (0.041)	0.442 (0.177)
	Ridge	0.000 (0.000)	0.990 (0.000)	0.000 (0.000)	0.992 (0.000)	0.000 (0.000)	0.980 (0.000)	0.000 (0.000)	1.000 (0.000)

		(0.000)	(0.024)	(0.000)	(0.030)	(0.000)	(0.046)	(0.000)	(0.000)
	NNG	0.010 (0.071)	0.131 (0.186)	0.010 (0.043)	0.467 (0.296)	0.000 (0.000)	0.170 (0.246)	0.018 (0.044)	0.480 (0.301)
	GrpLasso	0.000 (0.000)	0.956 (0.117)	0.000 (0.000)	0.933 (0.176)	0.000 (0.000)	0.933 (0.154)	0.000 (0.000)	0.883 (0.207)
	HLasso	0.000 (0.000)	0.706 (0.181)	0.000 (0.000)	0.813 (0.229)	0.200 (0.267)	0.356 (0.119)	0.000 (0.000)	0.167 (0.000)
	HNNG	0.000 (0.000)	0.091 (0.079)	0.007 (0.030)	0.207 (0.150)	0.000 (0.000)	0.164 (0.165)	0.005 (0.025)	0.285 (0.157)

Table A3. Overall Variable Selection Errors for 50 Replications

Sample Size	Methods	$\rho=0, \tau=0$		$\rho=0, \tau=0.3$		$\rho=0.6, \tau=0$		$\rho=0.6, \tau=0.3$	
		S=0.1	S=0.4	S=0.1	S=0.4	S=0.1	S=0.4	S=0.1	S=0.4
$n_{tr}=20$	LR	0.527 (0.114)	0.536 (0.118)	0.512 (0.098)	0.548 (0.121)	0.529 (0.090)	0.534 (0.084)	0.514 (0.101)	0.517 (0.103)
	Lasso	0.208 (0.118)	0.359 (0.076)	0.182 (0.107)	0.368 (0.091)	0.206 (0.120)	0.330 (0.069)	0.211 (0.114)	0.315 (0.095)
	Ridge	0.748 (0.254)	0.578 (0.066)	0.730 (0.275)	0.588 (0.034)	0.766 (0.221)	0.590 (0.027)	0.812 (0.159)	0.592 (0.026)
	NNG	0.179 (0.149)	0.379 (0.089)	0.158 (0.166)	0.404 (0.105)	0.163 (0.129)	0.392 (0.129)	0.168 (0.120)	0.383 (0.149)
	GrpLasso	0.539 (0.315)	0.422 (0.155)	0.566 (0.247)	0.486 (0.174)	0.545 (0.272)	0.461 (0.167)	0.554 (0.255)	0.411 (0.169)
	HLasso	0.379 (0.117)	0.266 (0.136)	0.223 (0.082)	0.362 (0.179)	0.441 (0.126)	0.174 (0.130)	0.175 (0.085)	0.191 (0.098)
	HNNG	0.182 (0.168)	0.387 (0.098)	0.156 (0.156)	0.385 (0.111)	0.175 (0.164)	0.330 (0.107)	0.186 (0.139)	0.327 (0.121)
$n_{tr}=100$	LR	0.549 (0.099)	0.691 (0.069)	0.538 (0.065)	0.691 (0.081)	0.553 (0.076)	0.678 (0.074)	0.543 (0.061)	0.653 (0.078)
	Lasso	0.311 (0.147)	0.336 (0.112)	0.324 (0.158)	0.327 (0.118)	0.242 (0.147)	0.293 (0.091)	0.310 (0.198)	0.288 (0.087)
	Ridge	0.876 (0.037)	0.593 (0.018)	0.889 (0.025)	0.598 (0.010)	0.873 (0.058)	0.594 (0.016)	0.873 (0.041)	0.594 (0.016)
	NNG	0.085 (0.075)	0.2420 (0.171)	0.118 (0.132)	0.242 (0.171)	0.121 (0.159)	0.322 (0.171)	0.175 (0.183)	0.344 (0.181)
	GrpLasso	0.785 (0.153)	0.560 (0.106)	0.840 (0.108)	0.535 (0.122)	0.770 (0.169)	0.545 (0.105)	0.820 (0.138)	0.525 (0.136)
	HLasso	0.238 (0.067)	0.111 (0.046)	0.500 (0.137)	0.128 (0.087)	0.590 (0.193)	0.185 (0.130)	0.340 (0.074)	0.105 (0.035)
	HNNG	0.082 (0.092)	0.203 (0.121)	0.125 (0.138)	0.188 (0.124)	0.079 (0.094)	0.160 (0.091)	0.133 (0.164)	0.194 (0.102)
$n_{tr}=200$	LR	0.556 (0.094)	0.696 (0.086)	0.544 (0.079)	0.708 (0.082)	0.558 (0.083)	0.698 (0.054)	0.562 (0.077)	0.702 (0.073)
	Lasso	0.334 (0.168)	0.376 (0.122)	0.295 (0.194)	0.357 (0.112)	0.339 (0.156)	0.319 (0.122)	0.314 (0.183)	0.271 (0.110)
	Ridge	0.887 (0.022)	0.598 (0.010)	0.881 (0.028)	0.596 (0.014)	0.891 (0.022)	0.595 (0.018)	0.882 (0.041)	0.600 (0.000)
	NNG	0.126 (0.202)	0.304 (0.229)	0.093 (0.098)	0.322 (0.196)	0.119 (0.167)	0.284 (0.175)	0.153 (0.221)	0.295 (0.177)
	GrpLasso	0.850 (0.124)	0.560 (0.093)	0.825 (0.126)	0.570 (0.096)	0.860 (0.106)	0.560 (0.106)	0.840 (0.139)	0.530 (0.124)
	HLasso	0.430 (0.082)	0.255 (0.142)	0.410 (0.050)	0.140 (0.106)	0.635 (0.163)	0.488 (0.137)	0.340 (0.080)	0.100 (0.000)
	HNNG	0.074 (0.069)	0.092 (0.078)	0.081 (0.123)	0.172 (0.124)	0.082 (0.071)	0.127 (0.090)	0.148 (0.148)	0.173 (0.094)

Supplemental Material B

Figure B1 shows the pairwise comparisons of the ROC curves for the benchmark models and the proposed HNNG model, with the point generated at the 0.5 cut-off probability. It is clear that HNNG performs better than benchmark models LR (Figure B1 (a)), GrpLasso (Figure B1 (e)) and HLasso (Figure B1 (f)) in general. If we further look at the point of 0.5 cut-off probability, HNNG will have larger true positive rate than Ridge (Figure B1 (b)), Lasso (Figure B1 (c)) and NNG (Figure B1 (d)). This is consistent with the CV results in Table 2, that is, the Type I error of HNNG is the smallest among all model candidates. However, the HNNG may have larger false positive rate than other methods, and Table 2 also indicates that sometimes the Type II error of the HNNG is not the smallest.

Figure B1. Pairwise Comparison of ROC Curves of Benchmark and Proposed Models (0.5 Cut-Off Probability Points are Shown on the Curves)

Figure B2 shows the ROC curves for all the models. From the figure, the selection of cut-off probability will influence the errors. In this paper, a conventional 0.5 cut-off probability is used. Other cut-off probabilities can be selected. The corresponding AUC values for the models are listed in Table B1. Note that the AUC value is used to characterize the average performance of a classifier. The larger the AUC value is, the better the classification performance is. From the table, the proposed method performs well. In practice, the cut-off probability will not be varied for a given prediction problem. Therefore, the CV Error is a better indicator to evaluate the models than AUC values. Given the 0.5 cut-off probability, the HNNG method has the smallest overall misclassification error in CV, and the HNNG performs better in various simulation scenarios, we conclude that the HNNG is indeed better than the benchmark models.

Figure B2. Combined ROC Curves for all Models

Table B1. AUC Values for Different Models

Methods	LR	Lasso	Ridge	NNG	GrpLasso	HLasso	HNNG
AUC	0.8495	0.9101	0.9255	0.9360	0.8773	0.8622	0.9187

Supplemental Material C

The predictors selected in the models are visualized in Figures C1-C4. Recall that the four groups of variables are heater power, SP value, pull speed and furnace pressure. We use the two detailed levels wavelet coefficients and the coarse level wavelet coefficients during the modeling. The number of wavelet coefficients for two detail levels and coarse level are 11, 9 and 7. In the figures, the vertical axis represents the index of features, and the horizontal axis represents the model used. A feature is marked if it is selected larger or equal to 7 times over 14 CV folds. From these figures, we can conclude that the proposed model selects features from the coarse levels of heater power and SP value, which implies that the changes in thermal field are responsible for polycrystalline defects in the production.

Figure C1. Selected Local Features for Heater Power

Figure C2. Selected Local Features for SP Value

Figure C3. Selected Local Features for Pull Speed

Figure C4. Selected Local Features for Pressure